

Contribuição ao reconhecimento automático de
padrões epileptiformes em sinais de
eletroencefalograma utilizando análise morfológica

Christine Fredel Boos

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA**

Christine Fredel Boos

**CONTRIBUIÇÃO AO RECONHECIMENTO AUTOMÁTICO DE
PADRÕES EPILEPTIFORMES EM SINAIS DE
ELETROENCEFALOGRAMA UTILIZANDO ANÁLISE
MORFOLÓGICA**

Florianópolis

2015

Christine Fredel Boos

**CONTRIBUIÇÃO AO RECONHECIMENTO AUTOMÁTICO DE
PADRÕES EPILEPTIFORMES EM SINAIS DE
ELETROENCEFALOGRAMA UTILIZANDO ANÁLISE
MORFOLÓGICA**

Tese de doutorado submetida ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal de Santa Catarina para a obtenção do título de Doutor em Engenharia Elétrica, área de concentração Processamento de Informação, subárea Engenharia Biomédica.

Orientador: Prof. Fernando Mendes de Azevedo, Dr.Sc.

Florianópolis

2015

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Boos, Christine Fredel

Contribuição ao reconhecimento automático de padrões epileptiformes em sinais de eletroencefalograma utilizando análise morfológica / Christine Fredel Boos ; orientador, Fernando Mendes de Azevedo - Florianópolis, SC, 2015.
149 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

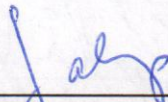
1. Engenharia Elétrica. 2. Eletroencefalograma. 3. Padrões epileptiformes. 4. Processamento digital de sinais. 5. Reconhecimento de padrões. I. Azevedo, Fernando Mendes de. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

Christine Fredel Boos

**CONTRIBUIÇÃO AO RECONHECIMENTO AUTOMÁTICO DE
PADRÕES EPILEPTIFORMES EM SINAIS DE
ELETROENCEFALOGRAMA UTILIZANDO ANÁLISE
MORFOLÓGICA**

Esta Tese de Doutorado foi julgada adequada para obtenção do Título de Doutor em Engenharia Elétrica, área de concentração Processamento de Informação, subárea Engenharia Biomédica, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Florianópolis, 19 de junho de 2015.



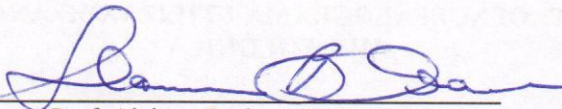
Prof. Carlos Galup Montoro, PhD.
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica.

Banca Examinadora:



Prof. Fernando Mendes de Azevedo, Dr.Sc.
Universidade Federal de Santa Catarina.
Presidente da banca e Orientador

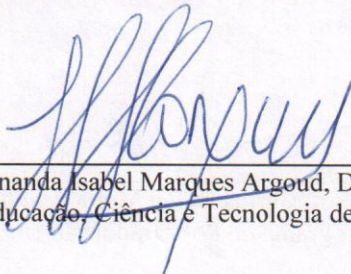
Banca Examinadora (continuação):



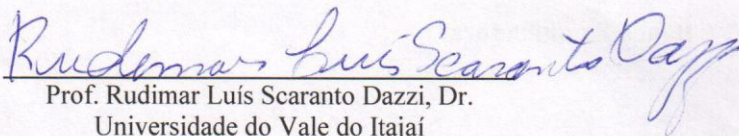
Prof. Alcimar Barbosa Soares, PhD.
Universidade Federal de Uberlândia



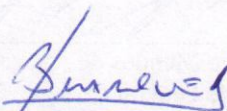
Prof.^a Edna Lúcia Flores, Dra.
Universidade Federal de Uberlândia



Prof.^a Fernanda Isabel Marques Argoud, Dra.
Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina



Prof. Rudimar Luís Scaranto Dazzi, Dr.
Universidade do Vale do Itajaí



Prof. Jefferson Luiz Brum Marques, PhD.
Universidade Federal de Santa Catarina

AGRADECIMENTOS

A todos os meus familiares. Em especial aos meus pais, Luiz Carlos Boos e Griseldes Fredel Boos, e minha irmã, Melina Fredel Boos, pelo amor incondicional, pelo apoio, pelas broncas quando preciso, pelo incentivo, motivação, compreensão, paciência e conselhos dados ao longo da minha vida e especialmente nestes últimos seis anos (dois de mestrado e quatro de doutorado). As longas conversas por telefone e Skype foram indispensáveis para que eu me mantivesse motivada e continuasse trabalhando, principalmente nos momentos em que tudo parecia difícil e a saudade de estar em casa com vocês era muito grande.

Ao meu orientador e amigo, professor Fernando Mendes de Azevedo, pela oportunidade de fazer o doutorado e por todo o apoio, paciência, orientação e conversas que foram fundamentais na realização deste trabalho e que contribuíram muito para o meu crescimento pessoal e profissional.

Aos professores do IEB-UFSC por todos os ensinamentos, oportunidades e trocas de ideias que foram muito importantes para o meu desenvolvimento acadêmico e pessoal.

Aos meus colegas no IEB-UFSC por suas palavras de apoio e incentivo, conversas e risadas. A amizade e a convivência com todos vocês foi muito importante para mim.

Aos neurologistas Dr. Roger Walz, Dra. Katia Lin, Dra. Lucia Sukys Caludino por fornecer a base de dados do Centro de Epilepsia do Estado de Santa Catarina e, em conjunto com a neurologista Aline Vieira Scarlatelli Lima, realizar as marcações dos registros de EEG e responder às minhas perguntas sobre os aspectos clínicos da análise dos registros de EEG.

Ao Programa de Pós-Graduação em Engenharia Elétrica da UFSC, IEB-UFSC e CNPq pela oportunidade e apoio financeiro para o desenvolvimento dos trabalhos realizados.

RESUMO

A análise do sinal de eletroencefalograma (EEG) pode ser uma importante ferramenta de auxílio à comprovação clínica do diagnóstico de epilepsia. Esta análise é uma verificação visual criteriosa de registros de EEG na busca por uma atividade eletrográfica específica denominada descarga ou padrão epileptiformes. Essa atividade, quando encontrada com determinada frequência, pode corroborar o diagnóstico de epilepsia. No entanto, como os registros analisados normalmente são resultado de longos períodos de monitoramento, a análise do EEG pode ser um processo demorado. Diversos estudos propuseram metodologias ou sistemas para automatizar este tipo de análise, mas apesar dos esforços e relativo sucesso ainda não existe um algoritmo ou sistema para este reconhecimento automático que seja amplamente difundido no ambiente clínico, possua um desempenho de acordo com as necessidades dos especialistas na área e, do ponto de vista técnico, não possua restrição quanto aos dados de entrada. Desta forma, com o intuito de contribuir para o estudo do reconhecimento automático de padrões epileptiformes o presente trabalho descreve uma proposta de metodologia baseada em análise morfológica e inspirada no comportamento dos especialistas humanos. A análise morfológica é realizada utilizando um conjunto de descritores morfológicos extraídos de sinais de EEG processados digitalmente. A avaliação da metodologia é realizada utilizando um grupo de classificadores computacionais e diferentes bases de dados de EEG. O desempenho da metodologia é analisado por métricas consolidadas na literatura e por uma comparação pareada com o desempenho de sensibilidade, especificidade e concordância de quatro neurofisiologistas clínicos. Resultados obtidos mostraram que o classificador que apresentou o melhor desempenho geral foi uma rede neural combinada com a Análise de Componentes Principais. A sensibilidade e a especificidade média atingida por essa rede foram, respectivamente 58,6 e 87,2%. A eficiência média obtida pela rede nas bases de dados foi de 89%. E a concordância dessa rede com os quatro especialistas foi de 45,7%. Os resultados obtidos para a especificidade mostraram-se satisfatórios, uma vez que os valores médios obtidos foram compatíveis àqueles dos especialistas. Apesar da sensibilidade e da concordância com os especialistas terem apresentado valores baixos, pode-se considerar que a metodologia proposta apresenta resultados promissores considerando a quantidade e natureza dos testes realizados.

Palavras-chave: eletroencefalograma, padrão epileptiforme, descritores morfológicos, processamento digital de sinais, reconhecimento de padrões.

ABSTRACT

The analysis of the EEG signal (EEG) can be an important support tool to the clinical confirmation of the diagnosis of epilepsy. This analysis is a thorough visual inspection of EEG recordings in the search for a specific electrographic activity called epileptiform pattern or discharge. This activity, when found with a certain frequency, can corroborate the diagnosis of epilepsy. However, as the analyzed recordings are usually the result of long term monitoring, the EEG analysis can be a time consuming process. Several studies have proposed methodologies and/or systems to automate this type of analysis however despite the efforts and relative success, there is still no solution for the automatic recognition that is widely disseminated in the clinical environment, has a performance according to the needs of specialists in the field and, from a technical point of view has no restriction about the input data. Thus, in order to contribute to the study of automatic recognition of epileptiform patterns this work describes a methodology proposal based on morphological analysis and inspired by the behavior of human experts. The morphological analysis is performed using a set of morphological descriptors extracted from digitally processed EEG signals. The evaluation of the methodology is performed using a group of computational classifiers and different EEG databases. The methodology's performance is analyzed by metrics consolidated in the literature and by a pairwise comparison of the sensitivity, specificity and agreement of four clinical neurophysiologists. Obtained results show that the classifier that presented the best overall performance was a neural network combined with the Principal Component Analysis. The average sensitivity and specificity achieved by this network were respectively 58.6 and 87.2%. The network's average efficiency obtained in the databases was 89%. And the agreement of the network with four experts was 45.7%. The results for specificity were satisfactory, since the average values obtained were consistent to those of the experts. Despite the low values of sensitivity and agreement with the experts presented by the classifiers, it can be considered that the proposed methodology shows promising results considering the amount and nature of the tests performed in this work.

Keywords: electroencephalogram, epileptiform pattern, morphological descriptors, digital signal processing, pattern recognition.

LISTA DE ILUSTRAÇÕES

Figura 1 – Sistema 10/20 para o posicionamento de eletrodos.	28
Figura 2 – Diferenças do sinal de EEG quando em (A) montagem bipolar e (B) unipolar ou referencial.	28
Figura 3 – Morfologia e respectiva denominação dos padrões epileptiformes mais encontrados em registros de EEG.	31
Figura 4 – Classificação dos métodos de detecção automática de paroxismos.	32
Figura 5 – Pontos referenciais utilizados para o cálculo dos descritores de Gotman (1980), com base em um exemplo de uma espícula e uma onda aguda.	34
Figura 6 – Modelo geral do neurônio artificial.	35
Figura 7 – Tipos de treinamento de redes neurais.	36
Figura 8 – Tabela de Contingência de um classificador genérico.	43
Figura 9 – Visão geral da metodologia proposta e as simulações realizadas neste trabalho.	49
Figura 10 – Descritores relacionados à amplitude e duração do padrão.	55
Figura 11 – Descritores relacionados ao ângulo de vértice dos picos positivo e negativo de um padrão.	56
Figura 12 – Descritores utilizados para a diferenciação entre padrões epileptiformes e artefatos de alta frequência presentes nos sinais.	57
Figura 13 – Primeira parte do teste da interpolação dos sinais de EEG, na qual são extraídas a amplitude máxima e mínima de todos os segmentos de sinal selecionados, tanto para o sinal original quanto para o interpolado.	58
Figura 14 – Diagrama de Bland-Altman gerado para a análise da concordância entre os valores da amplitude mínima para o sinal original, amostrado a 100 Hz, e para o sinal interpolado, amostrado a 1 kHz.	59
Figura 15 – Diagrama de Bland-Altman gerado para a análise da concordância entre os valores da amplitude máxima para o sinal original, amostrado a 100 Hz, e para o sinal interpolado, amostrado a 1 kHz.	60
Figura 16 – Resposta em frequência dos filtros <i>notch</i> utilizados para a remoção de ruído de 60 Hz e desvios da linha de base.	61
Figura 17 – Comportamento de um padrão epileptiforme (paroxismo) e um padrão não-epileptiforme (piscada) quando submetidos a um filtro passa alta de 5ª ordem com frequência de corte de 5 Hz.	62

Figura 18 – Função degrau (atenuação ideal I), a suavização do degrau (atenuação ideal II) utilizado e as curvas de conformação (curva modelo 1, 2 e 3) representadas pelas funções de 4ª ordem.....	63
Figura 19 – Comportamento de um padrão epileptiforme (paroxismo) e um padrão não-epileptiforme (piscada) quando submetidos à curva de conformação de espectro para a atenuação de 4,5 dB nas frequências até 5 Hz.	64
Figura 20 – Exemplo da aplicação do tratamento das bordas.....	65
Figura 21 – Função Gaussiana (a) e função wavelet (b), 1ª derivada da Função Gaussiana utilizada para o cálculo dos descritores.	66
Figura 22 – Função wavelet, 1ª derivada da função Gaussiana, aplicada a um sinal senoidal de frequências 5 e 50 Hz.....	66
Figura 23 – Análise qualitativa dos valores da matriz de correlação.	75
Figura 24 – Resultados da Análise de Componentes Principais: autovalores para as componentes com e sem rotação, identificando com marcador preenchido a máxima componente a ser selecionada.	77
Figura 25 – Resultados da Análise de Componentes Principais: variância explicada acumulada para as componentes com e sem rotação, identificando com marcador preenchido a máxima componente a ser selecionada.....	78
Figura 26 – Sensibilidade dos classificadores para os registros da Base B.	86
Figura 27 – Valores de especificidade dos classificadores para os registros de EEG da Base B.....	87
Figura 28 – Eficiência dos classificadores para os registros de EEG da Base B.	88
Figura 29 – Desempenho médio dos classificadores nas simulações com os registros da base D utilizando o especialista 1 como referência.	90
Figura 30 – Desempenho médio dos classificadores nas simulações com os registros da base D utilizando o especialista 2 como referência.	90
Quadro 1 – Características e desempenho de metodologias ou sistemas descritos na literatura sobre automatização reconhecimento de padrões epileptiformes.....	45

LISTA DE TABELAS

Tabela 1 – Valores de p_{in} e p_{out} utilizados para obtenção das funções.	71
Tabela 2 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtidos nas simulações com as redes neurais.....	76
Tabela 3 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações das redes neurais utilizando componentes principais como estímulo de entrada das redes.	79
Tabela 4 – Tamanho das funções obtidas com a Análise de Discriminantes, da correspondente redução no tamanho do conjunto de descritores utilizados na classificação e apresentação dos índices de desempenho das funções (variância explicada e significância estatística).	80
Tabela 5 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações utilizando as funções discriminantes.....	81
Tabela 6 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações das redes neurais utilizando as variáveis das funções discriminantes como estímulo de entrada das redes.	82
Tabela 7 – Resultado da comparação do desempenho dos classificadores em duas bases de dados diferentes (Base A e Base B).....	83
Tabela 8 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações com a base de dados (Base C).	89
Tabela 9 – Sensibilidade (probabilidade de marcação de um paroxismo) entre os quatro especialistas e os quatro classificadores.	92
Tabela 10 – Especificidade (probabilidade de marcação de um padrão não-epileptiforme) entre os quatro especialistas e os quatro classificadores.	92
Tabela 11 – Concordância da marcação, dos padrões epileptiformes, entre os quatro especialistas e os quatro classificadores.	93

PRODUÇÃO ACADÊMICA

BOOS, C. F.; SCOLARO, G. R.; AZEVEDO, F. M. Comparison between Artificial Neural Networks and Discriminant Functions for Automatic Detection of Epileptiform Discharges. In: ROA ROMERO, L. M. (Ed.). *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*. IFMBE Proceedings. Heidelberg: Springer International Publishing, 2014. v. 41, p. 770–773.

BOOS, C. F.; SCOLARO, G. R.; PEREIRA, M. C. V.; AZEVEDO, F. M. Analysis of Pre-Processing Methods for Artificial Neural Network Pattern Recognition of EEG Signals. In: LONG, M. (Ed.). *World Congress on Medical Physics and Biomedical Engineering 2012*. IFMBE Proceedings. Heidelberg: Springer Berlin Heidelberg, 2013. v. 39, p. 558–561.

BOOS, C. F.; DE AZEVEDO, F. M. Automatic pattern recognition of epileptiform discharges using morphological descriptors and linear discriminant analysis. In: 33rd IEEE International Scientific Conference Electronics and Nanotechnology, 2013, Kiev, Ucrânia, *Anais...*, New Jersey, 2013, p. 293–296.

BOOS, C. F.; AZEVEDO, F. M. Análise de Componentes Principais como Ferramenta para o Dimensionamento de Redes Neurais Artificiais. In: XXIII Congresso Brasileiro em Engenharia Biomédica, 2012, Porto de Galinhas, *Anais...*, Porto de Galinhas, 2012, p. 1–5.

BOOS, C. F.; AZEVEDO, F. M. DE; SCOLARO, G. R.; PEREIRA, M. DO C. V. Automatic Detection of Paroxysms in EEG Signals Using Morphological Descriptors and Artificial Neural Networks. In: LASKOVSKI, A. (Ed.). *Biomedical Engineering, Trends in Electronics, Communications and Software*. Heidelberg: InTech, 2011. p. 387–402.

SCOLARO, G. R.; AZEVEDO, F. M. DE; BOOS, C. F.; WALZ, R. Wavelet Filter to Attenuate the Background Activity and High Frequencies in EEG Signals Applied in the Automatic Identification of Epileptiform Events. In: ANDRADE, A. O.; PEREIRA, A. A.; NAVES, E. L. M.; SOARES, A. B. (Eds.). *Practical Applications in Biomedical Engineering*. [s.l.]: InTech, 2013. p. 81–102.

SCOLARO, G. R.; BOOS, C. F.; AZEVEDO, F. M. Wavelet Filter Proposal to Attenuate the Background Activity and High Frequencies in EEG Signals. (A.R. Boccaccini, Ed.) In: 10th IASTED International Conference on Biomedical Engineering, 2013, Innsbruck, Áustria, *Anais...*, Calgary, Canadá, 2013, p. 243–249.

SCOLARO, G. R.; AZEVEDO, F. M.; BOOS, C. F. Evaluation of Different Wavelet Functions Applied in the Development of Digital Filters to Attenuate the Background Activity in EEG Signals. In: LONG, M. (Ed.). *World Congress on Medical Physics and Biomedical Engineering 2012*. IFMBE Proceedings. Heidelberg: Springer Berlin Heidelberg, 2013. v. 39, p. 340–343.

SCOLARO, G. R.; BOOS, C. F.; AZEVEDO, F. M. Neural Classifier for Detection and Classification of Spikes and Sharp Waves. In: International Conference on Bio-inspired Systems and Signal Processing, 2011, Roma, Itália, *Anais...*, [s.l.]: SciTePress - Science and Technology Publications, 2011, p. 504–509.

SUMÁRIO

1 Introdução	21
1.1 Objetivos	25
1.1.1 Objetivo geral.....	25
1.1.2 Objetivos específicos.....	25
2 Pesquisa Bibliográfica.....	27
2.1 Sinal de eletroencefalografia e a epilepsia	27
2.2 Detecção automática de padrões epileptiformes	31
2.3 Classificadores computacionais	35
2.3.1 Redes Neurais Artificiais.....	35
2.3.2 Análise de Componentes Principais	37
2.3.3 Análise de Discriminantes Linear	40
2.4 Métricas de análise de desempenho	42
2.5 Estudo comparativo.....	44
3 Metodologia	49
3.1 Base de dados – sinais de EEG	Erro! Indicador não definido.
3.1.1 Seleção dos dados.....	52
3.2 Conjunto de descritores.....	54
3.2.1 Processamento dos sinais de EEG.....	57
3.2.2 Cálculo e análise dos descritores.....	64
3.3 Classificadores	68
3.3.1 Redes Neurais Artificiais.....	69
3.3.2 Redes Neurais combinadas com PCA	69
3.3.3 Análise de Discriminantes.....	70
3.3.4 Redes Neurais combinadas com LDA.....	71
3.4 Análise e seleção dos classificadores	72
4 Resultados e Discussão	75
5 Conclusões	97
Referências.....	103
Apêndice A – Protocolo para marcação das telas de EEG apresentado aos especialistas.	117

Apêndice B – Resumo dos elementos do conjunto de descritores	119
Apêndice C – Interface gráfica utilizada nos testes dos filtros	121
Apêndice D – Fluxograma do funcionamento do algoritmo desenvolvido para calcular o conjunto de descritores.....	123
Apêndice E – Interface gráfica utilizada nas simulações e testes para implementação das redes neurais	125
Apêndice F – Fluxograma do processo de treinamento realizado nas redes neurais artificiais desenvolvidas.....	127
Apêndice G – Interface gráfica utilizada nas simulações para teste e seleção dos classificadores.....	129
Apêndice H – Matriz de Correlação dos descritores da Base A	131
Apêndice I – Valores das constantes e pesos das Funções Discriminantes obtidas nas simulações.....	137
Apêndice J – Desempenho dos classificadores nas simulações com os registros da base D e marcações do especialista 1.....	139
Apêndice K – Desempenho dos classificadores nas simulações com os registros da base D e marcações do especialista 2.....	143
Apêndice L – Análise do desempenho de classificação dos registros de EEG da Base D utilizando o especialista 1 como referência e o especialista 2 no papel de classificador.....	147
Apêndice M – Análise do desempenho de classificação dos registros de EEG da Base D utilizando o especialista 2 como referência e o especialista 1 no papel de classificador.....	149

1 INTRODUÇÃO

A Epilepsia é uma condição crônica, ou um grupo de doenças, de alta prevalência, ainda pouco explicada pela ciência e cujo diagnóstico ainda é predominantemente clínico (MORMANN *et al.*, 2007). Apesar dos neurologistas poderem realizar o diagnóstico de epilepsia através da anamnese, a análise do eletroencefalograma (EEG) é uma importante ferramenta que pode ser utilizada para a confirmação clínica do diagnóstico, auxílio na definição do tipo de síndrome, apoio no planejamento de terapia medicamentosa e auxílio na decisão sobre a viabilidade cirúrgica (ASKAMP; VAN PUTTEN, 2014; FERNÁNDEZ-TORRE, 2010; SANEI; CHAMBERS, 2007).

A confirmação clínica do diagnóstico da epilepsia é realizada por uma criteriosa análise visual dos registros de EEG obtidos de um indivíduo com suspeita de possuir esta doença. Esta análise é realizada por um especialista na área com a finalidade de encontrar padrões específicos na atividade eletroencefalográfica registrada. Estes padrões são denominados padrões epileptiformes (PILLAI; SPERLING, 2006).

De acordo com a literatura, os padrões epileptiformes são sinais eletrográficos característicos com amplitude em torno de 200 μV e duração entre 20 e 200 ms. Estes padrões normalmente, são morfológicamente caracterizados por um pico (espícula ou onda aguda) imediatamente seguidos por atividade de baixa frequência (onda lenta). Quando os padrões epileptiformes são encontrados nos registros de EEG com determinada frequência eles podem indicar clinicamente que o indivíduo possui epilepsia (SCHOMER; SILVA, 2011).

Os registros analisados na investigação da epilepsia geralmente são resultados de longos períodos de monitoramento do indivíduo. Desta forma, a análise dos registros pode ser um processo demorado e desgastante para o especialista. Comumente, este tipo de análise requer do especialista a verificação criteriosa de registros de, no mínimo, 24 horas de EEG obtidos a partir de 24 a 64 canais¹ simultâneos e que, geralmente, são apresentados em telas de 10 a 15 segundos (AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY, 2008; PILLAI; SPERLING, 2006). Em outras palavras, o processo exige que, para algumas horas de registro, o especialista analise visualmente milhares de telas nas quais é preciso reconhecer padrões de atividade

¹ Canal pelo qual os potenciais elétricos provenientes dos neurônios são obtidos pela utilização de um equipamento de aquisição de sinal eletrofisiológico.

que possuem amplitude de alguns microvolts e, no máximo, 200 ms de duração.

Por exemplo, considerando a revisão de um registro de EEG de longa duração obtido com as configurações mínimas exigidas (24 canais e 24 horas contínuas de registro) e analisado utilizando telas de 10 segundos ter-se-ia 8.640 telas para análise ou 576 horas de sinal contínuo ($24 \text{ h} \cdot 24 \text{ canais} = 576 \text{ h}$). Porém, considerando que o tempo médio de monitoramento é de sete dias chega-se a um valor de 60.480 telas de EEG que são analisadas pelo especialista para cada indivíduo. Este monitoramento demanda muito tempo além de ser uma tarefa cansativa e suscetível a interpretações errôneas se não realizada com máxima atenção.

No estudo de Wilson *et al.* (1996) os especialistas participantes relataram ter gasto em média 60 horas para realizar a marcação dos primeiros 50 padrões epileptiformes encontrados em cada um dos 50 registros de EEG de curta duração analisados no estudo. O grande tempo despendido para realizar uma análise meticulosa, que requer do revisor um alto nível de atenção, pode ser uma das razões pela qual a grande parte dos centros de análise não identificam todos os padrões epileptiformes presentes nos registros. Mas realizam a marcação de um subconjunto representativo das morfologias e das origens observadas (WILSON *et al.*, 2003).

Nas últimas décadas, inúmeros estudos propuseram metodologias ou algoritmos para automatizar este tipo de análise e facilitar o trabalho dos especialistas, reduzindo o tempo necessário para realizar a revisão desses registros, além de aumentar a confiabilidade do processo. No entanto, apesar dos esforços e relativo sucesso de alguns trabalhos, ainda não foi consolidada uma metodologia ou algoritmo para esse reconhecimento automático que demonstre desempenho adequado às necessidades dos especialistas da área (ASKAMP; VAN PUTTEN, 2014; HALFORD *et al.*, 2013; HARNER, 2009; SCHOMER; SILVA, 2011).

Os trabalhos existentes têm seu desempenho prejudicado, fundamentalmente, porque grande parte não passou por um processo adequado de validação estatística ou então apresentaram algumas limitações vinculadas à metodologia empregada, que dificultam, e por vezes até impossibilitam a sua aplicação em registros de EEG diferentes daqueles utilizados durante o desenvolvimento metodológico (HALFORD, 2009; WILSON; EMERSON, 2002). Um dos motivos para a resistência à aplicação dos processos automatizados no ambiente clínico pode ser a insuficiência de testes das metodologias formuladas,

uma vez que os trabalhos desenvolvidos e publicados, apesar de demonstrarem resultados relativamente bons, são realizados testes utilizando constantemente bases de dados muito restritas (ASKAMP; VAN PUTTEN, 2014; GUERRERO-MOSQUERA; TRIGUEROS; NAVIA-VAZQUEZ, 2012; WILSON; EMERSON, 2002). As bases de dados usadas nos estudos descritos na literatura são pequenas quando comparadas com a quantidade de registros de EEG analisados pelos especialistas durante seu treinamento. E ainda, a avaliação realizada para que os especialistas sejam considerados aptos a analisar registros de EEG sem supervisão é muito mais rigorosa do que os testes realizados com os sistemas automatizados.

Outro aspecto negativo relativo ao desempenho de sistemas ou metodologias de revisão automática de EEG é devido à elevada ocorrência de resultados equivocados (falsos positivos) que, conseqüentemente, remetem a baixos valores de especificidade da classificação dos padrões. Além disso, a baixa especificidade implica em pouca, ou até mesmo nenhuma, economia efetiva de tempo, pois a marcação realizada de forma automatizada precisa ser revisada pelo especialista com a finalidade de eliminar os falsos positivos que podem prejudicar a análise clínica (ANDERSON; DOOLITTLE, 2010; HALFORD *et al.*, 2013; HARNER, 2009; LODDER; VAN PUTTEN, 2014; WILSON; EMERSON, 2002).

Por outro lado, alguns estudos mostraram que o grau de concordância no processo de revisão dos registros de EEG, entre diferentes especialistas, também pode ser considerado baixo, na faixa de 50% a 80%. Este fato pode ocorrer por vários motivos, com destaque para a fadiga do profissional que a realiza, pois o processo de análise é extenuante. Contribui, ainda, diferentes níveis de experiência ou tipo de treinamento dos especialistas envolvidos ou, ainda, a falta de tempo disponível para a realização do processo de revisão (ANDERSON; DOOLITTLE, 2010; BARKMEIER *et al.*, 2012; DÜMPPELMANN; ELGER, 1999; HALFORD *et al.*, 2013; SCHOMER; SILVA, 2011; WILSON *et al.*, 1996).

Além dos motivos citados anteriormente, a falta de concordância entre os especialistas pode ocorrer também em função da alta variabilidade do sinal de EEG tanto entre indivíduos diferentes, quanto no mesmo indivíduo quando este se encontra em diferentes estados comportamentais e de consciência (SHIBASAKI *et al.*, 2014).

Associada a baixa concordância entre os especialistas, a alta variabilidade do sinal de EEG e as variações morfológicas dos padrões epileptiformes resultam na ausência de um “padrão-ouro” que possa ser

utilizado no desenvolvimento, avaliação e validação das metodologias e dos processos de revisão automatizada (ASKAMP; VAN PUTTEN, 2014; GRANT *et al.*, 2014).

Outra dificuldade encontrada no desenvolvimento da automatização do processo de revisão dos registros de EEG é a inexistência de limites mínimos para o desempenho do processo automatizado para que este seja clinicamente aceito.

Assim, a problemática inerente ao desenvolvimento e implementação das metodologias para o processamento automático dos sinais de EEG e a posterior classificação de padrões eletrográficos requer estudos mais aprofundados, no sentido de superar as limitações encontradas, motivou a elaboração deste trabalho.

Além disso, a identificação automática dos padrões epileptiformes e a análise morfológica dos paroxismos podem ser aplicadas em estudos relacionados à determinação de um prognóstico cirúrgico (KRENDL; LURGER; BAUMGARTNER, 2008), na determinação da severidade das crises epiléticas (SELVITELLI *et al.*, 2010), na análise da correlação entre a frequência dos paroxismos e disfunções cognitivas (NICOLAI; KASTELEIJN-NOLST TRENITÉ, 2011; RODIN *et al.*, 2009), na análise da frequência de paroxismos em indivíduos sem histórico de crises epiléticas (SO, 2010) e em investigação da epilepsia (HARNER, 2009; LODDER; VAN PUTTEN, 2014; MILLER; GOTMAN, 2008; SCHOMER; SILVA, 2011).

Desta forma, justifica-se a realização de novos estudos, como o proposto neste trabalho, devido a necessidade de identificar uma metodologia que auxilie no processo de revisão dos registros de EEG de longa duração e que possibilite que as deficiências e as limitações apresentadas por métodos existentes sejam superadas e atinjam um desempenho aceitável para sua implementação e a sua utilização em um ambiente clínico.

1.1 OBJETIVOS

O presente trabalho norteia-se pelos seguintes objetivos.

1.1.1 Objetivo geral

Contribuir para o reconhecimento automático de padrões epileptiformes em sinais de eletroencefalograma através do desenvolvimento e avaliação de uma metodologia baseada em análise morfológica.

1.1.2 Objetivos específicos

- Determinar as bases de dados de registros de EEG que serão utilizadas;
- Definir o conjunto de características morfológicas da metodologia proposta;
- Desenvolver um algoritmo para a extração automática do conjunto de características;
- Definir as técnicas de Processamento Digital de Sinais aplicadas na metodologia proposta;
- Definir os classificadores computacionais utilizados nos testes da metodologia proposta;
- Determinar as métricas para a avaliação do desempenho dos classificadores.

2 PESQUISA BIBLIOGRÁFICA

Esse capítulo aborda os principais conceitos que fornecem sustentação a este trabalho. Esses conceitos permitem fundamentar a metodologia proposta nesta tese.

2.1 SINAL DE ELETROENCEFALOGRAFIA E A EPILEPSIA

O primeiro registro de atividade cerebral na forma de sinais elétricos foi realizado em 1875 pelo cientista inglês Richard Caton utilizando um galvanômetro e inserindo dois eletrodos sob o escalpo de um indivíduo. Desde então, o eletroencefalograma (EEG) tem sido usado para denotar a atividade neural elétrica do cérebro (Sanei e Chambers, 2007).

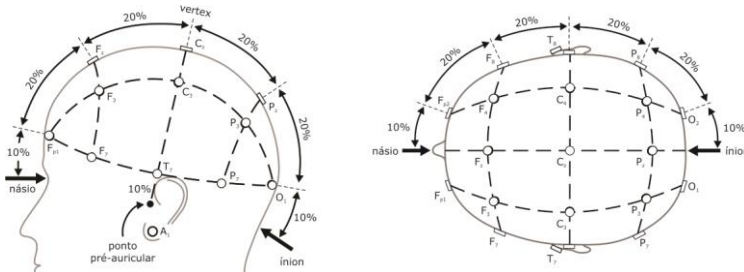
A aquisição do sinal de EEG pode ser realizada introduzindo eletrodos no interior do tecido cerebral (registro de profundidade), posicionando eletrodos diretamente na superfície exposta do córtex cerebral (Eletrocorticograma – ECoG) ou posicionando eletrodos, de forma não invasiva, na superfície do couro cabeludo (Eletroencefalograma de superfície – sEEG).

De acordo com algumas indicações da Federação Internacional de Neurofisiologia Clínica (DEUSCH; EISEN, 1999) e diretrizes da Sociedade Americana de Neurofisiologia Clínica (AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY, 2006a, 2006b, 2006c, 2008), o EEG superficial é o mais indicado para a realização da aquisição tanto do EEG de rotina, que é um registro ambulatorial com duração média de 20 minutos, como do EEG de longo período, que consiste em um registro normalmente adquirido quando o indivíduo encontra-se em observação de, no mínimo, 24 horas para confirmação clínica de diagnóstico de epilepsia (PANAYIOTOPOULOS, 2010).

Outras importantes recomendações realizadas referem-se à utilização de no mínimo 21 eletrodos para a aquisição registros de EEG de rotina e 24 ou 32 eletrodos para a aquisição de registros de longo período. Em ambos os casos é indicado que os eletrodos sejam dispostos de acordo com o sistema 10/20, ilustrado na Figura 1, e que a montagem utilizada na aquisição dos registros deve ser do tipo referencial. Este tipo de montagem, ilustrada na Figura 2, é indicada para a aquisição desses registros, pois o processo de aquisição dos potenciais elétricos de cada

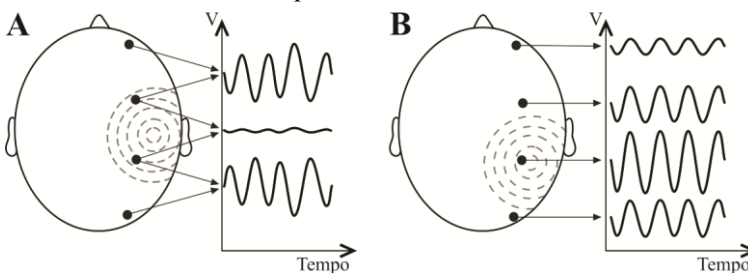
um dos eletrodos é realizado em relação a um mesmo ponto referencial. Desta forma, permite-se que os registros possam ser visualizados posteriormente em diferentes tipos de montagem como, por exemplo, a montagem bipolar, também ilustrada na Figura 2, na qual o potencial elétrico do sinal adquirido é resultado da diferença de potencial entre pares de eletrodos.

Figura 1 – Sistema 10/20 para o posicionamento de eletrodos.



Fonte: Adaptado de Malmivuo e Plonsey (1995).

Figura 2 – Diferenças do sinal de EEG quando em (A) montagem bipolar e (B) unipolar ou referencial.



Fonte: Adaptado de Malmivuo e Plonsey (1995).

O EEG superficial (de escalpo) em seres humanos, em relação às suas características, ocupa uma banda em frequência de 0 a 100 Hz e uma faixa de amplitudes de 2 a 200 μV . No entanto, de forma geral, o sinal concentra-se entre 0,5 e 60 Hz com amplitude média de 50 μV (SCHOMER; SILVA, 2011).

Estas características de amplitude e frequência do sinal são influenciadas pela localização em relação ao córtex cerebral a partir da qual foi adquirido o registro, bem como da idade, do estado físico (sono,

vigília, coma, e outros) e do estado comportamental (depressão, excitação, euforia, stress, entre outros) do indivíduo.

Além disso, o sinal também pode ser distorcido por interferências originadas das mais diversas fontes como, por exemplo, dos potenciais elétricos extra cerebrais do próprio paciente, dos eletrodos e do sistema de aquisição do sinal e de interferências eletromagnéticas externas. Isto significa que os registros podem apresentar variabilidade entre os indivíduos sob circunstâncias idênticas e também apresentar variações em um mesmo indivíduo ao longo do tempo. Isso contribui para dificultar muito uma possível modelagem matemática dos padrões eletrográficos comumente presentes no EEG.

Por outro lado, estes padrões possuem uma relativa regularidade em frequência, morfologia e amplitude, ou seja, o espectro de frequências do EEG é normalmente subdividido em faixas de frequência que podem estar relacionadas a diversos estados físicos e comportamentais. De acordo com Sanei e Chambers (2007), esses padrões são:

- Banda Alfa (α) – 8 a 13 Hz – Ritmo comum em pacientes normais e mais facilmente observável no indivíduo acordado, relaxado e com os olhos fechados;
- Banda Beta (β) – 13 a 22 Hz – Ritmo comum em pacientes adultos normais, durante o estado de vigília. Dominante na região pré-central do cérebro, mas aparecendo também em outras regiões cerebrais;
- Banda Teta (θ) – 4 a 8 Hz – Ritmo frequente em crianças, nas regiões central e temporal. Típico de estágios iniciais do sono. Alguns transientes de componentes deste ritmo têm sido encontrados em pacientes adultos normais;
- Banda Delta (δ) – 0,5 a 4 Hz – Ritmo comum em crianças, principalmente em recém-nascidos, em estado de sono profundo. Pode-se ressaltar que a presença desta onda em adultos sob estado de alerta pode indicar anormalidades.

O registro de EEG é usado, entre outros, no monitoramento do estado de vigília, coma e morte encefálica, na localização de áreas danificadas após traumatismo craniano, em acidente vascular cerebral e tumor, na realização de testes de vias aferentes, no monitoramento da profundidade anestésica, em investigação da fisiologia e dos distúrbios do sono e na investigação de epilepsia e localização do foco das crises epiléticas (SANEI; CHAMBERS, 2007).

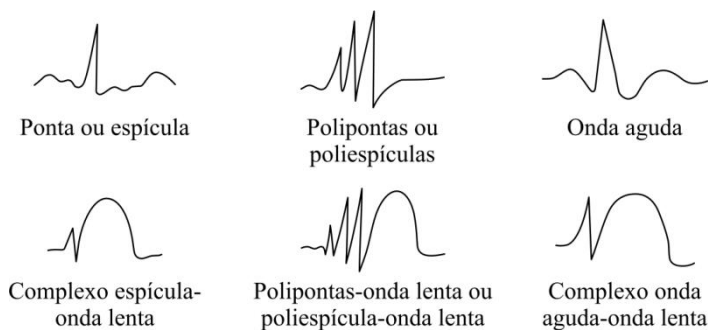
Dentre todas essas aplicações, a investigação da epilepsia representa uma das principais possíveis aplicações finais da metodologia proposta nesse trabalho, razão pela qual é selecionada como campo de aplicação nessa tese.

A epilepsia é uma doença caracterizada por uma persistente predisposição a gerar crises epiléticas e pelas consequências neurobiológicas, cognitivas, psicológicas e sociais desta condição (FISHER *et al.*, 2005). A definição de epilepsia requer a ocorrência de pelo menos uma crise epilética, ou seja, uma ocorrência transitória de sinais ou sintomas devidos à atividade anormal excessiva ou síncrona de neurônios no cérebro com ou sem fatores desencadeantes conhecidos (BERG, 2010).

Apesar dos neurologistas poderem realizar um diagnóstico de epilepsia pela anamnese, a análise do sinal de EEG é uma ferramenta comumente utilizada e importante que pode ser usada para a confirmação clínica do diagnóstico, no auxílio para a definição do tipo de síndrome epilética, para fornecer informação necessária ao planejamento da terapia medicamentosa e, ainda, para auxiliar a tomada de decisão sobre viabilidade de intervenção cirúrgica (ASKAMP; VAN PUTTEN, 2014; FERNÁNDEZ-TORRE, 2010; SANEI; CHAMBERS, 2007).

A ocorrência de certas manifestações eletrográficas no sinal de EEG, denominadas padrões epileptiformes, é um forte indicador da presença desta patologia. Um padrão epileptiforme, cujas diferentes morfologias e denominações são apresentadas na Figura 3, tem suas características morfológicas básicas relacionadas à sua amplitude e duração. As espículas possuem uma duração de 20 a 70 ms, enquanto que uma onda aguda possui duração de 70 a 200 ms, mas como ambas podem ser um padrão epileptiforme, uma distinção entre elas não tem aplicação do ponto de vista clínico na detecção automática de padrões. Por isso, considera-se que a duração de um paroxismo pode variar de 20 a 200 ms. Os valores das amplitudes de ambas (espículas e ondas agudas) são variados, porém, considerando os padrões epileptiformes, o valor da amplitude (em módulo) geralmente encontra-se entre 20 μV e 200 μV (SCHOMER; SILVA, 2011).

Figura 3 – Morfologia e respectiva denominação dos padrões epileptiformes mais encontrados em registros de EEG.



Fonte: Montenegro *et al.* (2012).

Um especialista na leitura de registros de EEG, denominado eletroencefalografista, emprega tempo considerável para analisar cada registro, especialmente quando eles são adquiridos por longos períodos de tempo (mais de 24 horas) e utilizando muitos canais, entre 24 e 128 canais simultâneos (HARNER, 2009; PILLAI; SPERLING, 2006). Por exemplo, considerando a revisão de um registro de EEG de monitoramento contínuo de um indivíduo por três dias completos (72 horas) usando telas de 15 segundos ter-se-ia 17.280 telas para análise, o que demanda tempo, além de ser uma tarefa cansativa e suscetível a interpretações incompletas se não realizada com a máxima atenção. Estudos que avaliaram e analisaram a concordância entre os especialistas na tarefa de revisão de registros de EEG para a identificação de padrões epileptiformes indicam que o tempo despendido para a revisão de cada hora de registro pode ser no melhor caso, de 40 minutos (NONCLERCQ *et al.*, 2012) ou variar entre 3,6 e 10 horas (HALFORD, 2009; WILSON *et al.*, 1996), dependendo do nível de exigência empregado na identificação dos padrões.

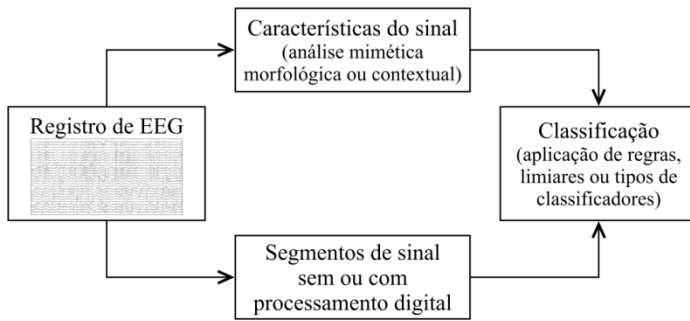
2.2 DETECÇÃO AUTOMÁTICA DE PADRÕES EPILEPTIFORMES

Nas últimas décadas, estudos propuseram metodologias ou sistemas automatizados, na tentativa de facilitar a tarefa de revisão dos registros de EEG. De forma geral, o funcionamento das metodologias e algoritmos desenvolvidos pode ser dividido em duas etapas: análise digital e classificação do sinal de EEG. Quanto à etapa de análise do

senal, os métodos utilizados podem ser divididos em duas grandes classes, ilustradas na Figura 4:

- métodos que utilizam segmentos dos sinais de EEG, com ou sem processamento digital adicional;
- métodos que utilizam características dos sinais de EEG.

Figura 4 – Classificação dos métodos de detecção automática de paroxismos.



Fonte: próprio autor.

Os métodos que utilizam segmentos do sinal de EEG funcionam, basicamente, aplicando o sinal de EEG diretamente à etapa de classificação. Os sinais podem ser utilizados da mesma forma como foram adquiridos (WEBBER *et al.*, 1994) ou após terem passado por alguma técnica de processamento digital mais específico. Os principais tipos de processamento digital utilizado nesses métodos são a Transformada Wavelet, a Análise de Componentes Independentes, os filtros morfológicos, os modelos paramétricos e o *Template Matching*.

O processamento com a Transformada Wavelet (TW) é utilizado para isolar componentes de frequência no sinal com uma melhor resolução temporal. Além disso, os diferentes tipos de função Wavelet e a quantidade de escalas que são utilizados na análise permitem que sejam realçadas diferentes características dos padrões presentes no sinal de EEG (ARGOUD *et al.*, 2006; INDIRADEVI *et al.*, 2008; SCOLARO, 2014).

A Análise de Componentes Independentes (ICA) realiza a separação das diferentes formas de onda presentes no sinal de acordo com a sua fonte. Esta análise é muito empregada para a eliminação de artefatos ruidosos, uma vez que o sinal resultante é composto apenas pelos padrões de interesse. É utilizada também para aplicações de localização do foco das descargas epileptiformes (DE LUCIA *et al.*,

2008; OSSADTCHI *et al.*, 2004). Quando a TW e a ICA são utilizadas, a maioria dos métodos aplica os resultados do processamento, ou seja, o sinal de EEG, à ferramentas de inteligência computacional para que seja realizada a classificação dos padrões.

Os modelos paramétricos, em sua maioria modelos auto-regressivos, e os filtros morfológicos são utilizados para destacar os padrões epileptiformes e atenuar a quase zero as demais formas de onda presentes. Nesses casos, a identificação dos padrões epileptiformes é realizada utilizando um limiar que, se for ultrapassado, indica a presença de um padrão epileptiforme (FLANAGAN *et al.*, 2003; LIU; ZHANG; YANG, 2002; OIKONOMOU; TZALLAS; FOTIADIS, 2007; XU *et al.*, 2007).

O método denominado *Template Matching* funciona através da seleção de um modelo de forma de onda para o padrão epileptiforme que é utilizado como um gabarito para a classificação dos padrões presentes no sinal analisado (EL-GOHARY; MCNAMES; ELSAS, 2008; LARSSON; WILSON; EEG-OLOFSSON, 2009; LODDER; ASKAMP; VAN PUTTEN, 2013; LODDER; VAN PUTTEN, 2014; NONCLERCQ *et al.*, 2009). A escolha desse modelo de padrão pode ser realizada pela simples promediação das várias formas de onda dos padrões epileptiformes ou o modelo pode ser obtido utilizando técnicas mais avançadas como, por exemplo, a Transformada Wavelet.

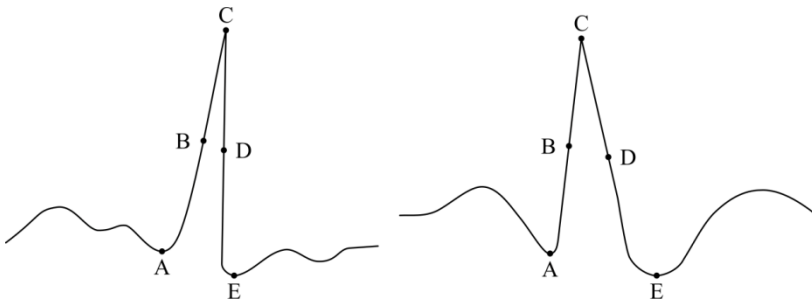
Os métodos que utilizam as características dos sinais de EEG são aqueles que extraem informações do sinal de EEG e utilizam esse conjunto de informações, normalmente valores numéricos, para a classificação dos padrões presentes no sinal. Segundo Tzallas *et al.* (2012), as características extraídas, que neste trabalho são denominados descritores, podem ser: morfológicas ou contextuais.

Os descritores morfológicos são valores numéricos referentes à morfologia do sinal como, por exemplo, amplitude, duração e angulação dos picos. O estudo de Gotman e Gloor (1976) é considerado um dos pioneiros na utilização desses descritores que são obtidos, inicialmente, pela decomposição do sinal em segmentos e sequências para, então serem calculados os valores de amplitude, duração e inclinação entre os pontos referenciais determinados pela decomposição inicial. Esses pontos referenciais são ilustrados na Figura 5. O conjunto gerado é formado pelos seguintes descritores:

- duração do intervalo entre os pontos A e C;
- duração do intervalo entre os pontos C e D;
- diferença de amplitude entre os pontos A e C;

- diferença de amplitude entre os pontos C e D;
- inclinação do segmento de reta B-C;
- inclinação do segmento de reta C-D.

Figura 5 – Pontos referenciais utilizados para o cálculo dos descritores de Gotman (1980), com base em um exemplo de uma espícula e uma onda aguda.



Fonte: próprio autor.

Uma série de métodos de detecção automática de eventos epileptiformes, como os utilizados por Guedes de Oliveira, Queiroz e Lopes da Silva (1983), Webber *et al.* (1994), Walczak e Nowack (2001), Acir *et al.* (2005), Tzallas *et al.* (2006), Kutlu, Kuntalp e Kuntalp (2009) e Shibasaki *et al.* (2014), foram desenvolvidos com base nos seis descritores de Gotman e Gloor (1976), ou ainda, adicionando alguns outros descritores morfológicos derivados deles.

Além dos descritores referentes às características morfológicas, outro tipo de descritor também muito empregado é o chamado descritor contextual. Este tipo de descritor representa características referentes ao contexto do sinal em análise como, por exemplo, o valor da variância, o valor da entropia, a localização no tempo e a identificação do canal onde ocorreu o padrão, os coeficientes extraídos da Transformada Wavelet, as informações sobre a densidade do espectro de frequência e a quantidade de cruzamentos pela linha de base (HALFORD *et al.*, 2013; PEREIRA, 2003; SOVIERZOSKI, 2009; ÜBEYLI, 2009).

Os métodos de detecção automática de padrões epileptiformes podem empregar conjuntos de descritores morfológicos, contextuais ou ainda um conjunto híbrido mesclando os dois tipos. No entanto, um ponto em comum entre todos esses métodos que utilizam descritores é a aplicação de métodos estatísticos de análise multivariada ou ferramentas de inteligência artificial para realizar a classificação dos padrões.

2.3 CLASSIFICADORES COMPUTACIONAIS

Uma etapa importante de qualquer sistema de identificação automática de padrões é a sua classificação, comumente representada por um conjunto de regras, algum método estatístico ou uma ferramenta de Inteligência Artificial.

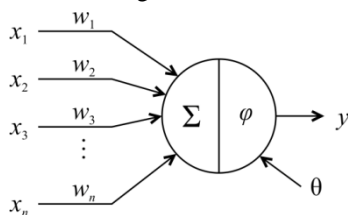
Alguns dos métodos e ferramentas mais aplicados em processos de automatização de processos de auxílio ao diagnóstico são os Sistemas Especialistas, as Máquinas de Vetores de Suporte, a Análise de Fatores, a Análise de Discriminantes, a Análise de Componentes Principais e as Redes Neurais Artificiais.

2.3.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) são uma ferramenta de Inteligência Artificial (ou Computacional) de abordagem conexionista que procura imitar o discernimento humano, mas sem o estabelecimento de alguma regra.

As redes neurais, de forma geral, são um modelo matemático inspirado na estrutura neural de organismos inteligentes que adquirem conhecimento por experiência. Elas são compostas por diversas unidades de processamento, denominadas neurônios artificiais e ilustradas na Figura 6, que se encontram conectadas entre si por conexões sinápticas.

Figura 6 – Modelo geral do neurônio artificial.



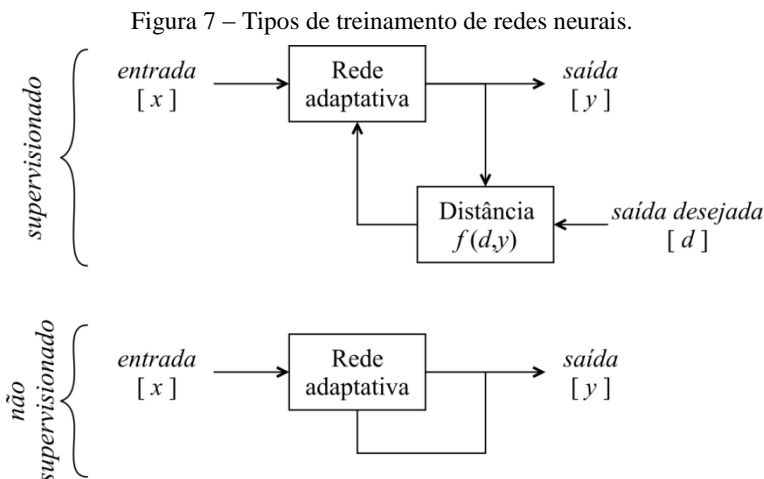
Fonte: adaptado de Zurada (1992).

Ao longo dos anos foram desenvolvidos alguns modelos de neurônios, com respectivas regras de aprendizado. Alguns exemplos dos mais comuns e populares são o McCulloch-Pitts, o *Adaptive Linear Neuron* (ADALINE) e o Perceptron.

A partir da definição da rede como uma associação de neurônios em grupos denominados camadas, é preciso definir como todos esses neurônios são conectados e como eles se comportarão para que a rede possa adquirir conhecimento.

Segundo Zurada (1992), de acordo com a associação dos seus neurônios, uma rede pode ser classificada em direta (*feedforward*) ou recursiva (*feedback*). E, de acordo com a quantidade de camadas de neurônios, as redes podem ser de uma única camada (*single layer*) ou de múltiplas camadas (*multilayer*).

Outra definição importante é o modo como a rede neural adquire o conhecimento. Isso é determinado pelo tipo de treinamento da rede e a regra de aprendizado aplicada. Os tipos de treinamento da rede, ilustrados na Figura 7, são o supervisionado e o não supervisionado.



Fonte: adaptado de Zurada (1992).

O treinamento supervisionado é aquele no qual para cada padrão de entrada utilizado no treinamento é apresentada à rede a resposta desejada correspondente à classe à qual pertence o padrão apresentado. O treinamento não supervisionado é aquele em que os padrões de treinamento são apresentados sem indicação de um valor correspondente para que a própria rede se encarregue de agrupar os padrões de entrada em classes de acordo com a semelhança de características entre os padrões de entrada.

As regras de aprendizagem utilizadas no treinamento das redes indicam de que forma os pesos sinápticos das conexões devem ter seus valores modificados para que ao final do treinamento a rede tenha obtido conhecimento suficiente para aprender a classificar corretamente os padrões de entrada utilizados no treinamento e também os novos padrões que sejam a ela apresentados. Existem várias regras de aprendizado como, por exemplo, a regra de Hebb, do Perceptron, de Widrow-Hoff, Delta e *Error Backpropagation* que estão descritas em detalhe na literatura.

Considerando o contexto deste trabalho, no qual os sistemas e algoritmos de automatização de detecção de padrões epileptiformes normalmente utilizam métodos de Inteligência Artificial para realizar os processos de classificação ou reconhecimento dos padrões, as redes neurais são uma ferramenta frequentemente utilizada (ADELI; GHOSH-DASTIDAR; DADMEHR, 2010; BOOS *et al.*, 2013; CASSON; LUNA; RODRIGUEZ-VILLEGAS, 2009; CHAVAKULA *et al.*, 2013).

2.3.2 Análise de Componentes Principais

A Análise de Componentes Principais (PCA) é um método estatístico que pode ser empregado com o objetivo de extrair novas informações acerca de um conjunto de dados específico ou minimizar a dimensão de um conjunto de dados sem resultar em perda significativa de informação (RAYKOV; MARCOULIDES, 2008).

No contexto deste trabalho, este método pode ser utilizado para, entre outros, reduzir o volume de informação apresentado às ferramentas de inteligência artificial empregadas na construção de sistemas de detecção automática (BOOS; AZEVEDO, 2012). Alguns exemplos das aplicações de PCA neste contexto são os trabalhos relacionados na sequência:

- Ghosh-Dastidar, Adeli e Dadmehr (2008), utilizaram a PCA para, além da redução de dimensionalidade, realizar o aprimoramento das nove características baseadas na Transformada Wavelet e na Teoria do Caos que são apresentadas a um classificador com a finalidade de melhorar o desempenho deste classificador;
- Wang *et al.* (2010), aplicaram a PCA com o objetivo de reduzir o tamanho dos sinais de EEG utilizados na pesquisa e

também realizaram a extração de características desses sinais;

- Sezer, Işık e Saracoğlu (2012), empregaram a PCA para reduzir a dimensão dos vetores de coeficientes da Transformada Wavelet que, neste caso, foi utilizada para extrair as características dos sinais de EEG que posteriormente serão classificados;
- Subasi e Gursoy (2010), apresentaram uma comparação entre os classificadores, chamados Máquinas de Vetor Suporte (*Support Vector Machines*), utilizando a Transformada Wavelet em combinação com a Análise de Discriminantes Linear (LDA), a Análise de Componentes Independentes (ICA) e a Análise de Componentes Principais (PCA) para a extração de características e redução de dimensionalidade dos dados apresentados aos classificadores.

Em termos gerais, conforme Myatt e Johnson (2009), a Análise de Componentes Principais é uma combinação linear de pesos das variáveis originalmente observadas no problema em análise. A PCA possibilita um melhor entendimento do conjunto de dados observados e a redução do número de variáveis apresentadas.

A descrição e o desenvolvimento matemático detalhado do funcionamento e da obtenção das componentes principais podem ser observados em diversas bibliografias, como Bilodeau e Brenner (1999), Härdle e Simar (2007), Raykov e Marcoulides (2008). No entanto, a síntese de como é utilizada a PCA pode ser exemplificada, de acordo com a aplicação desejada nesse trabalho, nos seguintes termos:

- a partir de uma matriz de dados iniciais “X” formada por “p” descritores obtidos de “n” sinais de EEG e representada pela Equação 1:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (1)$$

- considera-se que o vetor aleatório “x” apresentado na Equação 2 representa o conjunto de variáveis selecionadas para a análise:

$$\vec{x} = [x_1, x_2, \dots, x_p] \quad (2)$$

- calcula-se a matriz de covariância ou de correlação para todas as variáveis;
- e desenvolvendo matematicamente a matriz de covariância ou correlação calculada obtém-se, como mostrado na Equação 3, os autovalores “ λ ” e autovetores “ a ” de “ k ” componentes principais:

$$\lambda = [\lambda_1, \lambda_2, \dots, \lambda_k] \quad e \quad \vec{a} = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kp} \end{bmatrix} \quad (3)$$

- os autovetores são vetores com os pesos das combinações lineares utilizadas para calcular o valor das componentes principais “ y ” e os autovalores são o valor da variância de cada componente, ou seja, indicam a representatividade de cada componente principal calculada. A Equação 4 representa o valor das componentes principais “ y ”

$$\begin{aligned} y_1 &= a_1x = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ y_2 &= a_2x = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ &\dots \\ y_k &= a_kx = a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kp}x_p \end{aligned} \quad (4)$$

Neste exemplo descrito pelas Equações 1 a 4, após realizar a PCA em um determinado conjunto de variáveis, obtém-se um novo conjunto de variáveis composto de “ k ” elementos. Estes elementos são as “ k ” componentes principais, também denominadas fatores, cujo autovalor “ λ_k ” é o valor da variância apresentada pelo fator “ y_k ” e o somatório dos autovalores, representado por $\sum_{i=0}^k \lambda_i$, é utilizado para indicar a representatividade do conjunto obtido.

Em alguns casos, os valores dos pesos obtidos usando a PCA são muito semelhantes para todas as variáveis. Isso pode dificultar a observação e a interpretação das componentes principais resultantes. Para simplificar e facilitar a análise, uma solução muito utilizada é a rotação das componentes, que pode ser realizada de diferentes formas. O

método chamado *Varimax* é o mais aplicado popularmente. Este método maximiza o valor dos pesos de algumas variáveis enquanto minimiza, a valores próximos de zero, os pesos das variáveis restantes. Isso resulta em componentes mais simplificadas e facilita a interpretação da associação entre as variáveis.

2.3.3 Análise de Discriminantes Linear

A Análise de Discriminantes (DA) é um método estatístico multivariado comumente utilizado em aplicações onde se almeja a classificação de uma observação entre duas ou mais categorias, redução de dimensionalidade ou ainda a extração de informação acerca de um conjunto de dados conhecido (MYATT; JOHNSON, 2009; RAYKOV; MARCOULIDES, 2008).

Os exemplos de aplicação desta análise para a classificação são os estudos de Chapman *et al.* (2007), Boostani, Sadatnezhad e Sabeti (2009) e Romo Vázquez *et al.* (2012), onde a DA é utilizada como classificador para, respectivamente, o diagnóstico precoce de Alzheimer, o diagnóstico de Esquizofrenia e a identificação de padrões ruidosos em sinais de EEG. Na redução de dimensionalidade, a DA é empregada para reduzir a quantidade de parâmetros necessários para que um classificador computacional possa realizar, por exemplo, o diagnóstico de Doença Arterial Coronariana, o diagnóstico de Diabetes ou a detecção de crises epiléticas, conforme descrito nos estudos de Giri *et al.* (2013), Çalişir e Doğantekin (2011) e Subasi e Gursoy (2010). A outra aplicação dessa análise pode ser exemplificada pelo estudo de Palaniappan (2006), que a emprega para extrair as características do sinal de EEG de indivíduos durante a simulação de atividades, com a finalidade de que estas características extraídas sejam posteriormente utilizadas como identificadores biométricos de indivíduos em interfaces homem-máquina (*Brain Computer Interfaces*).

No contexto desse trabalho, a Análise de Discriminantes é utilizada para a classificação de padrões em sinais de EEG e, como são analisadas duas classes de padrões, o método estatístico denomina-se Análise de Discriminantes Linear (LDA). O objetivo ao utilizar este recurso é a geração de uma função que separe linearmente as duas classes de padrões (HÄRDLE; SIMAR, 2007).

A LDA utiliza um conjunto de variáveis independentes, cujos valores numéricos são obtidos para um grupo de observações, ou casos, sobre os quais é conhecida a classe categórica a qual pertencem.

Segundo Härdle e Simar (2007), as variáveis independentes em conjunto com uma variável dependente, que indica a classe à qual as observações pertencem, são utilizadas na obtenção de uma função chamada Função Discriminante (FD), que além de classificar os dados analisados, pode também ser utilizada como um preditor na classificação de novas observações.

De forma geral, e conforme a Equação 5, essa função discriminante (FD) é apresentada na forma de uma combinação linear das n variáveis independentes, representadas pela letra v , cada uma com um peso correspondente, representado por w , cujo valor é obtido no final da análise. Além disso, a função possui uma constante, representada por c , que pode ser ocultada da função pelo processo de normalização da equação obtida.

$$FD = c + w_1 \cdot v_1 + w_2 \cdot v_2 + w_3 \cdot v_3 + \dots + w_n \cdot v_n \quad (5)$$

A obtenção da FD é realizada fundamentalmente com base nos valores da variância das classes e a quantidade de funções obtidas depende da quantidade de classes analisadas (GARSON, 2012). Neste caso, como são analisadas apenas duas classes de padrões, no final da análise é obtida uma única função.

Um ponto interessante neste tipo de análise é que grande parte dos *softwares* de análise estatística utilizada permitem que a função seja obtida pela análise das variáveis independentes em conjunto ou gradualmente.

Quando analisadas em conjunto, as variáveis independentes são testadas primeiro contra um referencial de tolerância, cujo valor normalmente é muito próximo de zero e as variáveis aprovadas no teste são incluídas na função (RAYKOV; MARCOULIDES, 2008). A análise gradual das variáveis independentes é realizada pela da adição e remoção das variáveis na FD de acordo com dois valores de referência definidos no início da análise. Esses valores são correspondentes ao resultado do Teste F utilizado para identificar o desempenho da FD a cada etapa. E, para a grande maioria dos softwares utilizados para LDA é permitida a escolha entre a utilização do valor real (f_{in} e f_{out}) ou do valor da significância (p_{in} e p_{out}) do resultado deste Teste F.

Além disso, essa análise gradual pode ser realizada usando alguns métodos distintos como, por exemplo, a variância não explicada, a distância de Mahalanobis e o Lambda de Wilks. De acordo com Garson (2012) e Maesschalck, Jouan-Rimbaud e Massart (2000), a distância de Mahalanobis é uma métrica utilizada para calcular a distância entre o

valor médio de duas distribuições, também chamado de centroide das distribuições, ou entre um ponto e o centroide de uma distribuição. Quando aplicado à LDA, este método torna-se interessante pois, por definição, ele aceita a utilização de variáveis independentes cuja distribuição não é totalmente normal.

Assim, de forma geral, a obtenção da função discriminante de modo gradual e utilizando a distância de Mahalanobis ocorre da seguinte maneira: a cada etapa da análise é adicionada a variável independente que maximiza essa distância entre os centroides da distribuição das classes em análise.

2.4 MÉTRICAS DE ANÁLISE DE DESEMPENHO

A definição detalhada de todos os parâmetros que fazem parte da metodologia de processamento dos dados e no classificador utilizado para automatizar a identificação de padrões epileptiformes em sinais de EEG é de extrema importância para possibilitar um bom funcionamento do processo.

Do mesmo modo, a determinação das métricas utilizadas para a avaliação do desempenho da classificação também é de extrema importância, pois algumas delas podem mascarar o real desempenho, prejudicando a realização de possíveis análises comparativas com outros processos de classificação e dificultando a análise da validação do método desenvolvido (CASSON; LUNA; RODRIGUEZ-VILLEGAS, 2009).

Um dos métodos mais comuns e simples de analisar o resultado de um classificador é a extração dos índices da chamada Tabela de Contingência ou Teste Diagnóstico (MYATT; JOHNSON, 2009). A denominação deste último é utilizada na literatura da área da saúde, uma vez que os dados do classificador aplicados na tabela são referentes ao diagnóstico de alguma patologia (VAN BELLE *et al.*, 2004).

Pela inserção dos dados esperados do classificador e dos dados realmente obtidos do classificador, conforme ilustra a Figura 8, é possível a extração dos índices listados na sequência, cujas definições foram feitas conforme o contexto desse trabalho:

- Verdadeiro Positivo (VP), indica a quantidade de identificações corretas de padrões epileptiformes;
- Verdadeiro Negativo (VN), representa a quantidade de identificações corretas de padrões não-epileptiformes;

- Falso Positivo (FP), corresponde à quantidade de padrões não-epileptiformes identificados incorretamente como epileptiformes;
- Falso Negativo (FN), representa a quantidade de padrões epileptiformes que foram identificados como não-epileptiformes.

Figura 8 – Tabela de Contingência de um classificador genérico.

		RESPOSTA ESPERADA	
		+	-
RESPOSTA REAL	+	Verdadeiro Positivo (VP)	Falso Positivo (FP)
	-	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Fonte: próprio autor.

Segundo Myatt e Johnson (2009), partir destes índices, podem ser calculados os valores da sensibilidade, da especificidade e da eficiência do classificador. A sensibilidade, também chamada de taxa de verdadeiros positivos, foi definida por Casson, Luna e Rodriguez-Villegas (2009) como a probabilidade de um padrão epileptiforme ser identificado corretamente. A Equação 6 define a sensibilidade.

$$Sensibilidade = \frac{VP}{VP + FN} \quad (6)$$

A especificidade ou taxa de verdadeiros negativos é definida como a probabilidade de um padrão não-epileptiforme ser identificado corretamente. Conforme a Equação 7, ela é calculada utilizando os valores de verdadeiros negativos e falsos positivos.

$$Especificidade = \frac{VN}{VN + FP} \quad (7)$$

Segundo Myatt e Johnson (2009), a eficiência de um classificador descrita na Equação 8 pode ser definida como a taxa de acertos, uma vez que ela é calculada com base na razão entre o número de padrões classificados corretamente e o número total de padrões analisados.

$$Eficiência = \frac{VP + VN}{VP + VN + FP + FN} \quad (8)$$

Outras duas importantes métricas de análise do desempenho são o Valor Preditivo Positivo (VPP), que equivale à precisão do classificador, e o Valor Preditivo Negativo (VPN). Esses valores preditivos são definidos, respectivamente, como a proporção de padrões epileptiformes identificados corretamente e a proporção de padrões não-epileptiformes identificados corretamente. Os valores de VPP e VPN são obtidos utilizando as Equações 9 e 10, respectivamente.

$$VPP = \frac{VP}{VP + FP} \quad (9)$$

$$VPN = \frac{VN}{VN + FN} \quad (10)$$

2.5 ESTUDO COMPARATIVO

Em complementação à fundamentação teórica apresentada nos itens anteriores, esta seção apresenta um quadro comparativo de metodologias e sistemas descritos na literatura do estado da arte no qual se insere este trabalho.

Para sua elaboração foram selecionados, inicialmente, 32 estudos sendo 30 deles publicados em periódicos internacionais indexados e dois publicados em anais da conferência mundial do *Institute of Electrical and Electronics Engineers* (IEEE) na área de Engenharia Biomédica.

Dentre os estudos analisados, 13 foram excluídos devido à inexistência de informações técnicas necessárias à análise, como, por exemplo, a base de dados utilizada, ou por não apresentarem resultados que pudessem ser comparados com os demais trabalhos, isto é, os resultados foram apresentados apenas na forma gráfica (GUEDES DE OLIVEIRA; QUEIROZ; LOPES DA SILVA, 1983) ou porcentagem de falsas detecções (GOTMAN; GLOOR, 1976).

Destá forma, as características gerais dos 19 estudos analisados e os respectivos resultados são apresentados no Quadro 1. Destaca-se que as células em branco no quadro correspondem à informação não disponibilizada no estudo.

Quadro 1 – Características e desempenho de metodologias ou sistemas descritos na literatura sobre automatização reconhecimento de padrões epileptiformes.

Estudo	Esp ^a	Base de dados	FA ^b [Hz]	Montagem ^c	FPA ^d	FPB ^d	Metodologia	Classificação ^e	Se ^f	Es ^f	Ef ^f	VPP ^f
Webber <i>et al.</i> (1994)	1	10 EEGs (4min)	200	bipolar	0,30	70	descritores	MLP	73,6%	73,6%	73,6%	
Webber <i>et al.</i> (1994)	1	10 EEGs (4min)	200	bipolar	0,30	70	filtragem	MLP	45,7%	45,7%	45,7%	
Wilson <i>et al.</i> (1999)	1	15 segmentos (222s)	256	referencial	1,00	90	descritores	MMNN	89,9%	99,5%		80,8%
Pereira (2003)	3	45 segmentos (1s)	100	bipolar	0,50	40	signal	MLP	90,0%	80,0%	84,4%	78,3%
Acir <i>et al.</i> (2005)	2	10 EEGs (19min 48s)	256	referencial	1,00	70	descritores	RBF e SVM	89,1%			85,9%
Argoud <i>et al.</i> (2006)	3	270 segmentos (15s)	100	bipolar			Wavelet	MLP e SE	70,9%	100,0%		
Exarchos <i>et al.</i> (2006)	2	25 EEGs (15min)	256	bipolar	1,60	70	descritores	conjunto de regras	86,1%	91,8%		83,1%
Oikonomou; Tzallas; Fotiadis (2007)	1	13 segmentos (7min 30s)	256	bipolar	1,60	70	modelo AR	aplicação de <i>threshold</i>	85,3%			29,5%
De Lucia <i>et al.</i> (2008)	1	7 EEGs (17min 42s)	256	referencial	0,50	70	ICA	–	65,0%	80,0%		
Indiradevi <i>et al.</i> (2008)	2	20 EEGs	256	referencial	0,50	100	Wavelet	aplicação de <i>threshold</i>	91,7%	89,3%	90,5%	78,1%
Sovierzoski; Schwarz; Azevedo (2009)	1	200 segmentos (1s)	100	bipolar	0,50	40	signal	MLP	85,9%	87,1%		86,7%
Kutlu; Kuntalp; Kuntalp (2009)	2	119 segmentos (160ms)	256	referencial	1,00	70	descritores	MLP	90,8%	94,9%		
Sovierzoski (2009)	1	200 segmentos (1s)	100	bipolar	0,50	40	descritores	MLP	92,9%	90,1%		90,2%
Übeyli (2009)		500 segmentos (23,6s)	173,61	referencial	0,53	40	descritores	MLP	95,5%	96,5%	96,3%	
Scolaro; Boos; Azevedo (2011)	1	285 segmentos (1s)	512	referencial	0,50	40	Wavelet	MLP	97,1%	94,6%	96,1%	
Barkmeier <i>et al.</i> (2012)	3	9 ECoGs (10min)	200	referencial			filtragem e descritores	aplicação de <i>threshold</i>	50,1%		33,6%	
Sezer; İşik; Saracoğlu (2012)		200 segmentos (23,6s)	173,61	referencial	0,53	40	Wavelet	MLP	92,9%	100,0%	96,5%	100,0%
Chavakula <i>et al.</i> , (2013)	2	11 segmentos (6min 30s)	256	referencial			Wavelet e descritores	RNB	77,6%	87,6%		
Chavakula <i>et al.</i> (2013)	2	4 segmentos (5min)	256	referencial			Wavelet e descritores	RNB	43,0%	89,0%		
Chavakula <i>et al.</i> (2013)	2	11 segmentos (6min 30s)	256	referencial			Wavelet e descritores	SVM	69,6%	95,7%		
Chavakula <i>et al.</i> (2013)	2	4 segmentos (5min)	256	referencial			Wavelet e descritores	SVM	75,0%	82,0%		
Halford <i>et al.</i> (2013)	7 a 11	100 segmentos (30s)	256	referencial	1,00	70	descritores	MLP e RNB	58,5%	68,3%		
Halford <i>et al.</i> (2013)	4 a 7	100 segmentos (30s)	256	referencial	1,00	70	descritores	MLP e RNB	60,1%	66,5%		
Halford <i>et al.</i> (2013)	6 ou 7	100 segmentos (30s)	256	referencial	1,00	70	descritores	MLP e RNB	54,0%	72,8%		
Zhou <i>et al.</i> (2013)	7	100 segmentos (30s)	256	referencial	1,00	70	Wavelet e descritores	análise de <i>clusters</i>	84,7%	76,3%		
Zhou <i>et al.</i> (2013)	7	100 segmentos (30s)	256	referencial	1,00	70	Wavelet e descritores	análise de <i>clusters</i>	83,2%	75,9%		
Zhou <i>et al.</i> (2013)	7	100 segmentos (30s)	256	referencial	1,00	70	Wavelet e descritores	análise de <i>clusters</i>	85,6%	75,6%		
Zhou <i>et al.</i> (2013)	7	100 segmentos (30s)	256	referencial	1,00	70	Wavelet e descritores	análise de <i>clusters</i>	83,9%	75,9%		
Zhou <i>et al.</i> (2013)	7	100 segmentos (30s)	256	referencial	1,00	70	Wavelet e descritores	análise de <i>clusters</i>	81,1%	74,8%		

^a Quantidade de especialistas que realizaram a marcação dos registros de EEG utilizados.^b Valor da frequência de amostragem dos sinais de EEG utilizados.^c Indicação o tipo de montagem de eletrodos utilizada para a aquisição do EEG.^d Indicação da frequência de corte dos filtros passa-alta (FPA) e filtro passa-baixa (FPB) aplicados na aquisição do sinal de EEG.^e Tipo de classificador utilizado na metodologia ou sistema: rede neural Perceptron Multicamadas (MLP), rede neural de base radial (RBF), *support vector machine* (SVM), rede neural monotônica (MMNN), sistema especialista (SE), rede neural Bayesiana (RNB).^f Métrica de análise do desempenho: sensibilidade (Se), especificidade (Es), eficiência (Ef) e valor preditivo positivo (VPP).

A análise do Quadro 1 permite observar que grande parte dos estudos (76%) utiliza segmentos de EEG em suas análises. Os trabalhos com EEG completo também poderiam ser considerados segmentos, pois a duração dos registros é de apenas alguns minutos. A escassez de métodos testados com registros de longa duração, ou seja, de no mínimo 24 horas, é devido à dificuldade de encontrar tais registros que já tenham sido analisados e marcados por especialistas.

Em relação às marcações dos registros de EEG constatou-se que todos os estudos avaliaram seus métodos e classificadores utilizando marcações provenientes da concordância absoluta ou majoritária de um grupo de especialistas ou a marcação de apenas um especialista. Esta é uma questão relevante para a análise do desempenho alcançado porque os bons resultados obtidos podem ter sido resultado de influência da forma de marcação dos registros.

Considerando que os estudos realizaram testes com registros de EEG extraídos da mesma base de dados utilizada para o treinamento dos classificadores, e que um mesmo grupo de especialistas realiza as marcações na base de dados, deve-se levar em conta que o desempenho pode ter sido polarizado ao conhecimento apenas da base de dados. Entretanto, não é possível corroborar esta hipótese uma vez que não foram realizados testes com novas marcações e bases de dados. Apenas o estudo de Wilson *et al.* (1999), menciona o uso do sistema desenvolvido em ambiente clínico. Não foi mencionado nenhum valor do desempenho atingindo neste período de utilização clínica mas é descrita uma queda no desempenho do método.

Ainda em relação à base de dados e conforme pode-se observar no Quadro 1, 50% dos estudos utiliza a frequência de 256 Hz como amostragem na aquisição dos registros de EEG. Apesar das diferentes amostragens, não houve menção se as metodologias ou os sistemas podiam ter sido utilizados com registros amostrados com frequências diferentes como, por exemplo, 128 ou 512 Hz. Além disso pode-se destacar três trabalhos que foram excluídos do estudo comparativo, mas que apresentaram um diferencial em relação à frequência de amostragem dos registros. Os trabalhos de Lodder, Askamp e Van Putten (2013) e Lodder e Van Putten (2014) mencionaram a realização da diminuição da frequência de amostragem de 256 Hz para 100 Hz e o trabalho de El-Gohary, McNames e Elsas (2008) indicou a realização de interpolação da frequência de amostragem de 200 Hz para 1000 Hz. Entretanto, não ficou evidente a razão da realização dessas operações de alteração do valor da frequência ou se elas permitiram que quaisquer registros de EEG poderiam ter sido utilizados.

A análise do desempenho da grande maioria (90%) dos estudos apresentados no Quadro 1 foi realizada avaliando os valores de sensibilidade e especificidade dos classificadores implementados. Em alguns casos a avaliação do desempenho também considerou o valor da concordância entre os especialistas colaboradores do estudo ou comparou os resultados obtidos em outros estudos. Neste último caso, a análise foi indireta em função das diferentes bases de dados utilizadas. Entretanto, em nenhum dos estudos analisados, nem mesmo aqueles avaliados ao longo da pesquisa bibliográfica realizada ao longo desse trabalho, foi definido ou utilizado um limite mínimo para o desempenho. Em outras palavras, a avaliação é realizada de forma quantitativa, porém não existe uma indicação clara do limite mínimo a ser usado para considerar um desempenho aceitável. Desta forma, apesar de 38% dos estudos terem apresentado valores de sensibilidade e especificidade superiores a 80%, não é possível afirmar se os resultados podem ser considerados bons no contexto de aplicação dos trabalhos.

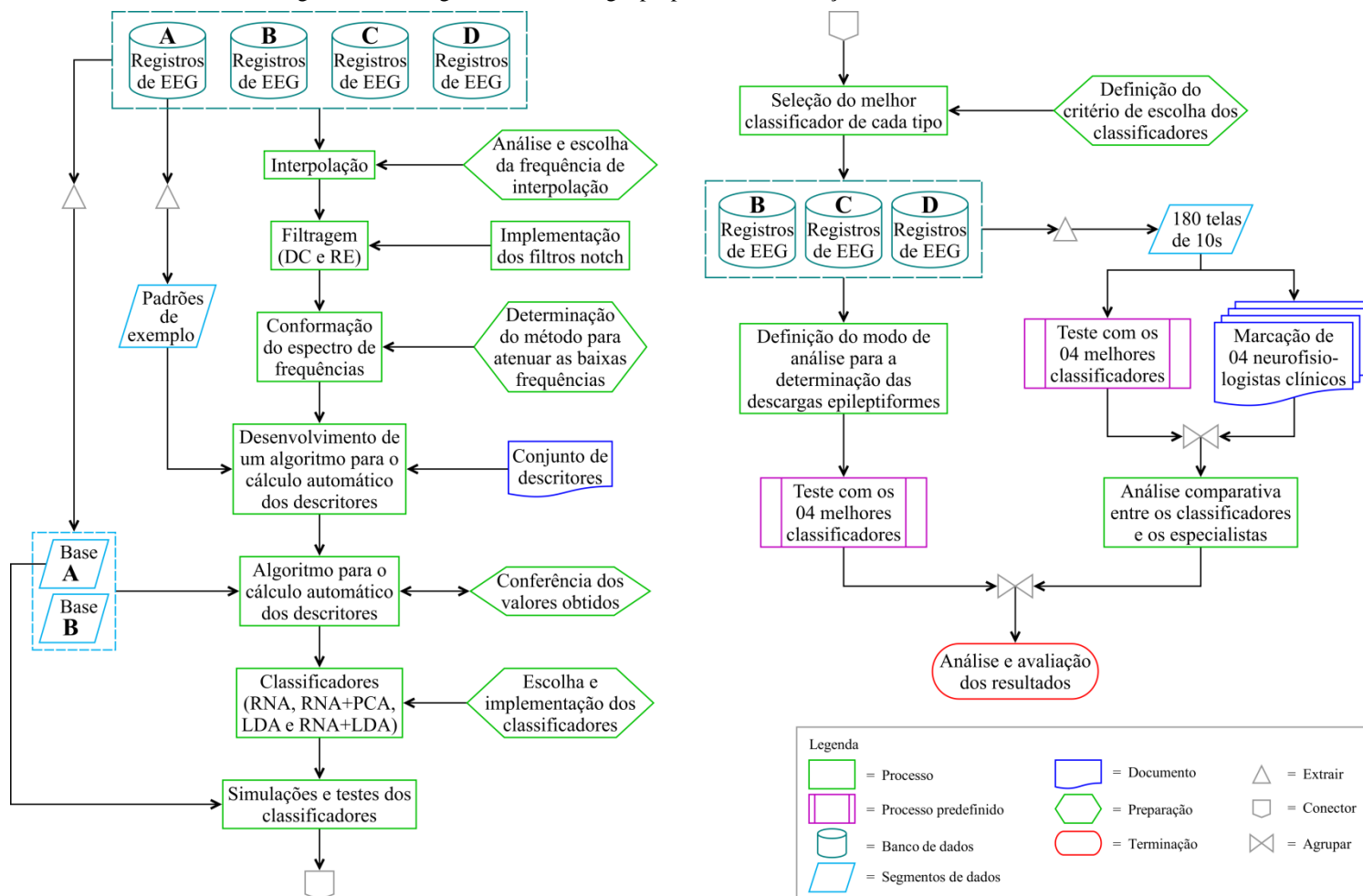
Ao considerar a metodologia utilizada para a detecção dos padrões epileptiformes, constatou-se que 62% dos estudos analisados utilizaram descritores morfológicos e contextuais. O tipo de classificador implementado é a rede neural artificial em pouco mais da metade dos estudos analisados, inclusive aqueles que foram excluídos do estudo comparativo. E dentre as redes neurais implementadas, a mais comum foi a *Perceptron* de múltiplas camadas.

A realização da pesquisa bibliográfica, complementada por este estudo comparativo, possibilitou a constatação de algumas limitações técnicas. Os aspectos referentes à escassez de marcações e bases de dados para a realização de testes, assim como a questão da frequência de amostragem e da análise de desempenho são abordados e considerados neste trabalho, quando da apresentação da metodologia e da discussão dos resultados.

3 METODOLOGIA

Este capítulo apresenta os dados, técnicas e ferramentas utilizadas neste trabalho, tanto para o cálculo do conjunto de descritores, quanto para a realização das simulações e testes utilizados como ferramenta de avaliação da metodologia proposta. A figura Figura 9 apresenta uma visão geral da metodologia proposta neste trabalho.

Figura 9 – Visão geral da metodologia proposta e as simulações realizadas neste trabalho.



Fonte: próprio autor.

3.1 BASE DE DADOS – SINAIS DE EEG

Uma característica comum nos métodos de detecção automática de padrões epileptiformes é a utilização de apenas uma base de dados, ou seja, um único grupo de registros de EEG para o desenvolvimento e o teste dos métodos. A utilização de dados de uma população pequena e restrita é uma limitação importante para a completa aceitação destes métodos dentro do ambiente clínico pois os padrões analisados, e os sinais nos quais estão presentes, apresentam uma grande variabilidade.

Teoricamente, a amostragem dos dados utilizados tanto no desenvolvimento como nos testes das metodologias, deveria refletir, ao máximo, a população do problema, ou seja, os indivíduos com epilepsia. No entanto, existe uma grande dificuldade para a aquisição de registros de EEG de longa duração pertencentes a diferentes grupos de indivíduos, bem como para a realização da revisão e da marcação meticulosa desses registros por parte de um grupo mínimo de especialistas.

Essas duas situações implicam o aparecimento de limitações presentes nos métodos desenvolvidos até o momento e que este trabalho pretende minimizar utilizando quatro bases de dados² diferentes. As bases a serem usadas são compostas por registros de EEG de diferentes indivíduos e foram adquiridas por dois equipamentos distintos, com frequência de amostragem, montagem de eletrodos e quantidade de canais distintos.

Uma das bases de dados, denominada base A, foi selecionada para a utilização no desenvolvimento da metodologia proposta, ou seja, nos testes e na implementação dos filtros, no desenvolvimento do algoritmo para o cálculo dos descritores, no treinamento e nos testes preliminares dos classificadores. As três bases restantes – base B, C e D – são usadas apenas para a realização dos testes da metodologia proposta. O objetivo desta divisão de dados foi possibilitar a verificação da generalidade do método proposto quanto ao tipo e à característica dos dados analisados.

A base utilizada no desenvolvimento da metodologia (base A) é composta por nove registros de indivíduos comprovadamente epiléticos. Os registros de EEG, totalizando 12 horas de sinais, foram adquiridos pelo sistema Harmonie® usando 24 canais (um registro) e

² A utilização dos registros de EEG contidos nas bases de dados foi aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos (CEPSH) da Universidade Federal de Santa Catarina (UFSC) conforme parecer consubstanciado número 144469 e 243001.

32 canais (oito registros), utilizando a frequência de amostragem de 100 Hz. E ainda, como parte do processo de aquisição, os sinais passaram por uma filtragem analógica para isolamento da faixa de frequência de 0,5 a 40 Hz.

A montagem dos eletrodos usados foi a bipolar do tipo zigomática-temporal, com 25 eletrodos nas posições Zy1, Zy2, Fp1, Fp2, F3, F4, F7, F8, F9, F10, CZ, C3, C4, T3, T4, T5, T6, T9, T10, P3, P4, P9, P10, O1, O2 do sistema 10/20 e dois eletrodos posicionados para aquisição do eletrooculograma (EOG).

As três bases utilizadas para os testes de metodologia fazem parte do banco de registros de EEG do Centro de Epilepsia de Santa Catarina (CEPESC). Este banco é composto por registros de EEG de 221 indivíduos monitorados continuamente por três dias para a avaliação da viabilidade cirúrgica. Pelo protocolo adotado no CEPESC, apenas seis horas de EEG de cada indivíduo são armazenadas, das quais três horas correspondem à primeira hora de sono de cada dia e três horas correspondem à primeira hora de vigília de cada um dos dias de monitoramento.

Do total de aproximadamente 1.200 registros de EEG, cada um com duração de uma hora, apenas 53 EEGs puderam ser utilizados, pois estes possuíam marcação digital do instante de tempo no qual ocorreram os padrões epileptiformes. Os demais registros do banco do CEPESC possuíam apenas a informação da quantidade de paroxismos.

Os 53 registros selecionados foram divididos nas três bases de dados de teste de acordo com as características de frequência de amostragem e montagem de eletrodos.

A base B possui 19 registros adquiridos no sistema 10-10, com frequência de amostragem de 512 Hz, usando 28 canais com montagem referencial de eletrodos e referência no eletrodo Pz.

A base C possui quatro registros com a mesma montagem e quantidade de eletrodos da Base B, porém os registros foram adquiridos com uma frequência de amostragem de 256 Hz.

Finalmente, a base D é composta de 30 registros adquiridos no sistema 10-20, com frequência de amostragem de 512 Hz e utilizando 27 canais, sendo 26 canais com sinal de EEG na montagem referencial no eletrodo Pz e um canal com o sinal de eletrocardiograma.

3.1.1 Seleção dos dados

A seleção dos segmentos de EEG utilizados no desenvolvimento da metodologia de detecção e nos testes iniciais foi realizada de forma aleatória, usando um algoritmo que selecionou quantidades iguais de segmentos de sinal com padrões epileptiformes e não-epileptiformes. A identificação dos padrões na base A foi realizada por três especialistas que analisaram todos os registros e a marcação final contida no EEG foi fruto do consenso entre esses especialistas.

Para as análises e as simulações relativas ao desenvolvimento da metodologia e aos testes iniciais, foram selecionados, dos registros da base A, um total de 802 segmentos de sinal, cada um com duração de três segundos. Uma vez selecionados, os segmentos foram divididos em dois conjuntos. O primeiro conjunto de sinais, denominados “padrões de exemplo”, é composto de oito segmentos de três segundos. Cada segmento desse subconjunto exemplifica um dos padrões típicos e de interesse, encontrados na base de dados:

- espícula com pico máximo positivo;
- espícula com pico máximo negativo;
- onda aguda com pico máximo positivo;
- onda aguda com pico máximo negativo;
- onda alfa;
- atividade normal de fundo;
- piscada;
- artefato ruidoso.

O segundo conjunto foi formado para a utilização nas simulações dos classificadores computacionais e foi subdividido em conjunto de treinamento, teste e validação. Os conjuntos de treinamento e validação possuem, cada um, 60 segmentos de EEG escolhidos manualmente por apresentarem características morfológicas bem definidas, enquanto que o conjunto de teste possui 200 segmentos escolhidos aleatoriamente. Todos os três conjuntos são balanceados, contendo quantidades iguais de padrões epileptiformes e não-epileptiformes.

Na base B foi possível selecionar 1.398 segmentos de sinal com três segundos de duração, com 657 segmentos contendo padrões não epileptiformes e 741 segmentos com padrões epileptiformes. Esses segmentos foram selecionados para serem usados nos testes de escolha dos classificadores para a realização das simulações com registros completos.

A base B, em conjunto com as bases C e D, foi utilizada nos testes com os registros de EEG completos, ou seja, a duração total dos registros é analisada e os canais de EEG adquiridos são analisados em paralelo. Neste caso, após consulta a neurofisiologistas clínicos, a identificação de um padrão epileptiforme ficou vinculada à presença de paroxismos em, no mínimo, três canais adjacentes.

A marcação dos padrões epileptiformes de cada um dos registros de EEG das bases B e C foi realizada por dois especialistas, diferentes dos que analisaram a base A, porém cada registro da base foi analisado apenas por um deles. Como a identificação dos padrões epileptiformes é feita de forma subjetiva, a marcação dos paroxismos em cada um dos registros da base D foi realizada por dois especialistas para que os resultados das simulações dos classificadores pudessem ser comparados com o parecer de mais de um especialista.

Além da falta de um “padrão-ouro” que possa ser utilizado como referência na análise do desempenho dos classificadores computacionais, os padrões epileptiformes duvidosos normalmente não são identificados nas marcações. A inserção dos padrões duvidosos como certeza negativa na comparação dos resultados de classificação gera falsos positivos que interferem no valor da especificidade e consequentemente mascaram o real desempenho (eficiência) do classificador.

A solução encontrada e aplicada no presente trabalho é uma adaptação do estudo de Wilson *et al.* (1996), no qual os especialistas realizaram a marcação de todos os padrões epileptiformes presentes nos registros de EEG utilizando um índice de percepção, no qual os padrões duvidosos recebem um peso menor na marcação.

Assim, para a avaliação final dos classificadores, foram selecionadas 180 telas de EEG com duração de 10 segundos. As telas foram escolhidas aleatoriamente a partir dos registros das bases de teste (B, C e D) e quatro especialistas foram instruídos a marcar todos os padrões epileptiformes presentes no EEG sem se preocuparem em indicar o padrão de máxima eletronegatividade, identificando as certezas com uma caneta verde e as dúvidas com uma caneta vermelha. O protocolo de marcação apresentado aos especialistas pode ser observado no Apêndice A.

3.2 CONJUNTO DE DESCRITORES

O conjunto de descritores escolhido neste trabalho para o desenvolvimento da metodologia de detecção dos padrões epileptiformes foi proposto por Pereira (2003). A justificativa dessa escolha é que o conjunto apresenta descritores que ilustram matematicamente a lista de características da atividade epileptiforme utilizada pelos especialistas, na prática clínica, para a identificação de um padrão epileptiforme. Este fato converge com o objetivo da metodologia proposta neste trabalho, que é o de reproduzir o procedimento de análise de um especialista.

O conjunto selecionado é composto por 43 descritores divididos entre aqueles existentes na literatura e outros propostos após uma extensa análise das características dos padrões apresentados pelos sinais de EEG.

Os elementos do conjunto extraídos da literatura são a amplitude máxima do padrão, a duração do padrão, o coeficiente de variação, o fator de crista e a entropia. A amplitude máxima e a duração do padrão são consideradas na sua definição clássica. O coeficiente de variação, utilizado para medir a regularidade do sinal, é a razão entre o desvio padrão e o valor da média do sinal.

O fator de crista é a diferença entre as amplitudes máxima e mínima dividida pelo desvio padrão (WEBBER *et al.*, 1994). Por sua vez, a entropia, relatada em diversos estudos, como os de Srinivasan, Eswaran e Sriraam (2007) e Naghsh-Nilchi e Aghashahi (2010), fornece um valor para a complexidade do sinal em análise.

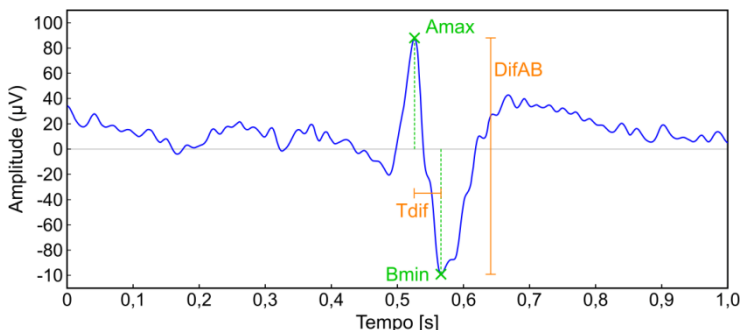
Os elementos do conjunto propostos por Pereira (2003) são descritores baseados nas características físicas ou morfológicas dos sinais em análise, levando em consideração as particularidades correspondentes ao tipo de montagem dos eletrodos utilizada e com o objetivo de realçar ainda mais os padrões epileptiformes dos outros tipos de padrões presentes nos registros.

As particularidades devido ao tipo de montagem referem-se ao fato de que, observando os registros adquiridos, pode-se notar que os padrões epileptiformes (espículas ou ondas agudas), devido à utilização da montagem bipolar, podem aparecer de quatro formas diferentes. Em outras palavras, em razão do tipo de montagem, as espículas e as ondas agudas podem aparecer com picos de amplitude tanto eletronegativos como eletropositivos.

Como ilustrado na Figura 10, os descritores relacionados à amplitude e a duração do padrão são:

- amplitude máxima (A_{max}) e mínima (B_{min});
- intervalo de tempo entre A_{max} e B_{min} (T_{dif});
- diferença de amplitude entre A_{max} e B_{min} ($DifAB$).

Figura 10 – Descritores relacionados à amplitude e duração do padrão.



Fonte: próprio autor.

Ainda em relação às amplitudes, determina-se um segmento de 300 ms centralizado no padrão contido nessa época e, analisando este segmento, identificam-se os seguintes descritores:

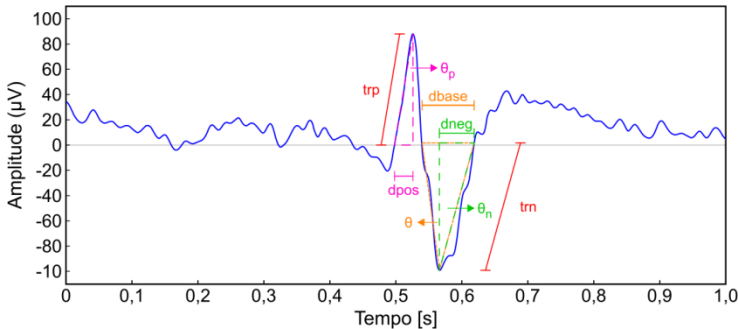
- amplitude máxima (A_{max_pts});
- amplitude mínima (B_{min_pts});
- diferença de amplitude ($DifAB_pts$) entre os extremos;
- intervalo de tempo (T_{dif_pts}) entre os extremos de amplitude.

Além dos descritores relacionados à amplitude e a duração, o conjunto contém elementos referentes ao ângulo do vértice dos picos do padrão (θ_p e θ_n). Esses vértices são formados pelas amplitudes máxima, mínima e os pontos de cruzamento por zero adjacentes no início e fim do padrão. Derivados destes ângulos, como ilustrado na Figura 11, identificam-se os seguintes descritores:

- cateto oposto dos ângulos – representando a base dos triângulos formados com as respectivas retas de inclinação e amplitudes máximas (d_{pos} e d_{neg});
- ângulo do vértice do pico do padrão analisado (θ);
- tangentes dos ângulos de picos (t_{gp} e t_{gn});

- retas de inclinação inicial e final do padrão (trp e trn);
- base do padrão (dbase).

Figura 11 – Descritores relacionados ao ângulo de vértice dos picos positivo e negativo de um padrão.

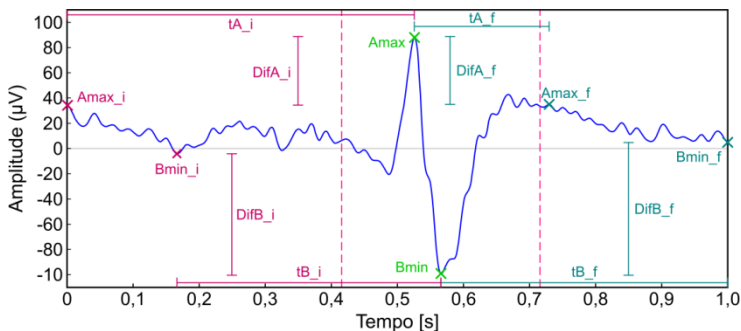


Fonte: próprio autor.

Pereira (2003) propôs também descritores baseados em relações de distâncias (intervalos de tempo) e diferenças de amplitude do sinal. Estes descritores são obtidos pela divisão do sinal contido na época sob análise em duas regiões (inicial e final), a partir do ponto de máxima amplitude absoluta do sinal. Uma vez determinadas estas regiões, têm-se, como ilustrado na Figura 12, respectivamente, os seguintes descritores:

- amplitude máxima e mínima da região inicial e final (A_{max_i} , B_{min_i} , A_{max_f} , B_{min_f});
- diferença de amplitude e intervalo de tempo entre A_{max} e A_{max_i} ($DifA_i$, tA_i) e entre A_{max} e A_{max_f} ($DifA_f$, tA_f);
- diferença de amplitude e intervalo de tempo entre B_{min} e B_{min_i} ($DifB_i$, tB_i) e entre B_{min} e B_{min_f} ($DifB_f$, tB_f).

Figura 12 – Descritores utilizados para a diferenciação entre padrões epileptiformes e artefatos de alta frequência presentes nos sinais.



Fonte: próprio autor.

Completando o conjunto de descritores estão os elementos com base nos índices estatísticos clássicos de média, desvio padrão e variância. Todos os descritores referentes aos índices estatísticos, o coeficiente de variação, o fator de crista e a entropia são calculados para a época sob análise (um segundo do sinal de EEG) e para o segmento de 300 ms centrado no padrão. O conjunto completo de 43 descritores encontra-se descrito resumidamente no quadro apresentado no Apêndice B.

3.2.1 Processamento dos sinais de EEG

Durante as etapas iniciais de desenvolvimento do algoritmo para o cálculo automático dos descritores, constatou-se a necessidade da realização do processamento do EEG adicionalmente a aplicação dos filtros realizada durante a aquisição dos sinais. Este processamento adicional é necessário para eliminar alguns problemas que podem prejudicar o funcionamento do algoritmo e limitar sua aplicação. Os problemas identificados nos sinais foram:

- baixa frequência de amostragem;
- desvios da linha de base do sinal;
- ruído caracterizado pela frequência de 60 Hz;
- interferência da atividade de baixa frequência (≤ 5 Hz).

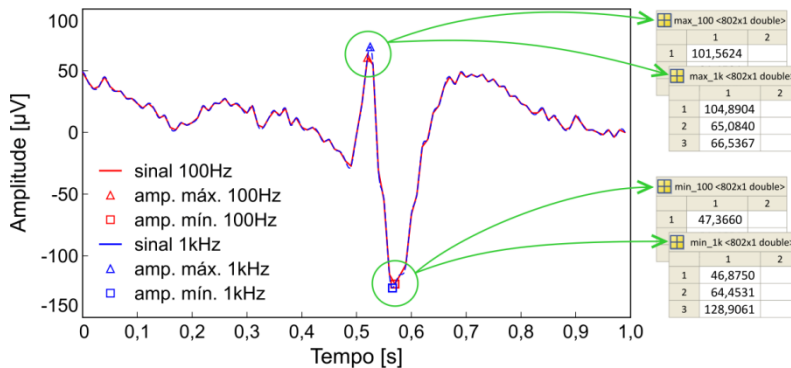
A baixa frequência de amostragem origina uma quantidade muito pequena de amostras para a representação dos padrões do sinal em

análise. Isso impossibilita a realização do cálculo de alguns descritores e diminui a precisão dos valores calculados. A solução encontrada para a baixa frequência de amostragem é a mesma encontrada para inserir na metodologia de detecção proposta, a independência da resolução temporal do sinal de EEG, ou seja, uma operação de interpolação.

A utilização da interpolação foi testada para garantir a ausência de possíveis prejuízos à morfologia do sinal ocasionados pela adição de amostras. Para o teste da interpolação foi utilizado o diagrama de Bland-Altman (ALTMAN; BLAND, 1983; BLAND; ALTMAN, 1986; MYLES; CUI, 2007; ZAKI *et al.*, 2012). Este diagrama é uma representação gráfica utilizada para a análise da concordância entre dois métodos diferentes de medição de um mesmo parâmetro. Esse diagrama apresenta a relação entre a diferença (eixo das ordenadas) e a média (eixo das abscissas) entre dois valores de um mesmo parâmetro obtidos de dois métodos distintos. Além disso, são apresentados os limites de concordância referentes ao intervalo de confiança (95%) dos valores dos parâmetros obtidos.

No contexto deste trabalho, como ilustra a Figura 13, o parâmetro analisado no diagrama é o valor da amplitude máxima e mínima dos segmentos de sinal obtido dos sinais original e interpolado.

Figura 13 – Primeira parte do teste da interpolação dos sinais de EEG, na qual são extraídas a amplitude máxima e mínima de todos os segmentos de sinal selecionados, tanto para o sinal original quanto para o interpolado.



Fonte: próprio autor.

Os valores de amplitude máxima e mínima foram escolhidos como parâmetros para a análise porque a grande maioria dos descritores

pertencentes ao conjunto selecionado utiliza de forma direta ou indireta, estes valores de amplitude para a obtenção dos valores dos demais descritores. Ou seja, qualquer imprecisão inserida nestes valores de amplitude seria propagada para os demais descritores.

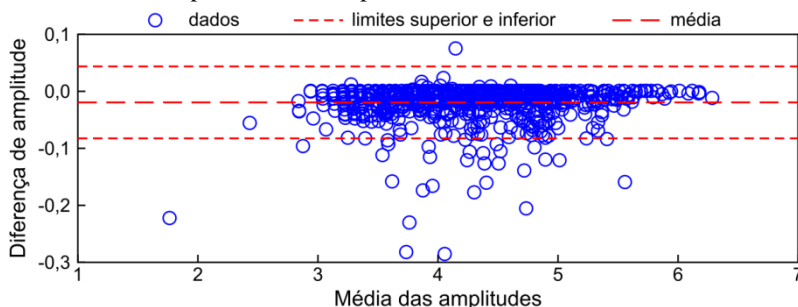
Estudos recentes como, por exemplo, o de Gotman (2010), Jacobs *et al.* (2012), Mari *et al.* (2012) e Zijlmans *et al.* (2012), apontaram que as chamadas oscilações de alta frequência (entre 80 e 500 Hz) podem fornecer informação clinicamente útil para a interpretação clássica do EEG na epilepsia. O interesse clínico em oscilações de alta frequência no EEG implica na necessidade de aumento da frequência de amostragem para, no mínimo, 1 kHz.

Considerando essa indicação de aumento da amostragem, a frequência de interpolação escolhida foi de 1 kHz. Com este valor a aplicação da interpolação na metodologia equivale a resolução temporal dos registros mais antigos e prepara a metodologia para ser utilizada nos novos registros de EEG.

Assim, o Diagrama de Bland-Altman foi utilizado para verificar a concordância entre o valor da amplitude máxima, e mínima, obtidas do sinal de 100 Hz e do sinal de 1 kHz, com a finalidade de determinar se o primeiro sinal poderia ser substituído pelo último sem que houvesse prejuízo no cálculo do conjunto de descritores morfológicos.

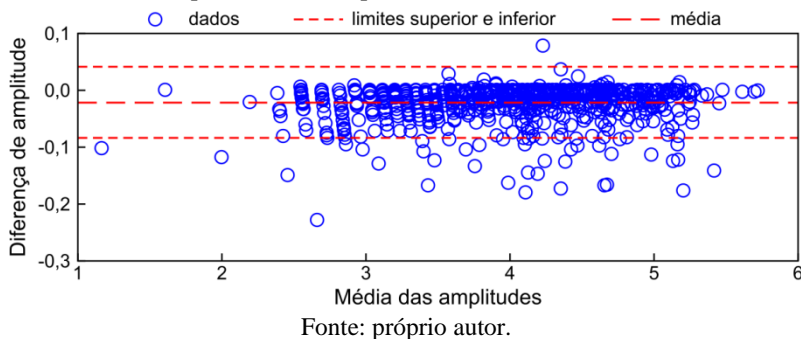
Para a geração do diagrama foi necessário realizar a adequação dos valores de amplitude, utilizando a transformação logarítmica, com a finalidade de desvincular a magnitude das diferenças de amplitude usadas no diagrama. Os diagramas gerados para análise das amplitudes podem ser observados na Figura 14 e Figura 15.

Figura 14 – Diagrama de Bland-Altman gerado para a análise da concordância entre os valores da amplitude mínima para o sinal original, amostrado a 100 Hz, e para o sinal interpolado, amostrado a 1 kHz.



Fonte: próprio autor.

Figura 15 – Diagrama de Bland-Altman gerado para a análise da concordância entre os valores da amplitude máxima para o sinal original, amostrado a 100 Hz, e para o sinal interpolado, amostrado a 1 kHz.



Analisando os diagramas, nota-se que os limites de concordância são relativamente pequenos e, conforme esperado, 95% das observações encontram-se dentro desses limites.

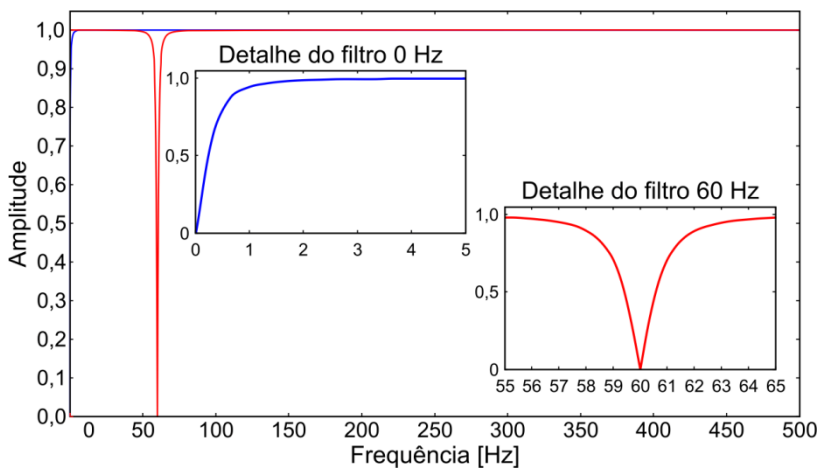
Além disso, foi possível extrair os fatores para a concordância entre as medidas, sendo que o fator para as medidas da amplitude máxima foi de 0,91903 e 1,0414, enquanto que, para a amplitude mínima, obteve-se 0,92042 e 1,0441.

Esses valores significam que, para a finalidade de substituir um sinal de baixa amostragem por um sinal interpolado com frequência de amostragem superior, as pequenas diferenças de amplitude apresentadas não resultariam em prejuízo para as análises posteriores. Em outras palavras, o sinal interpolado pode ser utilizado em substituição ao sinal amostrado a 100 Hz.

A interferência do ruído de 60 Hz e os desvios da linha de base do sinal, foram eliminados aplicando filtros *notch* de 4ª ordem. As respostas em frequência desses filtros podem ser observadas na Figura 16.

O filtro *notch* para linha de base possui frequência de corte de 0 Hz e largura de banda de 0,5 Hz e o filtro *notch* para ruído de 60 Hz possui frequência de corte de 60 Hz e largura de banda de 2 Hz para que as possíveis pequenas oscilações na frequência da rede elétrica também sejam incorporadas na faixa de rejeição do filtro, ainda que neste caso a atenuação seja mais leve.

Figura 16 – Resposta em frequência dos filtros *notch* utilizados para a remoção de ruído de 60 Hz e desvios da linha de base.



Fonte: próprio autor.

Conforme as orientações da *American Clinical Neurophysiology Society* (2008) e da *International Federation of Clinical Neurophysiology* (NUWER *et al.*, 1999) os filtros para a aquisição dos registros de EEG devem ser configurados para adquirir a maior quantidade de frequências possível ao invés de traçados limpos e a visualização do EEG registrado deve permitir a remoção das baixas frequências até 5 Hz.

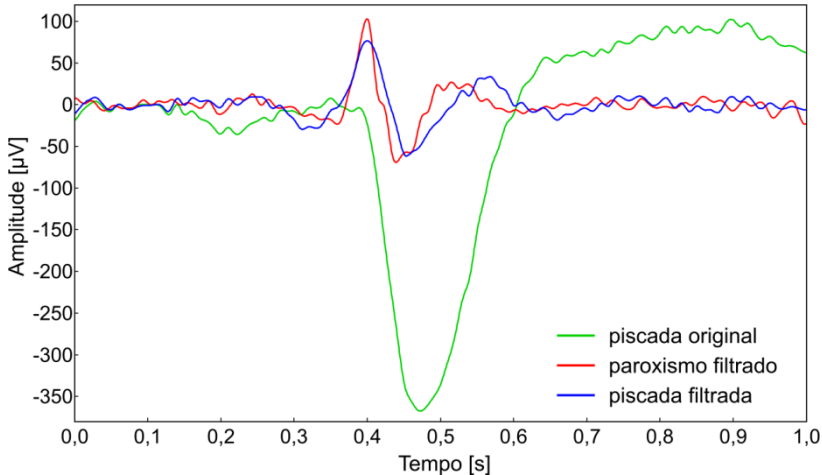
Alguns testes utilizando uma interface gráfica desenvolvida neste trabalho (Apêndice C) demonstraram que a eliminação das frequências até 5 Hz utilizando os métodos de filtragem – filtros passa-baixa e passa-banda, bancos de filtros Wavelet, entre outros – descritos na literatura, distorcem consideravelmente a morfologia de alguns padrões não-epileptiformes, em especial as piscadas.

Conforme ilustrado na Figura 17, a eliminação das baixas frequências até 5 Hz pode distorcer padrões como as piscadas de tal forma que o sinal filtrado se assemelha muito a um paroxismo.

Essa distorção morfológica, especialmente em uma metodologia baseada em morfologia do sinal, implica no aparecimento de padrões geradores de confusão no treinamento dos classificadores que tentarão, erroneamente, agrupar paroxismos e piscadas em uma mesma classe. Consequentemente, a confusão gerada durante o treinamento implica no

aparecimento de padrões classificados falsamente como paroxismos e diminuição da especificidade do classificador.

Figura 17 – Comportamento de um padrão epileptiforme (paroxismo) e um padrão não-epileptiforme (piscada) quando submetidos a um filtro passa alta de 5ª ordem com frequência de corte de 5 Hz.



Fonte: próprio autor.

Para resolver o problema da distorção morfológica causada pela remoção das frequências inferiores a 5 Hz seria necessário que esta banda do espectro sofresse uma atenuação leve de aproximadamente 4,5 dB. No entanto, essa diminuição da atenuação aplicada na banda de rejeição dos filtros é muito complexa, pois os filtros resultantes possuem pouca estabilidade e uma grande quantidade de coeficientes, deixando o processamento da filtragem muito pesado.

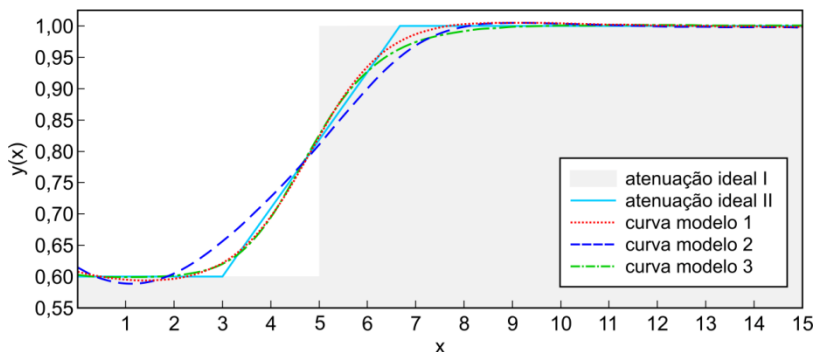
Como o conjunto de descritores implementado possui descritores desenvolvidos para diferenciar os padrões epileptiformes dos padrões mais lentos, tais como as piscadas e os ritmos alfa, e dos padrões ruidosos de alta frequência, não existia a necessidade de aplicar filtros para a redução da faixa de frequência analisada. A necessidade era apenas atenuar levemente as frequências até 5 Hz.

Assim, a solução encontrada e desenvolvida neste trabalho, foi a realização de uma adequação do espectro de frequência utilizando curvas, que foram denominadas de curvas de conformação.

As curvas de conformação foram obtidas em um processo com três etapas. A primeira consistiu na criação de uma curva degrau que descrevesse a situação ideal de atenuação do sinal, isto é, 40% de atenuação nas frequências até 5 Hz e nenhuma atenuação nas frequências superiores. A segunda etapa consistiu em suavizar esse degrau usando o método Savitzky-Golay (MADDEN, 1978; SAVITZKY; GOLAY, 1964; SCHAFER, 2011) e a terceira etapa foi obtenção das curvas de conformação utilizando o método dos Mínimos Quadrados Não Lineares e algoritmo *Trust-Region*.

Todas as curvas obtidas são representadas por funções racionais de 4ª ordem e as três melhores funções estão ilustradas na Figura 18. A curva escolhida para ser usada no processamento do sinal foi a curva modelo 2, que apresenta um erro médio quadrático de 0,001155 e o coeficiente de determinação de 0,998994.

Figura 18 – Função degrau (atenuação ideal I), a suavização do degrau (atenuação ideal II) utilizado e as curvas de conformação (curva modelo 1, 2 e 3) representadas pelas funções de 4ª ordem.



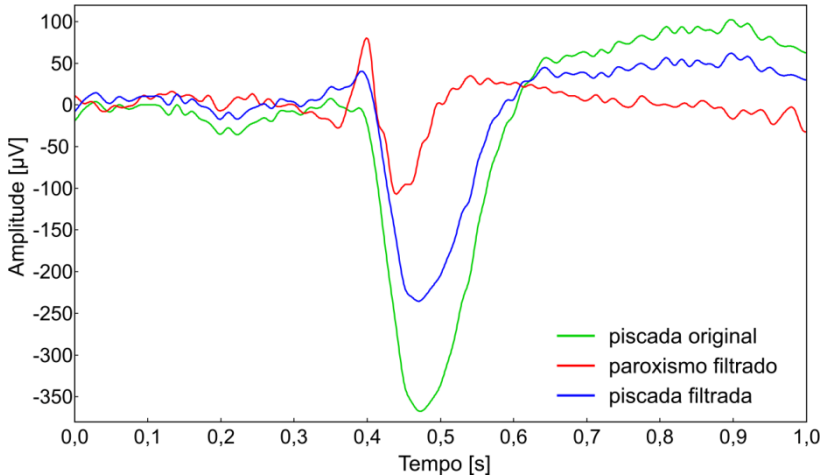
Fonte: próprio autor.

Finalmente, ao aplicar a solução da curva de conformação, nos mesmos sinais utilizados na Figura 17 é possível observar, conforme ilustrado na Figura 19, que a morfologia do paroxismo continua bem característica. E a conformação do espectro não causou a distorção morfológica da piscada de modo que a mesma se assemelha ao padrão epileptiforme.

Pode-se observar na Figura 19 que ao conformar o espectro, a piscada foi distorcida sem que a sua morfologia característica fosse alterada significativamente, ou seja, a piscada parece ter sido apenas

“achatada” mantendo sua característica eletrográfica de ser uma atividade lenta.

Figura 19 – Comportamento de um padrão epileptiforme (paroxismo) e um padrão não-epileptiforme (piscada) quando submetidos à curva de conformação de espectro para a atenuação de 4,5 dB nas frequências até 5 Hz.



3.2.2 Cálculo e análise dos descritores

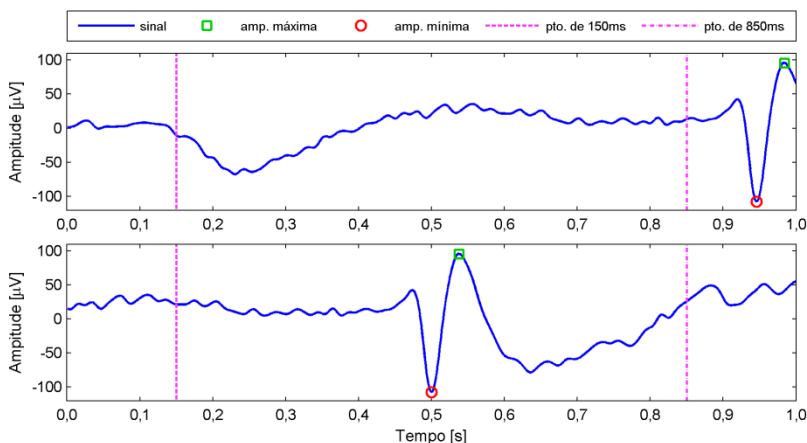
Uma vez escolhido o conjunto de descritores e determinado o processamento digital necessário para preparação dos sinais, foi desenvolvido um algoritmo para o cálculo automático dos descritores. O Apêndice D apresenta o fluxograma da visão geral do funcionamento do algoritmo desenvolvido.

O algoritmo inicia com a seleção dos segmentos de três segundos dos sinais que serão analisados. O próximo passo consistiu em realizar o processamento digital descrito no item 3.2.1 deste capítulo. Na sequência, o algoritmo inicia o cálculo dos descritores para cada época de sinal, que neste caso corresponde a um segmento de 1 segundo. A época de sinal foi selecionada da seguinte forma: dos três segundos do sinal coletado apenas o segundo central foi analisado (época) e os dois segundos adicionais funcionaram como margem de segurança para o tratamento das bordas realizado.

Os primeiros cálculos realizados pelo algoritmo são as amplitudes máxima (Amax) e mínima (Bmin) da época de sinal. Este dois valores devem ocorrer pela necessidade dos cálculos realizados para os demais descritores, entre os pontos de 150 e 850 ms da época analisada. Se as amplitudes ocorrerem fora desta faixa, como ilustra a Figura 20, é necessária a aplicação de um tratamento de bordas.

O tratamento das bordas é o simples deslocamento da época em alguns milissegundos com a finalidade de centralizar as amplitudes máximas e mínimas do padrão contidas na época. Caso este tratamento seja realizado, o algoritmo calcula novamente os valores de Amax e Bmin, e segue calculando os demais descritores do conjunto.

Figura 20 – Exemplo da aplicação do tratamento das bordas.



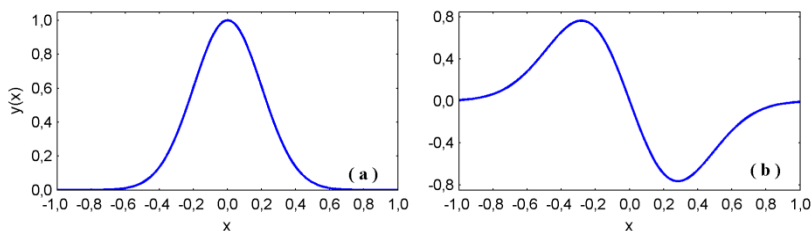
Fonte: próprio autor

Para a determinação do segmento de 300ms centralizado no padrão, e conseqüente o cálculo dos descritores referentes a ele, o pico de maior amplitude absoluta é utilizado como ponto central do segmento. Isto é, a partir do módulo do pico com maior valor de amplitude são contados 150ms para cada lado.

Considerando como referência o conjunto de descritores desenvolvido por Pereira (2003), a duração do padrão é obtida calculando a diferença entre o primeiro ponto de cruzamento pelo eixo das abscissas adjacente à Amax e Bmin no sentido do início e final do sinal, conforme a ordem em que as amplitudes aparecem.

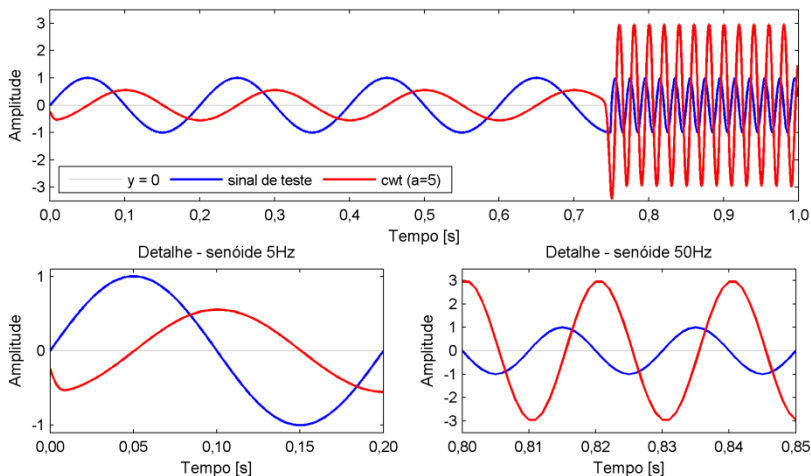
No entanto, como este cálculo apresenta dificuldades devido a característica bipolar dos sinais, foi necessário a utilização da Transformada Wavelet Contínua (*Continuous Wavelet Transform – CWT*) para possibilitar o cálculo da duração do padrão tanto para espículas e ondas agudas isoladas quanto para os complexos *spike-and-slow-wave* e *sharp-and-slow-wave*. A função *wavelet* escolhida para a implementação no algoritmo foi a 1ª derivada da Função Gaussiana com escala 5. A Figura 21 apresenta a forma de onda dessa função.

Figura 21 – Função Gaussiana (a) e função wavelet (b), 1ª derivada da Função Gaussiana utilizada para o cálculo dos descritores.



Fonte: próprio autor

Figura 22 – Função wavelet, 1ª derivada da função Gaussiana, aplicada a um sinal senoidal de frequências 5 e 50 Hz.



Fonte: próprio autor

Esta escolha foi baseada nos estudos de Sahambi, Tandon e Bhatt (1997), Garcia (2005) e Palodeto (2006). Esses estudos mostraram que esta função fornece informações singulares a respeito do sinal. Como pode ser observado na Figura 22, os pontos máximos e mínimos da CWT correspondem a pontos de cruzamento por zero do sinal e os cruzamentos da CWT por zero podem indicar os pontos de inflexão do sinal.

Para o cálculo do próximo grupo de descritores a época de sinal é dividida, a partir da duração do padrão, em duas regiões: inicial e final. A determinação destas duas regiões foi realizada de modo que o padrão analisado não pertencesse a nenhuma delas, ou seja, a região inicial é aquela entre o início da época e o início do padrão e a região final se estende do final do padrão ao final da época.

O cálculo do último grupo de descritores, relacionados a índices estatísticos e a entropia, é realizado com duas referências. São calculados a média, o desvio padrão, a variância, o fator de crista e a entropia do sinal tanto para o sinal presente na época em análise como para os 300 ms de sinal contidos no segmento determinado anteriormente.

Uma vez calculados todos os descritores, o algoritmo é finalizado com o armazenamento de todos os valores para a utilização dos mesmos em análises e simulações posteriores.

Como algoritmo foi desenvolvido com base nos padrões de exemplo, descritos no item 3.1.1 deste capítulo, foi realizada a conferência visual dos descritores calculados para 300 segmentos de sinal escolhidos, de forma aleatória, entre os conjuntos de treinamento, teste e validação. Após a realização desta conferência, constatou-se o funcionamento correto dos procedimentos instituídos, uma vez que houve concordância entre os valores de amplitude e duração calculados e as marcações dos registros de EEG.

Após o cálculo do conjunto de descritores foi realizada uma breve análise da correlação entre os descritores calculados a partir dos padrões de interesse (epileptiformes) e a partir dos demais padrões analisados. A análise da matriz de correlação foi realizada com a finalidade de excluir do conjunto os descritores que apresentassem correlação, positiva ou negativa, superior a 35%. Assim, os valores da correlação foram divididos em cinco faixas qualitativas:

- -0,05 e 0,05: valores ótimos – correlação com valor no intervalo [-0,05 ; 0,05];

- \pm (0,05 e 0,1): valores bons – correlação com valor no intervalo]-0,05 ; -0,10] ou]0,05 ; 0,10];
- \pm (0,10 e 0,20): valores regulares – correlação com valor no intervalo]-0,10 ; -0,20] ou]0,10 ; 0,20];
- \pm (0,20 e 0,30): valores altos – correlação com valor no intervalo]-0,20 ; -0,30] ou]0,20 ; 0,30];
- \pm (0,30 e 0,35): valores críticos – correlação com valor no intervalo]-0,30 ; -0,35[ou]0,30 ; 0,35[;
- \pm (0,35 e 1,00): valores excluídos – correlação com valor no intervalo [-0,35 ; -1,00] ou [0,35 ; 1,00].

3.3 CLASSIFICADORES

Uma parte importante das metodologias de reconhecimento de padrões epileptiformes em sinais de EEG é a etapa de classificação dos padrões. Neste trabalho, com o objetivo de verificar a generalidade da metodologia proposta, utilizou-se alguns classificadores para analisar a existência da equivalência dos resultados obtidos da classificação.

Para as simulações necessárias para a análise da metodologia proposta foram desenvolvidos quatro conjuntos de classificadores, baseados principalmente em Redes Neurais Artificiais. Esses classificadores são:

- Redes Perceptron Multicamadas (RNA);
- Redes Perceptron Multicamadas combinadas com a Análise de Componentes Principais (RPCA);
- Funções Discriminantes (FD) obtidas pela Análise de Discriminantes Linear;
- Redes Perceptron Multicamadas combinadas com a Análise de Discriminantes Linear (RLDA).

Para todos os classificadores foram utilizados os mesmos conjuntos de padrões para desenvolvimento – treinamento e validação das redes neurais e modelagem das funções discriminantes – e teste de cada um dos classificadores.

Além disso, para cada tipo de classificador foi aplicada uma espécie de protocolo de simulação com a finalidade de garantir que o desenvolvimento dos classificadores fosse sempre igual, dentro das especificidades e as necessidades provenientes de cada um deles. Este protocolo de simulação foi implementado para que os pontos iniciais das

simulações e as variações nos parâmetros de configuração de um mesmo tipo de classificador fossem na medida do possível padronizados para cada simulação realizada.

3.3.1 Redes Neurais Artificiais

Para a implementação das Redes Perceptron Multicamadas foi desenvolvida uma interface gráfica, ilustrada no Apêndice E. Nessa interface todas as simulações para a implementação das redes foram realizadas e os parâmetros de configuração e os resultados dos testes puderam ser armazenados para as análises posteriores.

O treinamento das redes foi supervisionado e realizado utilizando o algoritmo de Retropropagação de Erro (*Error Backpropagation*) com atualização de pesos sinápticos a cada nova época do treinamento. Além disso, foi utilizada a validação cruzada para determinação do ponto ótimo de parada do treinamento da rede. O Apêndice F apresenta um fluxograma do processo de treinamento das redes.

As três camadas da rede foram configuradas com 43 neurônios de entrada correspondendo ao tamanho do conjunto de descritores. Apenas um neurônio de saída para facilitar as comparações dos resultados e um número variável de neurônios na camada oculta. A variação dos neurônios ocultos, em uma faixa entre 6 e 14 neurônios, foi utilizada para possibilitar a determinação da melhor configuração possível.

Os critérios adotados para a parada do treinamento das redes foi o número máximo de épocas de treinamento (100.000 épocas) e o máximo erro médio quadrático da época de treinamento (definido em 0,0001).

O restante dos parâmetros de configuração das redes – pesos sinápticos iniciais, taxa de aprendizado, funções de ativação e coeficiente de momento – foram variados de forma a obter a melhor configuração possível para as redes.

3.3.2 Redes Neurais combinadas com PCA

A utilização da Análise de Componentes Principais (PCA) em conjunto com as redes foi idealizada neste trabalho com a finalidade de reduzir a dimensionalidade dos classificadores sem que houvesse uma perda significativa na informação contida no conjunto de descritores. Sendo assim, as componentes geradas pela análise substituem os descritores como estímulo de entrada das redes e, sendo escolhidas

corretamente, a diminuição da entrada da rede não resulta em prejuízo significativo para o desempenho final da classificação.

Para a obtenção das componentes principais foram agrupados os dados dos conjuntos de treinamento e de validação da rede, que, por sua vez, foram divididos de acordo com o tipo de padrão de sinal que representam, e foi utilizada a matriz de correlação dos dados. Uma vez obtido um conjunto de componentes, como parte das simulações realizadas, também foi aplicada a rotação do tipo *Varimax*.

Os critérios de seleção das componentes principais, com ou sem rotação, são:

- análise dos autovalores – na qual são selecionadas todas as componentes com autovalor, no máximo, unitário;
- variância explicada – na qual é calculada a variância explicada acumulada para as componentes dispostas em ordem decrescente de autovalor, e selecionado o conjunto de componentes com variância acumulada de 95%, ou o mais próximo possível desse valor.

Finalizada a obtenção das componentes, as simulações realizadas com as redes neurais foram feitas com Perceptron Multicamadas, utilizando o mesmo algoritmo, interface de usuário, parâmetros de configuração e protocolo de simulação. As duas modificações realizadas foram o tipo de estímulo de entrada, componentes principais em substituição aos descritores, e a quantidade de neurônios na camada de entrada, que foi sempre correspondente à quantidade de componentes.

3.3.3 Análise de Discriminantes

Neste trabalho, a Análise de Discriminantes foi utilizada para criar funções discriminantes, derivadas do conjunto de descritores, para diferenciar os padrões do sinal de EEG. As simulações foram realizadas com o auxílio de um software de análise estatística (IBM® SPSS Statistics).

A obtenção das funções discriminantes foi feita utilizando a distância de Mahalanobis e adicionando os descritores gradualmente à função discriminante – que no software utilizado corresponde ao método *stepwise*. Este método consiste em adicionar e remover os descritores na função discriminante, gradualmente, de acordo com dois valores de referência definidos no início da análise. Esses valores são constantes

relacionadas ao resultado do Teste F utilizado para identificar a função que melhor se adapta ao conjunto de padrões analisados. O software utilizado permite a escolha entre utilizar o valor real (f_{in} e f_{out}) ou o valor da significância (p_{in} e p_{out}) do resultado do Teste F. Com o objetivo de simplificar as simulações, optou-se por utilizar o valor da significância do teste, uma vez que os valores possíveis para p_{in} e p_{out} variam apenas entre zero e um.

Tabela 1 – Valores de p_{in} e p_{out} utilizados para obtenção das funções.

p_{in}		p_{out}				
0,01	0,25	0,50	0,75	1,00		
0,05	0,10	0,25	0,50	0,55	0,75	1,00
0,08	0,10					
0,09	0,10					
0,10	0,20	0,25	0,50	0,75	1,00	
0,15	0,75					
0,25	0,45	0,50	0,75	0,85	0,95	1,00
0,45	0,60					
0,50	0,75	1,00				
0,75	0,95	1,00				
0,95	1,00					

Durante as simulações da Análise de Discriminantes os valores de p_{in} e p_{out} determinados pelo software ($p_{in} = 0,05$ e $p_{out} = 0,10$) foram alterados para que fosse possível a obtenção de funções com diferentes tamanhos, ou seja, constituídas por diferentes grupos de descritores. Os valores utilizados nas simulações podem ser observados na Tabela 1.

3.3.4 Redes Neurais combinadas com LDA

As simulações das redes neurais em combinação com a Análise de Discriminantes foram realizadas utilizando os mesmos algoritmos, métodos e configurações das simulações descritas nos itens 3.3.1 e 3.3.3 deste capítulo. A Análise de Discriminantes foi usada como um critério de seleção dos descritores que são utilizados como entrada das redes neurais.

Foram utilizadas as redes Perceptron Multicamadas com algoritmo de retropropagação do erro e atualização dos pesos sinápticos a cada nova época de treinamento. Foi utilizada, também, a mesma interface para as simulações, assim como os mesmos critérios de parada do treinamento das redes e a mesma lógica de variação dos parâmetros

iniciais de configuração a fim de identificar um conjunto de parâmetros ótimos.

A quantidade de neurônios na camada de saída permaneceu igual, isto é, um neurônio, mas o número de neurônios de entrada e ocultos variou conforme as funções discriminantes utilizadas. O número de neurônios ocultos é alterado de tal forma que a camada oculta tenha aproximadamente a metade da dimensão da camada de entrada da rede e a quantidade de neurônios de entrada corresponde diretamente ao número de descritores presentes em cada uma das funções discriminantes utilizadas.

As funções discriminantes são as mesmas obtidas como resultado das simulações descritas anteriormente, isto é, as funções obtidas usando a distância de Mahalanobis, o método *stepwise* e variando os valores de P_{in} e P_{out} .

3.4 ANÁLISE E SELEÇÃO DOS CLASSIFICADORES

A análise do desempenho dos classificadores é feita utilizando os índices extraídos da Tabela de Contingência e das taxas de sensibilidade, especificidade e eficiência. A descrição e equacionamento dessas métricas de análise estão descritos no item 2.4. As simulações para teste e seleção dos classificadores são realizadas utilizando uma interface gráfica (Apêndice G) desenvolvida neste trabalho para permitir:

- a seleção do conjunto de segmentos de sinal para análise;
- a seleção de qualquer um dos classificadores desenvolvidos;
- a visualização dos parâmetros relativos ao processamento do sinal aplicado e às características de configuração do classificador em teste;
- a visualização gráfica dos segmentos de sinal analisados e o respectivo resultado da classificação;
- realização da análise dos valores de *threshold* das redes neurais;
- realização da análise e do armazenamento dos resultados dos testes dos classificadores.

A análise dos classificadores é realizada de forma a obter valores de falsos negativos mínimos, mantendo valores semelhantes para a sensibilidade e a especificidade para atingir uma eficiência ótima.

Não foi encontrado na literatura, uma definição clara de qual é o valor mínimo aceitável para o desempenho de um classificador computacional utilizado para o reconhecimento automático de padrões epileptiformes. Ainda que se pudesse argumentar, de forma lógica, que quanto mais próximo de 100% for o desempenho, melhor é o classificador computacional, esta indicação é incompleta no contexto deste trabalho. Como não existe um “padrão ouro” absoluto para a identificação dos paroxismos devido à natureza subjetiva da análise, a avaliação desses classificadores é mais complexa. É necessário considerar que o classificador deve atingir a eficiência máxima levando em consideração que o valor de referência utilizado na análise não é uma unanimidade e que dificilmente será uma representação que se assemelhe a um consenso da maioria dos especialistas.

Desta forma, a seleção dos classificadores utilizados na análise comparativa entre as diferentes bases de dados é realizada de acordo com um limite mínimo de 85% de desempenho nas taxas de sensibilidade, especificidade e eficiência. Idealmente todas as taxas precisariam superar este limite para o classificador ser selecionado, entretanto, se isto não for possível, o limite de 85% é considerado apenas para o valor da eficiência.

Esse limite mínimo de desempenho foi inspirado nos valores da concordância entre os especialistas reportados por Wilson *et al.* (1996), Dümpelmann e Elger (1999), Brown *et al.* (2007), Halford *et al.* (2011), Schomer e Silva (2011), Barkmeier *et al.* (2012) e Halford *et al.* (2013). A literatura descreve que, nos melhores casos estudados, os especialistas concordam em 81% das marcações. Considerando que a metodologia ou sistema desenvolvido deve auxiliar o especialista na tarefa de revisão dos registros de EEG, este trabalho considera que o desempenho mínimo do classificador computacional deve ser maior ou igual a 85%.

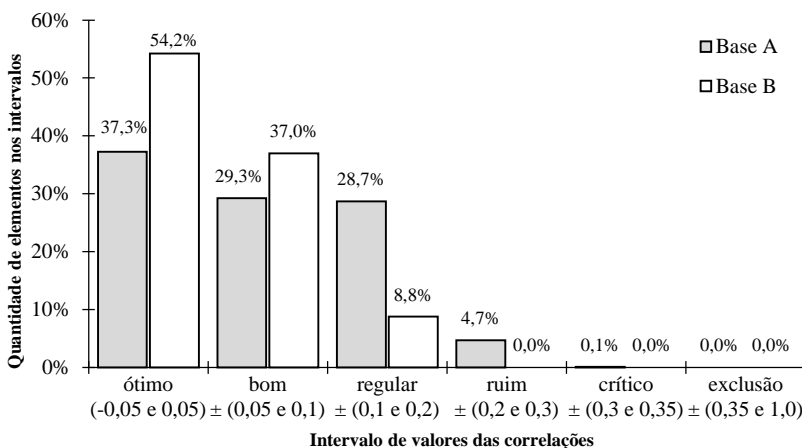
Para cada grupo de classificadores, o melhor deles é selecionado para a utilização na análise do desempenho da metodologia com as bases de dados de teste (bases B, C e D). Para analisar esse desempenho, conforme mencionado no item 3.1.1 deste capítulo, foi aplicada uma adaptação da metodologia desenvolvida e aplicada, respectivamente, nos estudos de Wilson *et al.* (1996) e Wilson *et al.* (1999), no qual a classificação computacional é comparada a um conjunto de marcações considerando a opinião de diferentes especialistas.

4 RESULTADOS E DISCUSSÃO

Neste capítulo, os resultados obtidos com as simulações e testes realizados são apresentados inicialmente e comentados brevemente para, no final, serem discutidos considerando todo o contexto das simulações realizadas.

Os resultados da análise da matriz de correlação para a base de dados de desenvolvimento (Base A) encontram-se detalhados na tabela apresentada no Apêndice H. Os resultados da análise da correlação para a Base A podem ser considerados satisfatórios. Com a finalidade de verificar se a correlação continuaria satisfatória entre descritores extraídos de sinais com uma montagem de eletrodos diferente a análise da matriz de correlação foi repetida na Base B. Os resultados obtidos com esta base, apresentados na Figura 23 podem ser considerados bons. Essa qualificação dos resultados é porque, como pode ser observado na Figura 23, respectivamente 95,3% e 100% dos valores de correlação para a Base A e B estão nas categorias ótimo, bom e regular.

Figura 23 – Análise qualitativa dos valores da matriz de correlação.



Fonte: próprio autor.

Assim, como a base dados de desenvolvimento (Base A) apresentou valores de correlação baixos e nenhum elemento com correlação na faixa indicada para exclusão, não foi necessário excluir

nenhum dos descritores do conjunto. Além disso, com a análise realizada na Base B, utilizada para os testes de classificadores, obteve-se valores de correlação ainda mais baixos.

Terminada a análise dos descritores, o próximo passo foi o de realizar as simulações com os classificadores. Após a realização de muitas simulações são apresentados os dez melhores resultados obtidos com cada grupo de classificadores.

O primeiro grupo analisado foi o das redes Perceptron Multicamadas. Pode-se observar na Tabela 2 que os valores de falsos positivos e negativos foram baixos, com nove falsos negativos (4,5%) apresentados pela rede rna_04 e oito falsos positivos (4%) apresentados pelas redes rna_06 e rna_09.

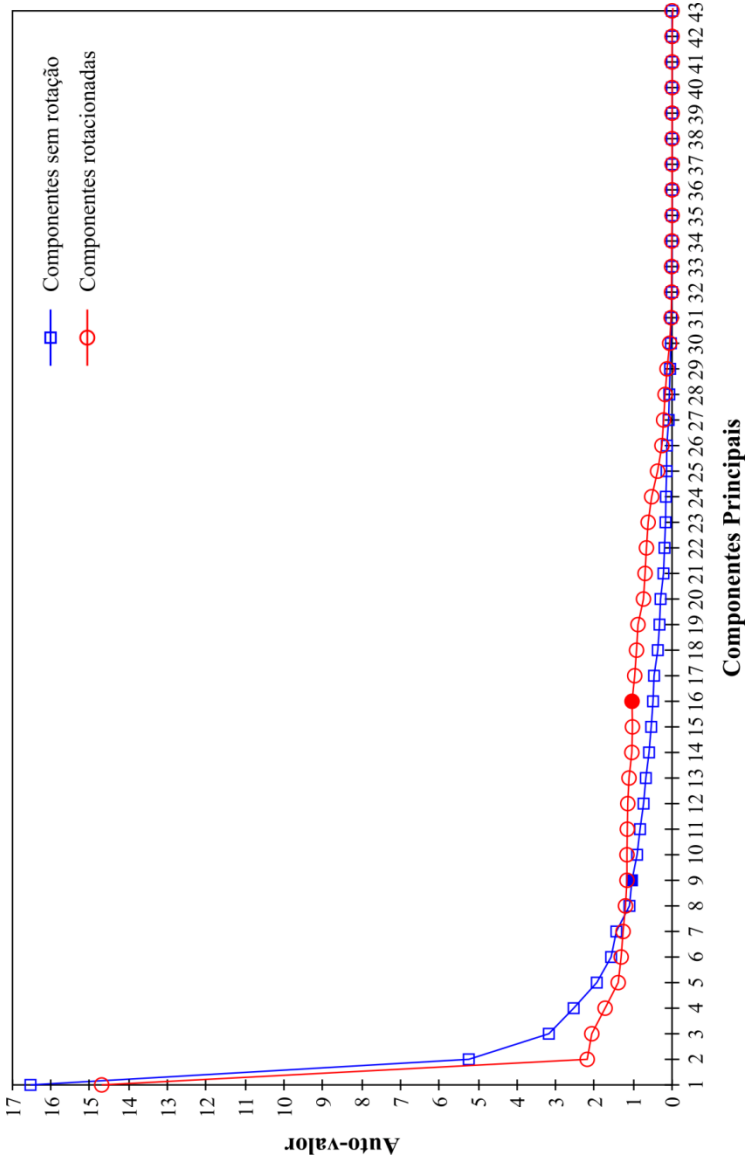
Os valores de sensibilidade e especificidade obtidos e apresentados na Tabela 2, foram todos iguais ou superiores a 84%. O melhor valor de sensibilidade foi 91% e a melhor especificidade foi de 92%. A menor eficiência obtida foi 87% e os melhores resultados foram das redes rna_08, rna_09 e rna_10, que obtiverem uma eficiência de aproximadamente 89%.

Tabela 2 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtidos nas simulações com as redes neurais.

Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
rna_01	89	85	15	11	89,00%	85,00%	87,00%
rna_02	85	90	10	15	85,00%	90,00%	87,50%
rna_03	89	86	14	11	89,00%	86,00%	87,50%
rna_04	91	84	16	9	91,00%	84,00%	87,50%
rna_05	89	87	13	11	89,00%	87,00%	88,00%
rna_06	84	92	8	16	84,00%	92,00%	88,00%
rna_07	88	88	12	12	88,00%	88,00%	88,00%
rna_08	88	89	11	12	88,00%	89,00%	88,50%
rna_09	85	92	8	15	85,00%	92,00%	88,50%
rna_10	87	90	10	13	87,00%	90,00%	88,50%

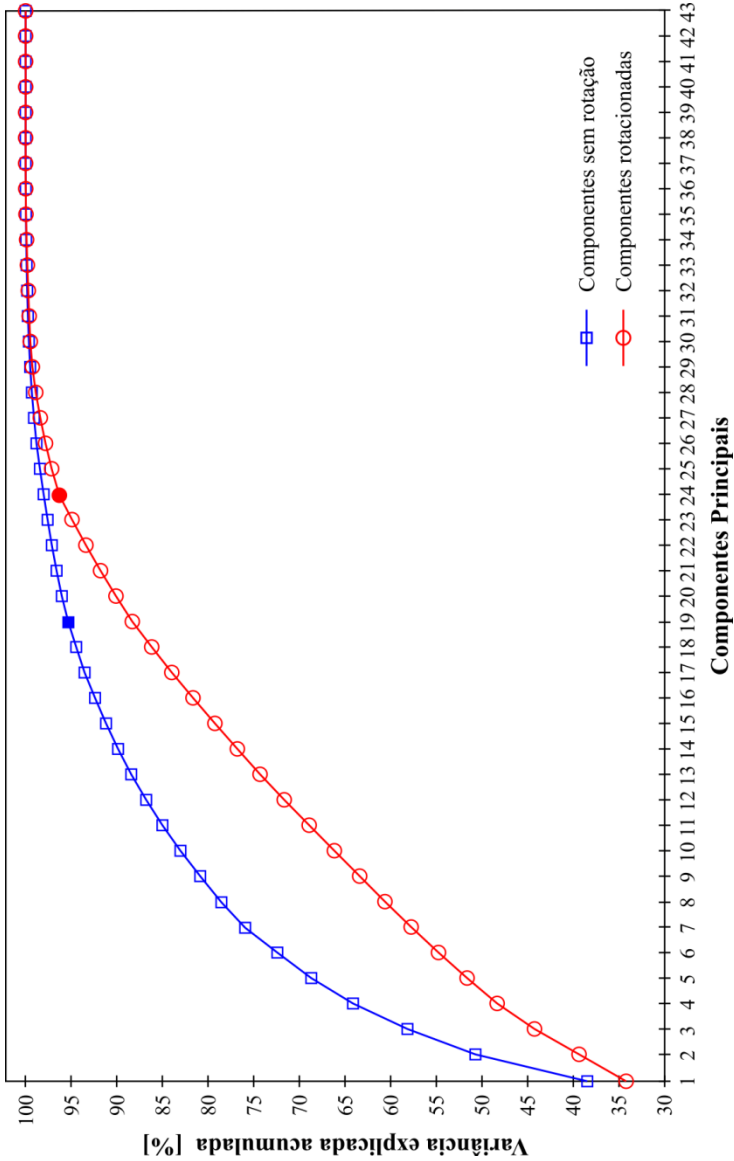
Utilizando a Análise de Componentes Principais, de acordo com o que foi proposto na metodologia, o conjunto de 43 descritores foi transformado em 43 componentes principais. A análise dessas componentes foi baseada nos dois pontos definidos no item 3.3.2 do capítulo anterior: autovalor e variância explicada acumulada.

Figura 24 – Resultados da Análise de Componentes Principais: autovalores para as componentes com e sem rotação, identificando com marcador preenchido a máxima componente a ser selecionada.



Fonte: próprio autor

Figura 25 – Resultados da Análise de Componentes Principais: variância explicada acumulada para as componentes com e sem rotação, identificando com marcador preenchido a máxima componente a ser selecionada.



Fonte: próprio autor

Analisando os resultados das componentes principais com base no autovalor, que podem ser observados na Figura 24, foram selecionadas as primeiras 9 e 16 componentes. Com base nos valores acumulados da variância explicada, ilustrados na Figura 25, foram selecionadas apenas as primeiras 19 e 24 componentes.

Os resultados obtidos com as componentes selecionadas aplicadas como estímulo de entrada de rede Perceptron Multicamadas são apresentados na Tabela 3. Pode-se observar nessa tabela que os valores dos falsos positivos e negativos também foram baixos, porém superiores aos atingidos pelas redes neurais com o conjunto de descritores como estímulo de entrada.

Todos os valores de sensibilidade, especificidade e eficiência obtidos foram superiores a 80%. Os melhores valores correspondentes a estas taxas foram 90, 89 e 88%, obtidos, respectivamente, pelas redes rna_07, rna_10 e rna_07. Considerando o desempenho médio da rede, pode ser considerado que a rede rpca_07 apresentou os melhores resultados dentre as redes com componentes principais como estímulo de entrada.

Tabela 3 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações das redes neurais utilizando componentes principais como estímulo de entrada das redes.

Classificador[§]	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
rpca_01	85	86	14	15	85,00%	86,00%	85,50%
rpca_02	86	85	15	14	86,00%	85,00%	85,50%
rpca_03	84	86	14	16	84,00%	86,00%	85,00%
rpca_04	90	80	20	10	90,00%	80,00%	85,00%
rpca_05	87	88	12	13	88,00 %	87,00 %	87,50 %
rpca_06	87	87	13	13	87,00 %	87,00 %	87,00 %
rpca_07	90	86	14	10	90,00%	86,00%	88,00%
rpca_08	87	87	13	13	87,00%	87,00%	87,00%
rpca_09	87	88	12	13	87,00%	88,00%	87,50%
rpca_10	86	89	11	14	86,00%	89,00%	87,50%

[§] As redes 1 a 4 possuem 9 componentes como estímulo de entrada, as redes 5 e 6, 7 e 8, e 9 e 10 possuem respectivamente 23, 19 e 16 componentes como estímulo de entrada.

Sequencialmente, foram realizados testes utilizando as Funções Discriminantes. Os resultados das simulações podem ser observados na Tabela 4 e Tabela 5. O tamanho das funções obtidas após as simulações,

isto é, a quantidade de descritores incluídos na função, variou entre 3 e 36. Os valores dos pesos das variáveis de cada função discriminante são apresentados no Apêndice I.

Tabela 4 – Tamanho das funções obtidas com a Análise de Discriminantes, da correspondente redução no tamanho do conjunto de descritores utilizados na classificação e apresentação dos índices de desempenho das funções (variância explicada e significância estatística).

Classificador	Tamanho[§]	Centroides	Variância explicada	Significância estatística
lda_01	3	± 1,0899	54,71%	7,00x10 ⁻²⁰
lda_02	7	± 1,2963	63,08%	1,29x10 ⁻²¹
lda_03	7	± 1,2963	62,94%	1,59x10 ⁻²¹
lda_04	7	± 1,2924	66,87%	1,12x10 ⁻²⁰
lda_05	9	± 1,3591	65,26%	1,34x10 ⁻²¹
lda_06	12	± 1,4087	71,15%	7,18x10 ⁻¹⁸
lda_07	22	± 1,5572	54,71%	7,09x10 ⁻²⁰
lda_08	24	± 1,5707	71,50%	4,31x10 ⁻¹⁷
lda_09	31	± 1,6076	72,44%	1,92x10 ⁻¹⁴
lda_10	36	± 1,6105	54,71%	7,09x10 ⁻²⁰

[§] O tamanho corresponde à quantidade de descritores incluídos na função discriminante.

A variância explicada pelas funções discriminantes apresentou valor médio de 63% e todas as funções obtidas apresentaram valores de significância estatística válidos com valores de p muito pequenos. O menor valor de variância foi obtido pela função lda_10 (54%).

Um aspecto que pode ser destacado é a existência de simetria nos valores dos centroides de todas as funções que, segundo a literatura, pode favorecer a classificação dos padrões apresentados. No entanto, a distribuição dos padrões em torno do centroide foi bastante dispersa, devido à alta variância apresentada pelos sinais de EEG, e consequentemente pelos valores dos descritores. Essa dispersão e a pequena distância entre os centroides das funções obtidas fez com que a simetria do valor dos centroides não influenciasse na classificação dos padrões.

Tabela 5 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações utilizando as funções discriminantes.

Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
lda_01	85	84	16	15	85,00%	84,00%	84,50%
lda_02	89	77	23	11	89,00%	77,00%	83,00%
lda_03	90	80	20	10	90,00%	80,00%	85,00%
lda_04	90	74	26	10	90,00%	74,00%	82,00%
lda_05	85	75	25	15	85,00%	75,00%	80,00%
lda_06	83	81	19	17	83,00%	81,00%	82,00%
lda_07	88	77	23	12	88,00%	77,00%	82,50%
lda_08	90	81	19	10	90,00%	81,00%	85,50%
lda_09	89	77	23	11	89,00%	77,00%	83,00%
lda_10	86	77	23	14	86,00%	77,00%	81,50%

Analisando os resultados obtidos na Tabela 5, pode-se observar que os valores de falsos positivos e negativos foram maiores do que os outros dois tipos de classificadores utilizados anteriormente. Os valores médios foram de 22 falsos positivos e 13 falsos negativos.

O melhor valor de falsos positivos foi obtido pela lda_01 e três funções – lda_03, lda_04 e lda_08 – apresentaram o menor valor de falsos negativos. O valor médio da sensibilidade, de 87,5%. Esse valor foi semelhante aos valores observados na Tabela 2 e Tabela 3, com destaque para as funções lda_03, lda_04 e lda_08, que apresentaram 90% de sensibilidade.

Por outro lado, o valor médio de 78,3% de especificidade foi inferior às médias de 88,3 e 85,5% observados nos classificadores anteriores. Como consequência da baixa especificidade, a eficiência média de 82,9% também foi mais baixa que os valores apresentados na Tabela 2 e Tabela 3. O melhor valor de eficiência, de 85,5%, foi obtido pela função lda_07. Esse pode ser considerado um valor bom quando comparado aos 88% atingidos na Tabela 2 e aos 88,5% apresentados na Tabela 3.

O último grupo de classificadores analisado neste trabalho foi uma combinação das funções discriminantes e das redes neurais. Os conjuntos de descritores de todas as 10 funções obtidas foram utilizados como estímulo de entrada e os resultados obtidos com as 10 melhores redes são apresentados na Tabela 6.

Tabela 6 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações das redes neurais utilizando as variáveis das funções discriminantes como estímulo de entrada das redes.

Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
rlda_01	85	85	15	15	85,00%	85,00%	85,00%
rlda_02	85	86	14	15	85,00%	86,00%	85,50%
rlda_03	85	81	19	15	85,00%	81,00%	83,00%
rlda_04	85	86	14	15	85,00%	86,00%	85,50%
rlda_05	82	90	10	18	82,00%	90,00%	86,00%
rlda_06	83	84	16	17	83,00%	84,00%	83,50%
rlda_07	87	85	15	13	87,00%	85,00%	86,00%
rlda_08	90	83	17	10	90,00%	83,00%	86,50%
rlda_09	88	88	12	12	88,00%	88,00%	88,00%
rlda_10	80	91	9	20	80,00%	91,00%	85,50%

Os valores dos falsos positivos e negativos obtidos pelas melhores redes foram relativamente baixos, com média de 14 falsos positivos (7%) e 15 falsos negativos (7,5%). Os melhores valores foram obtidos pela rede rlda_10, com nove falsos positivos (4,5%) e rlda_08 com dez falsos negativos (5%).

Os valores de sensibilidade e especificidade obtidos foram todos iguais ou maiores que 80%, com média de 85% de sensibilidade e 85,9% de especificidade. Podem ser destacadas as redes rlda_08 com 90% de sensibilidade, rlda_10 com 91% de especificidade e rlda_09 com eficiência de 88%.

Finalizada esta primeira etapa de testes, o próximo passo foi a escolha dos melhores classificadores. Foram selecionados rna_08, rpca_07, lda_08 e rlda_09.

As primeiras simulações com a base de dados Base B foram realizadas apenas com os segmentos de sinais de EEG – com 741 padrões epileptiformes e 657 padrões não epileptiformes. A finalidade dessas simulações foi verificar a adequação do critério de escolha definido.

Os resultados obtidos nessas simulações podem ser observados na Tabela 7. Nessa tabela são apresentados, além dos valores da tabela de contingência, sensibilidade, especificidade e eficiência, o valor do limiar (*threshold*) de classificação utilizado em cada classificador e a variação

(aumento ou diminuição) do valor da eficiência apresentado na classificação dos segmentos de sinais da Base B.

Tabela 7 – Resultado da comparação do desempenho dos classificadores em duas bases de dados diferentes (Base A e Base B).

Classificador	Base	Threshold	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência (taxa de acerto)	
rna_08	A	0,610	88	89	11	12	88,00%	89,00%	88,50%	–
	B	0,610	712	338	319	29	96,09%	51,45%	75,11%	↓ 15,1%
	B	0,924	630	559	98	111	85,02%	85,08%	85,05%	↓ 3,9%
rpca_07	A	0,560	90	86	14	10	90,00%	86,00%	88,00%	–
	B	0,560	700	301	356	41	94,47%	45,81%	71,60%	↓ 18,6%
	B	0,900	620	540	117	121	83,67%	82,19%	82,98%	↓ 5,7%
rlda_09	A	0,560	88	88	12	12	88,00%	88,00%	88,00%	–
	B	0,560	720	281	376	21	97,17%	42,77%	71,60%	↓ 18,6%
	B	0,850	624	553	104	117	84,21%	84,17%	84,19%	↓ 4,3%
lda_08	A	–	90	81	19	10	90,00%	81,00%	85,50%	–
	B	–	574	468	189	167	77,46%	71,23%	74,54%	↓ 12,8%

Analisando os resultados da Tabela 7 podem ser observadas e destacadas duas questões: o valor do *threshold* e a variação da eficiência.

Ao longo de todas as simulações com as redes neurais, a determinação do valor do *threshold* foi realizada com o intuito de atingir a máxima eficiência da rede. Ao realizar os testes dos classificadores com os descritores extraídos dos 1.398 segmentos de sinais da Base B o valor do *threshold* não foi alterado.

Como observado na Tabela 7, a utilização do mesmo valor resultou em diminuição da eficiência de 15,1% na rna_08 e 18,6% na rpca_07 e rlda_09. Considerando a média de todas as redes da Tabela 2, Tabela 3 e Tabela 6, houve uma diminuição de 7,22% na eficiência das redes “rpca”, de 16,11% nas redes “rna” e de 16,97% nas redes “rlda”.

Após observar os registros de EEG e os valores dos descritores de cada conjunto de dados (Base A e Base B) percebeu-se que a amplificação do sinal na aquisição da Base B foi superior àquela na aquisição da Base A. Como os descritores baseiam-se na análise morfológica do EEG, a maior amplificação resultou em descritores com os valores numéricos mais elevados na Base B. Consequentemente, em muitos segmentos analisados, os valores dos descritores para os padrões

não epileptiformes se assemelharam àqueles dos paroxismos e resultaram em erros na classificação dos padrões. Desta forma, na tentativa de recuperar o desempenho da rede, foi necessário aumentar o valor do *threshold* para que a máxima eficiência das redes neurais fosse atingida. O aumento desse valor resultou em uma variação da eficiência mais aceitável, implicando na diminuição média de 3,1% para “rna”, de 2,6% para “rpca” e de 4,0% para “rlda”.

Particularmente, como as redes neurais utilizando componentes principais como estímulo de entrada (“rpca”) inicialmente apresentaram uma menor redução da eficiência, em 60% das redes “rpca”, a alteração no valor do *threshold* resultou no aumento da eficiência.

O melhor desempenho das redes “rpca” pode ser atribuído as componentes principais corresponderem, por natureza, a uma combinação linear das variáveis que explicam (ou representam), cada uma, uma porcentagem da variância do conjunto de dados inicial. Em outras palavras, um conjunto reduzido de componentes principais funciona como uma generalização do conjunto de descritores, gerando um bom desempenho da classificação, porque a rede está “aprendendo” um conjunto de padrões genérico – o que a deixa mais robusta e preparada para a classificação de novos padrões diferentes daqueles utilizados em seu treinamento.

Por definição, a Análise de Componentes Principais pode ser aplicada como ferramenta para extrair características de um determinado grupo de variáveis que aparentemente não possuem correlação entre si (MYATT; JOHNSON, 2009). Neste contexto, a aplicação da PCA no conjunto de descritores resulta em uma combinação deles que pode estar realçando características que auxiliam a classificação dos padrões realizada pelas redes neurais.

A variação da eficiência encontrada nas simulações com a Base B possibilitou a percepção de que, para a identificação dos padrões epileptiformes em sinais de EEG, a melhor rede neural não é necessariamente aquela com o maior valor de eficiência. Também neste caso, a determinação da melhor rede com a utilização dos valores preditivos não foi a mais adequada porque as melhores redes, determinadas por este critério, também apresentaram uma redução significativa no desempenho da Base B.

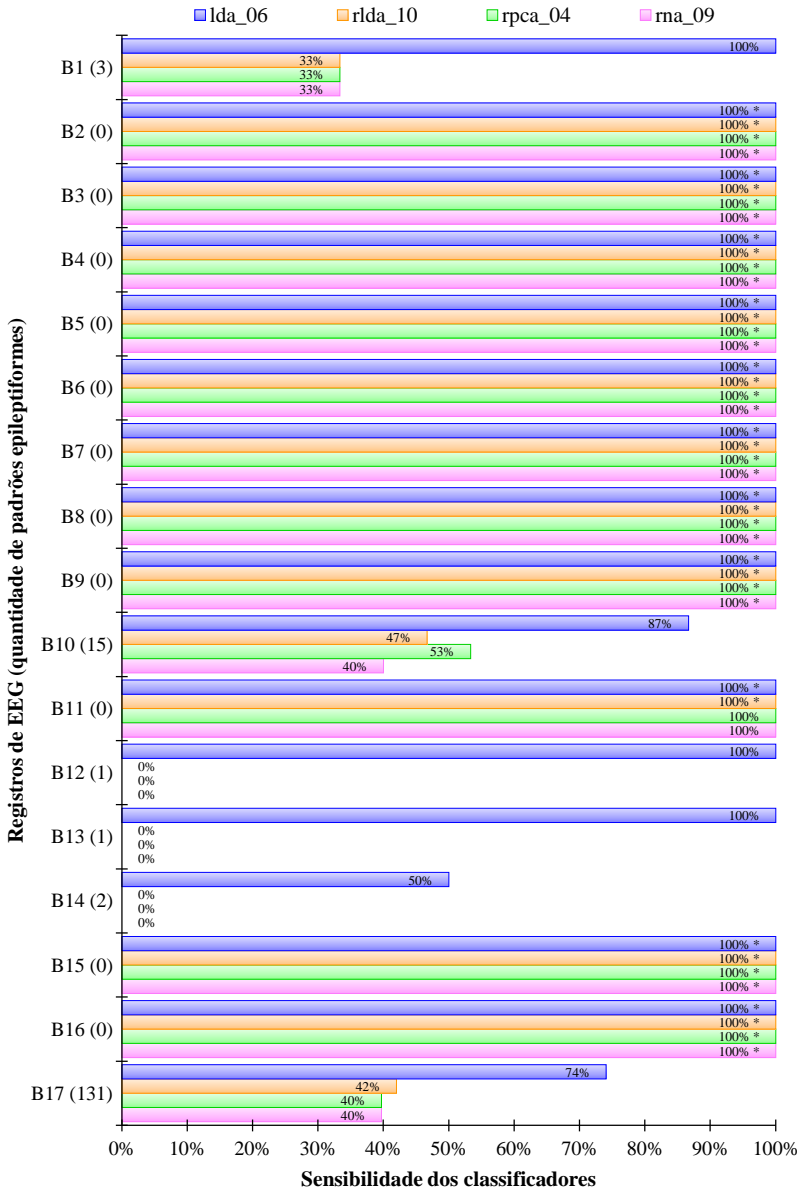
Após as simulações realizadas não foi possível identificar um padrão de correlação entre alguma métrica de desempenho e o melhor desempenho do classificador quando novos e diferentes padrões são apresentados a ele. Possivelmente, com um número maior de simulações

ou com a utilização de conjuntos de treinamento e testes muito maiores, a identificação de uma métrica mais robusta pode ser identificada.

Com o intuito de verificar a generalização e avaliar a metodologia proposta para identificar os padrões epileptiformes, a seleção dos quatro melhores classificadores foi variada e baseada em um equilíbrio entre valores de desempenho com as bases A e B. Desta forma, foram selecionados os seguintes classificadores: rna_09, rpca_04, lda_06 e rlda_10.

Após a seleção dos quatro classificadores foram realizadas as simulações com os registros das bases B e C. Nestas simulações utilizaram-se os registros de EEG completos e considerou-se o contexto de todos os canais adquiridos, ou seja, a propagação dos paroxismos para os canais adjacentes. Os resultados obtidos dessas simulações são apresentados na Figura 26 a Figura 28 e na Tabela 8.

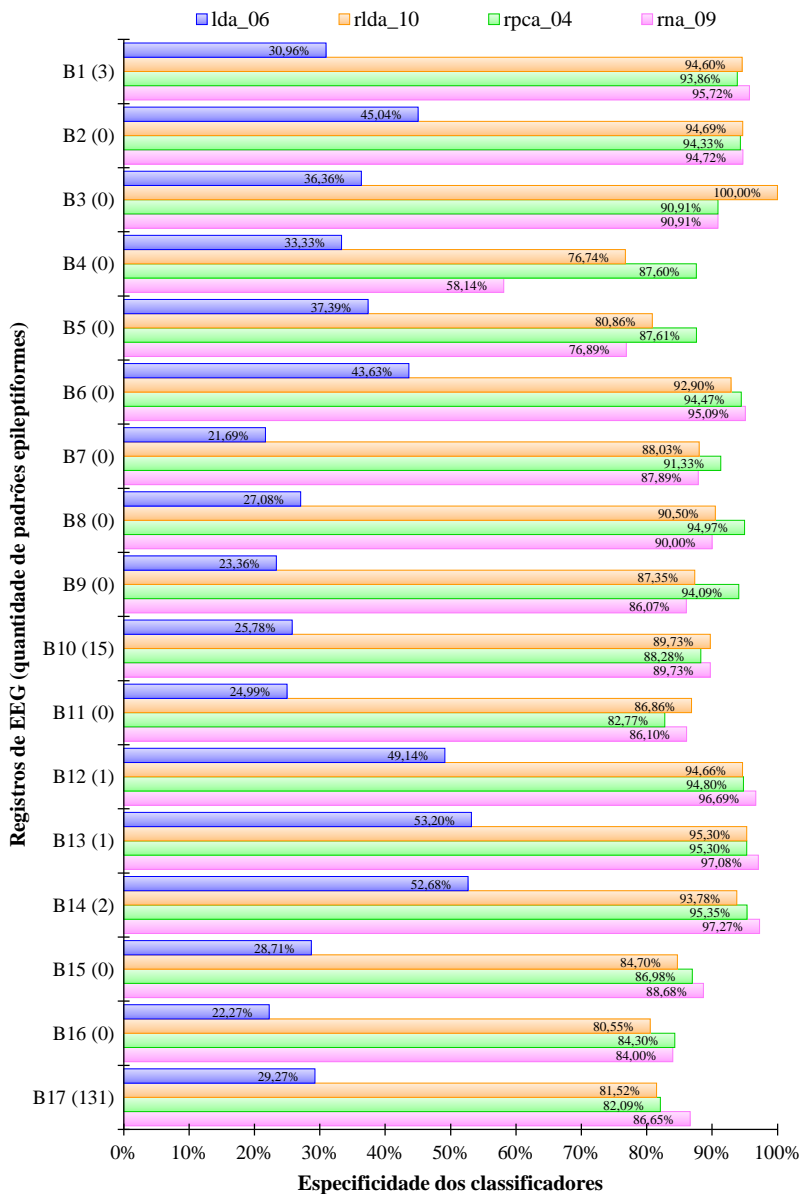
Figura 26 – Sensibilidade dos classificadores para os registros da Base B.



* Assume-se um valor de 100% de sensibilidade tanto o especialista quanto o classificador não identificaram paroxismos no registro de EEG analisado.

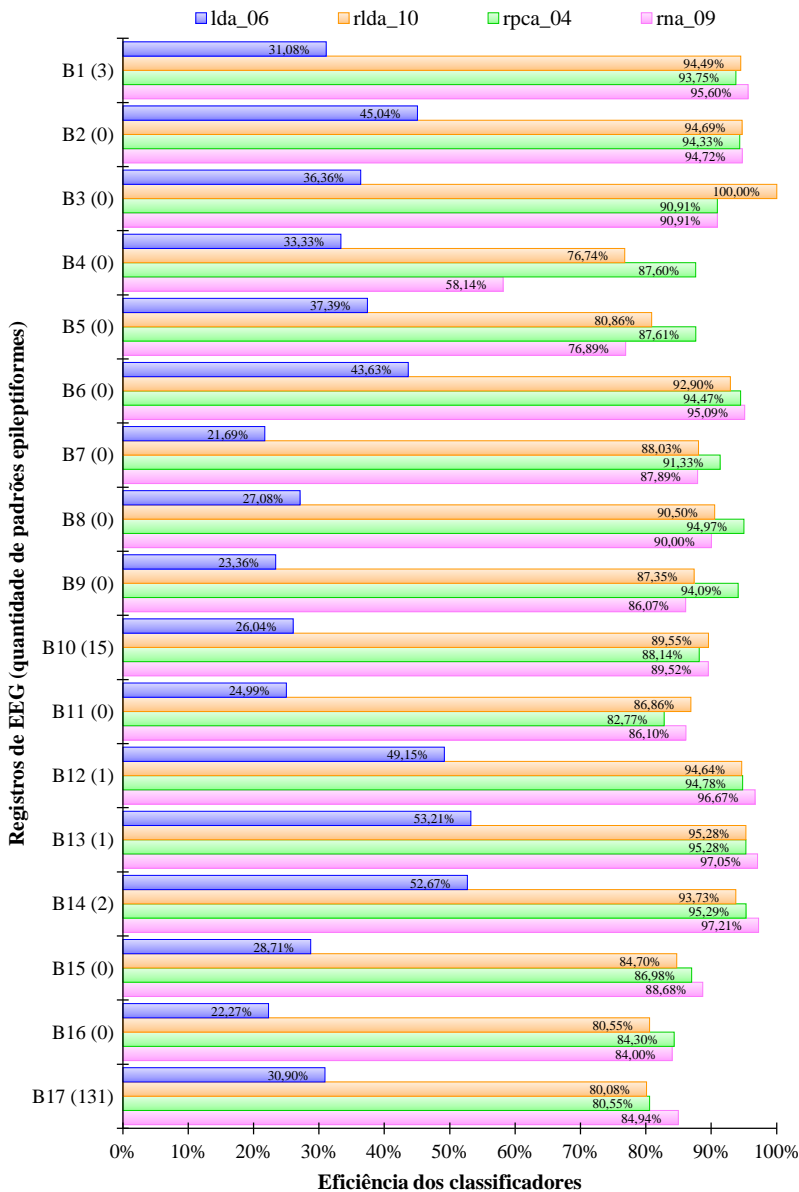
Fonte: próprio autor

Figura 27 – Valores de especificidade dos classificadores para os registros de EEG da Base B.



Fonte: próprio autor

Figura 28 – Eficiência dos classificadores para os registros de EEG da Base B.



Fonte: próprio autor

Na Figura 26, Figura 27 e Figura 28 são apresentados os resultados obtidos utilizando a Base B. O valor médio da eficiência para todos os 17 registros de EEG analisados pode ser considerado bom para as redes *rpca_04*, *rna_09* e *rlda_10* – que obtiveram valores de 90,42%, 88,21%, 88,88% de eficiência média, respectivamente – e ruim para *lda_06* que obteve apenas 34,52%. É importante destacar que as redes neurais obtiveram um desempenho “oposto” às funções discriminantes uma vez que enquanto as redes obtiveram altos valores de especificidade e baixa sensibilidade, a função discriminante apresentou sensibilidade média de 94,75% e apenas 34,40% de especificidade média.

Esse comportamento de oposição do desempenho das redes e da função discriminante também foi observado nas simulações com a Base C. Como pode ser observado na Tabela 8, as redes obtiveram sensibilidade e especificidade médias de 29,12 e 85,56%, respectivamente, enquanto que a *lda_06* obteve 80,14% de sensibilidade e 35,62% de especificidade. No entanto, os valores obtidos com essa base de dados foram menores que àqueles obtidos com a Base B. Considerando os quatro classificadores, houve uma diminuição de 46% no valor da sensibilidade, a especificidade foi 3% menor e a eficiência diminuiu aproximadamente 4%.

Tabela 8 – Valores da tabela de contingência, sensibilidade, especificidade e eficiência obtida nas simulações com a base de dados (Base C).

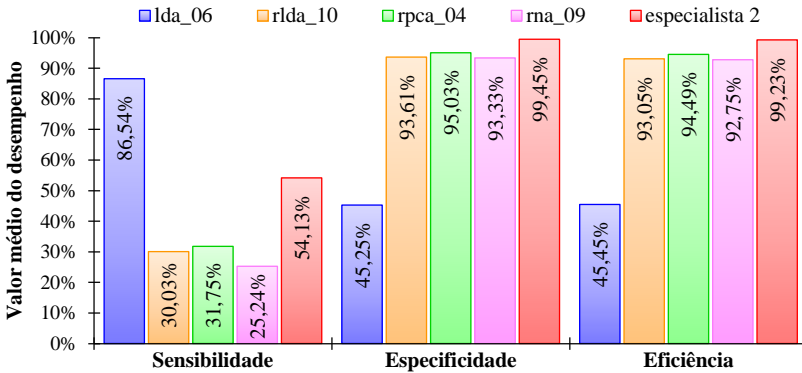
EEG	Classificador	Marcação [§]		VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
		EEP	NEP							
C1	<i>rpca_04</i>	0	300	0	271	29	0	100,00%*	90,33%	90,33%
	<i>rna_09</i>			0	234	66	0	100,00%*	78,00%	78,00%
	<i>rlda_10</i>			0	234	66	0	100,00%*	78,00%	78,00%
	<i>lda_06</i>			0	35	265	0	100,00%*	11,67%	11,67%
C2	<i>rpca_04</i>	23	796	4	766	53	19	17,39%	93,53%	91,45%
	<i>rna_09</i>			1	798	21	22	4,35%	97,44%	94,89%
	<i>rlda_10</i>			1	786	33	22	4,35%	95,97%	93,47%
	<i>lda_06</i>			17	398	421	6	73,91%	48,60%	49,29%
C3	<i>rpca_04</i>	3	3.593	0	2.650	946	3	0,00%	73,69%	73,63%
	<i>rna_09</i>			0	2.645	951	3	0,00%	73,55%	73,49%
	<i>rlda_10</i>			0	2.609	987	3	0,00%	72,55%	72,49%
	<i>lda_06</i>			3	1.487	2.109	0	100,00%	41,35%	41,40%
C4	<i>rpca_04</i>	30	7.139	2	6.544	625	28	6,67%	91,28%	90,93%
	<i>rna_09</i>			2	6.597	572	28	6,67%	92,02%	91,67%
	<i>rlda_10</i>			3	6.478	691	27	10,00%	90,36%	90,03%
	<i>lda_06</i>			14	2.929	4.240	16	46,67%	40,86%	40,88%

[§] Marcação refere-se à identificação dos paroxismos (EEP) e dos padrões não-epileptiformes (NEP).

* Assume-se um valor de 100% de sensibilidade tanto o especialista quanto o classificador não identificaram paroxismos no registro de EEG analisado.

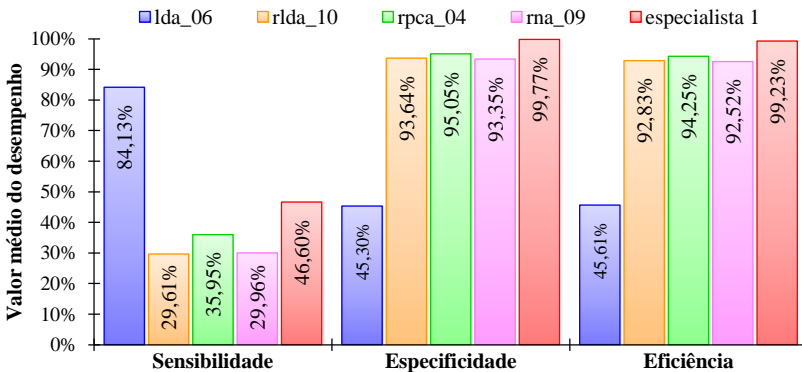
Os resultados obtidos após a utilização da Base D podem ser observados na Figura 29 e Figura 30, sendo que a primeira apresenta o desempenho médio dos 30 registros de EEG marcados pelo especialista 1 e a segunda apresenta o desempenho médio dos 30 registros de EEG marcados pelo especialista 2. Os resultados detalhados da Base D podem ser observados no Apêndice J e no Apêndice K

Figura 29 – Desempenho médio dos classificadores nas simulações com os registros da base D utilizando o especialista 1 como referência.



Fonte: próprio autor

Figura 30 – Desempenho médio dos classificadores nas simulações com os registros da base D utilizando o especialista 2 como referência.



Fonte: próprio autor

Os valores de eficiência obtidos nas simulações da Base D classificadas pelas três redes neurais podem ser considerados relativamente bons tanto para os registros de EEG marcados pelo especialista 1, com 93,43% de eficiência média, quanto para os registros marcados pelo especialista 2, com 93,20% de eficiência média.

Assim como o obtido nas bases B e C, a função discriminante apresentou resultados bem abaixo do que aqueles apresentados pelas redes neurais. A eficiência e a especificidade média foram de aproximadamente 45% para os registros marcados tanto pelo especialista 1 como pelo especialista 2. Esse resultado é considerado um resultado insuficiente. Novamente a função *lda_06* obteve melhores valores de sensibilidade do que as redes com média de 86,54 e 84,14% para as marcações do especialista 1 e 2, respectivamente, enquanto que as redes obtiveram 29 e 31,84% de sensibilidade média, resultado também insuficiente.

Levando em conta estes dados, ao comparar os resultados da classificação na Base D com os resultados da concordância entre os especialistas, apresentados no Apêndice L e Apêndice M, é possível considerar que a classificação computacional obteve resultados semelhantes. A média da sensibilidade e da especificidade da classificação dos registros realizada pelo especialista 2 (considerando o especialista 1 como referência) foi de 54,13% e 99,45%, respectivamente. Analisando o inverso, ou seja, utilizando o especialista 2 como referência, foram obtidos valores médios de 46,60% de sensibilidade e 99,77% de especificidade. Em ambos os casos a eficiência média da classificação foi de 99,23%. Isso representa uma concordância muito boa. Realizando uma comparação os valores médios considerando ambos os especialistas constatou-se que os valores da eficiência e da especificidade das redes neurais foram 5,96 e 5,63% menores do que os valores atingidos pelos especialistas e a sensibilidade foi 39,6% menor. A função *lda_06* atingiu aproximadamente 54% da especificidade e eficiência dos especialistas mas, por outro lado, obteve sensibilidade 69,4% maior.

É importante destacar um aspecto observado na análise dos resultados das simulações da Base D que corrobora observações descritas na literatura (HARNER, 2009; SCHOMER; SILVA, 2011; WILSON *et al.*, 1996; WILSON; EMERSON, 2002): a alta concordância entre os especialistas na definição dos verdadeiros negativos.

Analisando os resultados de especificidade obtidos, é possível comprovar que, apesar de existirem diferenças nas marcações entre os

especialistas e os padrões epileptiformes duvidosos não terem sido marcados, ambos os especialistas concordaram muito (99%) na identificação dos padrões não-epileptiformes.

Além disso, apesar das diferenças de desempenho, percebeu-se que a tendência de comportamento das redes neurais desenvolvidas era semelhante ao dos especialistas. Isto é, os valores de sensibilidade são muito menores que os valores de especificidade.

Finalmente, com as simulações usando as 180 telas de avaliação foram obtidos resultados de sensibilidade, especificidade e concordância na marcação dos padrões epileptiformes apresentados respectivamente na Tabela 9, na Tabela 10 e na Tabela 11. Como pode ser observado nas tabelas, os resultados foram obtidos utilizando oito classificadores – quatro especialistas (E1, E2, E3 e E4), três redes neurais (rna_09, rpca_04 e rlda_10) e uma função discriminante (lda_06) – tanto como referência (Ref.) quanto como classificador (Class.).

Tabela 9 – Sensibilidade (probabilidade de marcação de um paroxismo) entre os quatro especialistas e os quatro classificadores.

Class. Ref.	E1	E2	E3	E4	rpca_04	rna_09	rlda_10	lda_06
E1	1,000	1,004	0,728	0,900	0,754	0,142	0,201	0,844
E2	0,276	1,000	0,318	0,434	0,462	0,092	0,126	0,735
E3	0,642	1,019	1,000	0,814	0,747	0,182	0,226	0,901
E4	0,597	1,047	0,612	1,000	0,666	0,128	0,175	0,866
rpca_04	0,247	0,552	0,278	0,329	1,000	0,212	0,240	0,860
rna_09	0,196	0,462	0,285	0,266	0,891	1,000	0,739	0,989
rlda_10	0,230	0,523	0,293	0,302	0,838	0,613	1,000	1,000
lda_06	0,103	0,326	0,125	0,159	0,320	0,087	0,107	1,000

Tabela 10 – Especificidade (probabilidade de marcação de um padrão não-epileptiforme) entre os quatro especialistas e os quatro classificadores.

Class. Ref.	E1	E2	E3	E4	rpca_04	rna_09	rlda_10	lda_06
E1	1,000	0,789	0,956	0,931	0,827	0,956	0,950	0,441
E2	0,997	1,000	0,990	0,989	0,869	0,963	0,961	0,473
E3	0,980	0,803	1,000	0,938	0,834	0,960	0,953	0,446
E4	0,992	0,833	0,975	1,000	0,840	0,958	0,952	0,452
rpca_04	0,975	0,810	0,959	0,929	1,000	0,993	0,987	0,499
rna_09	0,932	0,742	0,913	0,876	0,821	1,000	0,975	0,444
rlda_10	0,936	0,749	0,917	0,881	0,826	0,986	1,000	0,449
lda_06	0,968	0,821	0,955	0,930	0,929	0,999	1,000	1,000

Tabela 11 – Concordância da marcação, dos padrões epileptiformes, entre os quatro especialistas e os quatro classificadores.

Class.	E1	E2	E3	E4	rpca_04	rna_09	rlda_10	lda_06
Ref.								
E1	1,000	0,526	0,665	0,746	0,424	0,160	0,205	0,289
E2	0,526	1,000	0,544	0,652	0,507	0,206	0,255	0,485
E3	0,665	0,544	1,000	0,684	0,436	0,235	0,252	0,322
E4	0,746	0,652	0,684	1,000	0,462	0,185	0,227	0,352
rpca_04	0,424	0,507	0,436	0,462	1,000	0,435	0,449	0,525
rna_09	0,160	0,206	0,235	0,185	0,435	1,000	0,673	0,294
rlda_10	0,205	0,255	0,252	0,227	0,449	0,673	1,000	0,327
lda_06	0,289	0,485	0,322	0,352	0,525	0,294	0,327	1,000
Média[§]	0,646	0,574	0,631	0,694	0,457	0,196	0,235	0,362

[§] A média refere-se ao valor médio considerando apenas a concordância com os especialistas.

Os resultados obtidos na Tabela 9 e na Tabela 10 são compatíveis com aqueles obtidos nas simulações com as bases de dados B, C e D. Os valores de sensibilidade atingidos pelas redes foram menores que os de especificidade, enquanto que a função discriminante apresentou um comportamento contrário – sensibilidade maior do que a especificidade.

Utilizando as marcações de quatro especialistas, que identificaram todos os possíveis paroxismos, incluindo os duvidosos, foi possível obter valores de desempenho mais precisos – uma vez que um dos problemas da avaliação de metodologias de identificação automática de padrões epileptiformes é a inexistência de um padrão-ouro e a baixa concordância entre os especialistas.

Os valores médios da sensibilidade para os quatro classificadores computacionais (rpca_04, rna_09, rlda_10 e lda_06) foram 0,657; 0,136; 0,182 e 0,837, respectivamente. A diferença de sensibilidade entre as redes neurais foi mais evidente e pôde-se observar uma menor diferença entre a rpca_04 e a lda_06. Esta última continuou apresentando o maior valor de sensibilidade entre os classificadores analisados.

A especificidade, que pôde ser calculada com maior precisão, apresentou média de 0,843 para a rede rpca_04, de 0,959 para rna_09, de 0,954 para rlda_10 e de 0,453 para lda_06. O comportamento da especificidade foi semelhante aos resultados anteriores, nos quais a função discriminante apresentou os menores valores. Entretanto foi observado que, ao contrário dos resultados das simulações anteriores, o valor da especificidade da rpca_04 foi menor do que o das outras duas redes (rna_09 e rlda_10).

Em função da baixa prevalência e da alta especificidade apresentadas nos resultados das bases B, C e D, o valor da concordância na marcação apresentou-se como um índice para avaliação do desempenho mais robusto, uma vez que foi obtido no reconhecimento à uma alta concordância dos especialistas na definição de um padrão epileptiforme (SCHOMER; SILVA, 2011; WILSON *et al.*, 1996). Pode-se observar na Tabela 11, a concordância entre os especialistas não foi superior a 0,75 e a média entre todos eles foi 0,636.

Analisando a concordância dos classificadores computacionais com a marcação dos especialistas, pôde-se verificar que os valores atingidos foram baixos, uma vez que eram esperados valores mais próximos ao desempenho dos especialistas.

A concordância média dos classificadores de 0,313 é 49,13% do valor atingido entre os especialistas (0,636). Entre os classificadores, a maior concordância, de 0,457, foi obtida pela rpca_04.

O desempenho abaixo do esperado é devido à baixa quantidade de segmentos de sinais de EEG utilizados no treinamento dos classificadores e, ainda, pelo fato das marcações utilizadas no treinamento não considerarem os paroxismos duvidosos, o que pode ter gerado falsos positivos.

Em ambos os casos, a solução depende da disponibilidade de um grande número de especialistas – quanto maior o número, maior é a qualidade do conjunto de marcações – para realizar exatamente a análise que há décadas tenta-se facilitar: a rigorosa identificação de todos os padrões epileptiformes presentes em um grande número de registros de EEG.

As marcações das 180 telas utilizadas para os testes finais podem ser um começo, porém é necessário um banco de marcações muito maior para que os classificadores possam ter uma qualidade de treinamento minimamente semelhante àquela dos especialistas – que chegam a analisar centenas de registros antes de serem considerados plenamente aptos a realizar a correta identificação dos paroxismos.

No entanto, apesar do desempenho alcançado ter sido abaixo do esperado, pode-se considerar que a metodologia proposta apresentou resultados promissores, pois o comportamento das redes neurais desenvolvidas, que apresentaram sensibilidade baixa e especificidade alta, foi semelhante ao comportamento descrito na literatura sobre a concordância entre os especialistas.

Não foi possível realizar análise comparativa dos resultados obtidos em outros trabalhos porque nesses estudos diferentes padrões e segmentos de sinal de uma mesma base dados são utilizados, tanto para

treinamento quanto para o teste dos classificadores. Uma comparação direta dos resultados significa uma reprodução da metodologia ou do sistema proposto e a realização de testes utilizando as bases de dados deste trabalho. Porém, a reprodução dos trabalhos não é possível porque, por diversos motivos, os mesmos não detalham todas as informações necessárias para a sua reprodução.

Dentre os trabalhos encontrados na literatura ao longo do estudo do estado da arte da detecção automática de paroxismos e aqueles citados neste trabalho, apenas o estudo de Wilson *et al.* (1999) pode ser comparado parcialmente, uma vez que este utilizou bases de dados diferentes. Porém, enquanto Wilson *et al.* (1999) utilizou registros de curta duração e um dos especialistas realizou a marcação dos registros de treinamento e teste, este trabalho utilizou registros de longa duração com características distintas entre a base de dados de treinamento e teste, além de especialistas diferentes para realizar a marcação de cada base de dados.

A comparação que pode ser feita é aquela em que, conforme os resultados e discussão apresentados por Wilson *et al.* (1999), os resultados do classificador computacional também foram menores do que a concordância dos especialistas. Além disso, a utilização do sistema em um ambiente clínico ao longo de um ano mostrou que a eficiência foi menor do que a obtida com os testes do sistema (não tendo sido mencionado o valor dessa redução).

5 CONCLUSÕES

O objetivo principal deste trabalho foi a apresentação de uma proposta de metodologia para a detecção automática de padrões epileptiformes em sinais de EEG baseada em análise morfológica foi. A metodologia proposta constituiu-se por uma primeira etapa de processamento digital do sinal, na qual foram realizadas as operações de interpolação, filtragem e conformação do espectro de frequências, e uma segunda etapa de extração do conjunto de descritores utilizando um algoritmo desenvolvido neste trabalho, especificamente para esta tarefa.

A interpolação dos sinais para uma frequência de amostragem de 1 kHz foi aplicada para padronizar os dados de entrada que, nesse caso, são os registros de EEG que podem ser adquiridos com diferentes frequências de amostragem.

O objetivo da filtragem realizada foi a remoção de desvios da linha de base dos sinais e a eliminação das interferências causadas pela rede elétrica. A decisão de não aplicar um filtro passa-baixa para selecionar apenas as frequências até 70 Hz, que representam a faixa considerada atualmente como faixa de trabalho para a detecção dos padrões epileptiformes, justifica-se no fato do conjunto de descritores utilizados na metodologia proposta conseguir diferenciar os padrões epileptiformes dos padrões de alta frequência (≥ 70 Hz).

Por sua vez, a obtenção e a aplicação de uma curva para a conformação do espectro de frequências foi a melhor solução encontrada para resolver o problema da presença das componentes de baixa frequência que, se removidas do espectro, interferem de forma prejudicial na morfologia dos padrões. A remoção dessas frequências, inferiores a 5 Hz altera a morfologia de alguns tipos de padrões não-epileptiformes de tal forma que a sua semelhança com os padrões epileptiformes é suficiente para ocasionar dúvidas na etapa de classificação.

Utilizando a conformação, o espectro de frequências foi processado de tal forma que as componentes de frequência inferiores a 5 Hz são atenuadas de forma sutil e continuam presentes no espectro, porém sem danos à morfologia dos padrões. Assim, a maneira como a interpolação e a filtragem foram aplicadas e a obtenção e utilização da curva de conformação do espectro de frequências podem ser consideradas contribuições relevantes deste trabalho.

A segunda etapa da metodologia proposta foi seleção do conjunto de descritores selecionado. O conjunto adotado neste trabalho foi

desenvolvido como parte de uma tese de doutorado do IEB-UFSC e encontra-se disponível na literatura. No entanto, o conteúdo adotado no desenvolvimento dessa metodologia ficou restrito à definição teórica de cada um dos descritores. A obtenção dos valores dos descritores foi um processo realizado por um algoritmo desenvolvido neste trabalho, com o objetivo específico de extrair os valores dos descritores dos sinais de EEG. Para tanto, foram necessárias adaptações da teoria do conjunto de descritores utilizada como base no desenvolvimento do algoritmo.

Para classificar os dados obtidos com a metodologia proposta, foram selecionados dois tipos de classificadores: um matemático, na forma de Funções Discriminantes, e outro computacional, na forma de Redes Neurais Artificiais. Além disso, no que diz respeito às redes neurais, foram aplicadas redes alimentadas pelo próprio conjunto de descritores, por componentes principais extraídas dos descritores e por subconjuntos dos descritores originais, formados aplicando a Análise de Discriminantes.

Após a realização de todos os testes – nas quatro bases de registros e nas 180 telas de EEG – foi possível constatar que o classificador que apresentou os melhores resultados foi o da Rede Neural alimentada por Componentes Principais (rpca_04). A sensibilidade atingida por essa rede nas bases de dados A, B, C, D e nas 180 telas de EEG foi, respectivamente 90, 72, 31, 34 e 66%. Os valores de especificidade obtidos pela rede rpca_04 foram 80, 90, 87, 95 e 84% para as bases de dados e telas de EEG. A eficiência média obtida pela rede nas bases de dados foi de 89%. E a concordância dessa rede com os quatro especialistas foi de 45,7%. Alguns aspectos observados durante o desenvolvimento deste trabalho podem ser considerados como contribuições relevantes para o contexto da detecção automática de padrões epileptiformes em sinais de EEG.

A metodologia proposta foi desenvolvida considerando o comportamento esperado do especialista ao analisar os sinais e registros de EEG em busca de padrões epileptiformes, ou seja, o especialista é capaz de identificar os padrões em sinais minimamente filtrados, adquiridos com diferentes frequências de amostragem e exibidos em diferentes montagens de eletrodos. Essa motivação surgiu da percepção de que essas questões, quando não incorporadas nas metodologias, podem ser limitações de aplicabilidade que não são claramente abordadas nos trabalhos descritos na literatura.

Também relacionado ao procedimento dos especialistas, mas com foco na quantidade e diversidade dos registros analisados ao longo do seu processo de aprendizagem e das diferentes referências utilizadas na

geração do conhecimento básico necessário para a apta identificação dos paroxismos, três aspectos observados ao longo deste trabalho podem ser destacados: o conjunto de treinamento, a aplicação do *threshold* e a análise do desempenho das redes neurais.

Após a realização dos testes da metodologia proposta em diferentes bases de dados e a utilização da marcação de vários especialistas como ponto de referência nos testes, constatou-se a importância de considerar as marcações de vários especialistas não apenas nos testes, mas também durante o treinamento da rede neural. O conjunto de treinamento utilizado neste trabalho é resultado do consenso absoluto entre os especialistas, o que pode ter sido uma das causas da baixa eficiência da metodologia, uma vez que as redes ficaram “especializadas” nas marcações de um conjunto específico de especialistas. Além disso, a inserção dos padrões duvidosos que são identificados pelos especialistas, mas não são anotados nas bases de dados (BARKMEIER *et al.*, 2012; GRANT *et al.*, 2014) podem fornecer um conhecimento diferenciado no treinamento do classificador que adquiriria um conhecimento mais semelhante àquele utilizado pelo especialista no seu treinamento, possibilitando a melhora no desempenho. Um grande obstáculo para utilização de um conjunto de treinamento com essas características é a indisponibilidade de registros com marcações suficientes para gerar o conjunto necessário. No entanto, considerando que para marcar as 180 telas (30 minutos) de EEG utilizadas neste trabalho os especialistas levaram, em média, cerca de 15 minutos ininterruptos. O mesmo protocolo de análise pode ser aplicado em uma quantidade de telas maior para a geração do conjunto de treinamento.

Após o treinamento é necessário que a rede neural tenha seu desempenho analisado, mas antes disso é necessário aplicar um limiar de decisão para a finalização da classificação dos padrões. Para o caso especial da análise morfológica a determinação desse limiar, chamado *threshold*, não foi uma tarefa tão simples quanto parece. Normalmente, o *threshold* escolhido é aquele que resulta na maior eficiência, porém, como os registros de EEG podem apresentar diferentes níveis de amplificação do sinal, a morfologia (amplitude) do EEG sofre alterações e a fixação de um *threshold*, quando se trata de análise morfológica, resulta em grandes variações de eficiência. Assim, esse valor precisa ser atualizado cada vez que o registro analisado apresentar níveis de amplificação diferentes ou, assim como faz o especialista ao analisar os registros, o nível de amplificação do sinal (também chamada de sensibilidade) precisa ser alterado. Neste trabalho, optou-se pela

atualização do *threshold*, mas isso demandou uma análise da curva de valores de *threshold* que só foi possível porque os registros de EEG estavam marcados. Para a implementação em um sistema automatizado, o ideal é realizar o ajuste da amplificação do sinal, contudo é necessária uma análise mais aprofundada para determinar se essa é a melhor solução.

Uma vez treinada a rede e determinado o *threshold*, um aspecto importante é o que diz respeito à forma de avaliação de desempenho da rede. De forma geral, quanto maior a eficiência de classificação, melhor é a rede neural, entretanto as redes implementadas para a detecção dos padrões epileptiformes podem ser consideradas como um caso especial por dois motivos principais: não existe um “padrão-ouro” e a variabilidade do sinal de EEG é muito alta, uma vez que os padrões eletrográficos presentes são diferentes entre os indivíduos, entre os tipos de epilepsia e entre os estados (como sono e vigília) de um mesmo indivíduo. A grande variabilidade do sinal implica na necessidade de um conjunto de testes composto por registros de longa duração ou registros de curta duração que representem, minimamente, essa variabilidade. Ainda, a inexistência de um “padrão-ouro” implica em analisar os resultados utilizando marcações de vários especialistas. Contudo, conforme mencionado anteriormente, existe uma grande dificuldade em conseguir todos esses registros e marcações. A solução encontrada neste trabalho, de utilizar telas de EEG que demandam menos tempo para sua marcação e aplicar uma adaptação da metodologia desenvolvida por Wilson *et al.* (1996), permitiu uma análise dos resultados de forma global. Isso possibilitou uma comparação direta dos resultados dos classificadores com as marcações de todos os especialistas e a visualização da concordância entre os especialistas.

Finalmente, além das questões discutidas anteriormente, dois temas devem ser estudados para continuar o trabalho realizado.

Um deles é à inserção de uma reciclagem das redes neurais ao longo de sua utilização de forma semelhante à retroalimentação implementada no sistema de *Template Matching* desenvolvido por Lodder e Putten van (2014). De forma simplificada, o sistema desenvolvido utiliza um banco de modelos de paroxismos que são comparados aos padrões presentes nos registros de EEG. Os padrões de EEG mais correlacionados aos modelos apresentados são então apresentados ao usuário do sistema que confirma a classificação com “sim”, “não” ou “talvez”. O resultado da confirmação é então utilizado para atualizar o conjunto de modelos, reforçando o treinamento e melhorando o desempenho da classificação ao longo do tempo e

personalizando-a de acordo com o conhecimento do usuário. Como os descritores utilizados também podem ser considerados modelos dos padrões, a aplicação da retroalimentação em redes neurais semelhantes às utilizadas neste trabalho pode melhorar o desempenho desses classificadores.

Outro tema considerado apropriado para aprofundar este presente trabalho, refere-se à utilização dos descritores morfológicos como variáveis para desenvolver um modelo matemático que descreva os paroxismos e possa ser utilizado em aplicações de *Template Matching* ou filtragem morfológica. Sugere-se a utilização dos descritores para o desenvolvimento de um modelo matemático após a observação de que, ao longo das simulações de Análise de Componentes Principais e Análise de Discriminantes, houve uma tendência de alguns descritores sempre estarem destacados ou presentes nos resultados da análise e esses mesmos descritores representarem algumas das características utilizadas na literatura para a definição dos padrões epileptiformes.

REFERÊNCIAS

ACIR, N.; OZTURA, I.; KUNTALP, M.; BAKLAN, B.; GÜZELİŞ, C. Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks. **IEEE Transactions on Biomedical Engineering**, v. 52, n. 1, p. 30–40, jan. 2005.

ADELI, H.; GHOSH-DASTIDAR, S.; DADMEHR, N. **Automated EEG-Based Diagnosis of Neurological Disorders: Inventing the Future of Neurology**. Boca Raton: CRC Press, 2010.

ALTMAN, D. G.; BLAND, J. M. Measurement in Medicine: The Analysis of Method Comparison Studies. **The Statistician**, v. 32, n. 3, p. 307, set. 1983.

AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY. Guideline 1: Minimum technical requirements for performing clinical electroencephalography. **Journal of Clinical Neurophysiology**, v. 23, n. 2, p. 86–91, abr. 2006a.

AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY. Guideline 4: Standards of practice in clinical electroencephalography. **Journal of Clinical Neurophysiology**, v. 23, n. 2, p. 105–6, abr. 2006b.

AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY. Guideline 8: Guidelines for recording clinical EEG on digital media. **Journal of Clinical Neurophysiology**, v. 23, n. 2, p. 122–4, abr. 2006c.

AMERICAN CLINICAL NEUROPHYSIOLOGY SOCIETY. Guideline twelve: guidelines for long-term monitoring for epilepsy. **Journal of Clinical Neurophysiology**, v. 25, n. 3, p. 170–80, jun. 2008.

ANDERSON, N. R.; DOOLITTLE, L. M. Automated analysis of EEG: opportunities and pitfalls. **Journal of Clinical Neurophysiology**, v. 27, n. 6, p. 453–457, 2010.

ARGOUD, F. I. M.; DE AZEVEDO, F. M.; NETO, J. M.; GRILLO, E. SADE3: an effective system for automated detection of epileptiform events in long-term EEG based on context information. **Medical &**

Biological Engineering & Computing, v. 44, n. 6, p. 459–70, jun. 2006.

ASKAMP, J.; VAN PUTTEN, M. J. A. M. Mobile EEG in epilepsy. **International Journal of Psychophysiology**, v. 91, n. 1, p. 30–35, jan. 2014.

BARKMEIER, D. T.; SHAH, A. K.; FLANAGAN, D.; ATKINSON, M. D.; AGARWAL, R.; FUERST, D. R.; JAFARI-KHOUSANI, K.; LOEB, J. A. High inter-reviewer variability of spike detection on intracranial EEG addressed by an automated multi-channel algorithm. **Clinical Neurophysiology**, v. 123, n. 6, p. 1088–95, jun. 2012.

BERG, A. T. Epilepsies. In: PANAYIOTOPOULOS, C. P. (Ed.). **Atlas of Epilepsies**. London: Springer London, 2010. p. 5–9.

BILODEAU, M.; BRENNER, D. **Theory of Multivariate Statistics**. New York: Springer, 1999.

BLAND, J. M.; ALTMAN, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. **Lancet**, v. 1, n. 8476, p. 307–10, 8 mar. 1986.

BOOS, C. F.; AZEVEDO, F. M. Análise de Componentes Principais como Ferramenta para o Dimensionamento de Redes Neurais Artificiais. In: XXIII Congresso Brasileiro em Engenharia Biomédica, 2012, Porto de Galinhas, **Anais...**, Porto de Galinhas, 2012, p. 1–5.

BOOS, C. F.; SCOLARO, G. R.; PEREIRA, M. C. V; AZEVEDO, F. M. Analysis of Pre-Processing Methods for Artificial Neural Network Pattern Recognition of EEG Signals. In: LONG, M. (Ed.). **World Congress on Medical Physics and Biomedical Engineering 2012**. IFMBE Proceedings. Heidelberg: Springer Berlin Heidelberg, 2013. v. 39, p. 558–561.

BOOSTANI, R.; SADATNEZHAD, K.; SABETI, M. An efficient classifier to diagnose of schizophrenia based on the EEG signals. **Expert Systems with Applications**, v. 36, n. 3, p. 6492–6499, abr. 2009.

BROWN, M. W.; PORTER, B. E.; DLUGOS, D. J.; KEATING, J.; GARDNER, A. B.; STORM, P. B.; MARSH, E. D. Comparison of novel computer detectors and human performance for spike detection in intracranial EEG. **Clinical Neurophysiology**, v. 118, n. 8, p. 1744–52, ago. 2007.

ÇALIŞIR, D.; DOĞANTEKİN, E. An automatic diabetes diagnosis system based on LDA-Wavelet Support Vector Machine Classifier. **Expert Systems with Applications**, v. 38, n. 7, p. 8311–8315, jul. 2011.

CASSON, A. J.; LUNA, E.; RODRIGUEZ-VILLEGAS, E. Performance metrics for the accurate characterisation of interictal spike detection algorithms. **Journal of Neuroscience Methods**, v. 177, n. 2, p. 479–87, 15 mar. 2009.

CHAPMAN, R. M.; NOWLIS, G. H.; MCCRARY, J. W.; CHAPMAN, J. A.; SANDOVAL, T. C.; GUILLILY, M. D.; GARDNER, M. N.; REILLY, L. A. Brain event-related potentials: Diagnosing early-stage Alzheimer's disease. **Neurobiology of Aging**, v. 28, n. 2, p. 194–201, mar. 2007.

CHAVAKULA, V.; SÁNCHEZ FERNÁNDEZ, I.; PETERS, J. M.; POPLI, G.; BOSL, W.; RAKHADE, S.; ROTENBERG, A.; LODDENKEMPER, T. Automated quantification of spikes. **Epilepsy & Behavior**, v. 26, n. 2, p. 143–52, fev. 2013.

DE LUCIA, M.; FRITSCHY, J.; DAYAN, P.; HOLDER, D. S. A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis. **Medical & Biological Engineering & Computing**, v. 46, n. 3, p. 263–272, mar. 2008.

DE MAESSCHALCK, R.; JOUAN-RIMBAUD, D.; MASSART, D. L. The Mahalanobis distance. **Chemometrics and Intelligent Laboratory Systems**, v. 50, n. 1, p. 1–18, jan. 2000.

DEUSCH, G.; EISEN, A. (EDS.). Recommendations for the practice of clinical neurophysiology: guidelines of the International Federation of Clinical Neurophysiology. **Electroencephalography and clinical neurophysiology. Supplement**, v. 52, p. 1–304, jan. 1999.

DÜMPELMANN, M.; ELGER, C. E. Visual and automatic investigation of epileptiform spikes in intracranial EEG recordings. **Epilepsia**, v. 40, n. 3, p. 275–85, mar. 1999.

EL-GOHARY, M.; MCNAMES, J.; ELSAS, S. User-guided interictal spike detection. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, Vancouver, Canadá, **Anais...**, New Jersey, 2008, p. 821–824.

EXARCHOS, T. P.; TZALLAS, A. T.; FOTIADIS, D. I.; KONITSIOTIS, S.; GIANNOPOULOS, S. EEG transient event detection and classification using association rules. **IEEE Transactions on Information Technology in Biomedicine**, v. 10, n. 3, p. 451–7, jul. 2006.

FERNÁNDEZ-TORRE, J. L. Interictal EEG. In: PANAYIOTOPOULOS, C. P. (Ed.). **Atlas of Epilepsies**. London: Springer London, 2010. p. 701–712.

FISHER, R. S.; VAN EMDE BOAS, W.; BLUME, W.; ELGER, C.; GENTON, P.; LEE, P.; ENGEL, J. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). **Epilepsia**, v. 46, n. 4, p. 470–2, abr. 2005.

FLANAGAN, D.; AGARWAL, R.; WANG, Y. H.; GOTMAN, J. Improvement in the performance of automated spike detection using dipole source features for artefact rejection. **Clinical Neurophysiology**, v. 114, n. 1, p. 38–49, jan. 2003.

GARSON, G. D. **Discriminant Function Analysis**. Asheboro: Statistical Associates Publishers, 2012.

GHOSH-DASTIDAR, S.; ADELI, H.; DADMEHR, N. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. **IEEE Transactions on Biomedical Engineering**, v. 55, n. 2 Pt 1, p. 512–8, fev. 2008.

GIRI, D.; RAJENDRA ACHARYA, U.; MARTIS, R. J.; VINITHA SREE, S.; LIM, T.-C.; AHAMED, T.; SURI, J. S. Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and

Discrete Wavelet Transform. **Knowledge-Based Systems**, v. 37, p. 274–282, jan. 2013.

GOTMAN, J. High frequency oscillations: the new EEG frontier? **Epilepsia**, v. 51 Suppl 1, p. 63–5, mar. 2010.

GOTMAN, J.; GLOOR, P. Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. **Electroencephalography and Clinical Neurophysiology**, v. 41, n. 5, p. 513–29, nov. 1976.

GRANT, A. C.; ABDEL-BAKI, S. G.; WEEDON, J.; ARNEDO, V.; CHARI, G.; KOZIORYNSKA, E.; LUSHBOUGH, C.; MAUS, D.; MCSWEEN, T.; MORTATI, K. A.; REZNIKOV, A.; OMURTAG, A. EEG interpretation reliability and interpreter confidence: A large single-center study. **Epilepsy and Behavior**, v. 32, p. 102–107, mar. 2014.

GUEDES DE OLIVEIRA, P.; QUEIROZ, C.; LOPES DA SILVA, F. Spike detection based on a pattern recognition approach using a microcomputer. **Electroencephalography and Clinical Neurophysiology**, v. 56, n. 1, p. 97–103, jul. 1983.

GUERRERO-MOSQUERA, C.; TRIGUEROS, A. M.; NAVIA-VAZQUEZ, A. EEG Signal Processing for Epilepsy. In: STEVANOVIC, D. (Ed.). **Epilepsy - Histological, Electroencephalographic and Psychological Aspects**. Rijeka: InTech, 2012. p. 49–74.

HALFORD, J. J. Computerized epileptiform transient detection in the scalp electroencephalogram: obstacles to progress and the example of computerized ECG interpretation. **Clinical Neurophysiology**, v. 120, n. 11, p. 1909–15, nov. 2009.

HALFORD, J. J.; PRESSLY, W. B.; BENBADIS, S. R.; TATUM, W. O.; TURNER, R. P.; ARAIN, A.; PRITCHARD, P. B.; EDWARDS, J. C.; DEAN, B. C. Web-based collection of expert opinion on routine scalp EEG: software development and interrater reliability. **Journal of Clinical Neurophysiology**, v. 28, n. 2, p. 178–184, abr. 2011.

HALFORD, J. J.; SCHALKOFF, R. J.; ZHOU, J.; BENBADIS, S. R.; TATUM, W. O.; TURNER, R. P.; SINHA, S. R.; FOUNTAIN, N. B.;

ARAIN, A.; PRITCHARD, P. B.; KUTLUAY, E.; MARTZ, G.; EDWARDS, J. C.; WATERS, C.; DEAN, B. C. Standardized database development for EEG epileptiform transient detection: EEGnet scoring system and machine learning analysis. **Journal of Neuroscience Methods**, v. 212, n. 2, p. 308–16, 30 jan. 2013.

HÄRDLE, W.; SIMAR, L. **Applied Multivariate Statistical Analysis**. 2. ed. New York: Springer, 2007.

HARNER, R. Automatic EEG spike detection. **Clinical EEG and Neuroscience**, v. 40, n. 4, p. 262–270, 1 out. 2009.

INDIRADEVI, K. P.; ELIAS, E.; SATHIDEVI, P. S.; DINESH NAYAK, S.; RADHAKRISHNAN, K. A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram. **Computers in Biology and Medicine**, v. 38, n. 7, p. 805–16, jul. 2008.

JACOBS, J.; STABA, R.; ASANO, E.; OTSUBO, H.; WU, J. Y.; ZIJLMANS, M.; MOHAMED, I.; KAHANE, P.; DUBEAU, F.; NAVARRO, V.; GOTMAN, J. High-frequency oscillations (HFOs) in clinical epilepsy. **Progress in Neurobiology**, v. 98, n. 3, p. 302–315, set. 2012.

KRENDL, R.; LURGER, S.; BAUMGARTNER, C. Absolute spike frequency predicts surgical outcome in TLE with unilateral hippocampal atrophy. **Neurology**, v. 71, n. 6, p. 413–8, 5 ago. 2008.

KUTLU, Y.; KUNTALP, M.; KUNTALP, D. Optimizing the performance of an MLP classifier for the automatic detection of epileptic spikes. **Expert Systems with Applications**, v. 36, p. 7567–7575, 2009.

LARSSON, P. G.; WILSON, J.; EEG-OLOFSSON, O. A new method for quantification and assessment of epileptiform activity in EEG with special reference to focal nocturnal epileptiform activity. **Brain topography**, v. 22, n. 1, p. 52–9, jun. 2009.

LIU, H. S.; ZHANG, T.; YANG, F. S. A multistage, multimethod approach for automatic detection and classification of epileptiform EEG. **IEEE Transactions on Biomedical Engineering**, v. 49, n. 12 Pt 2, p. 1557–66, dez. 2002.

LODDER, S. S.; ASKAMP, J.; VAN PUTTEN, M. J. A M. Inter-ictal spike detection using a database of smart templates. **Clinical Neurophysiology**, v. 124, n. 12, p. 2328–2335, dez. 2013.

LODDER, S. S.; VAN PUTTEN, M. J. A M. A self-adapting system for the automated detection of inter-ictal epileptiform discharges. **PLoS ONE**, v. 9, n. 1, p. e85180, jan. 2014.

MADDEN, H. H. Comments on the Savitzky-Golay convolution method for least-squares-fit smoothing and differentiation of digital data. **Analytical Chemistry**, v. 50, n. 9, p. 1383–1386, 1978.

MALMIVUO, J.; PLONSEY, R. **Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields**. New York: Oxford University Press, 1995.

MARI, F.; ZELMANN, R.; ANDRADE-VALENCA, L.; DUBEAU, F.; GOTMAN, J. Continuous high-frequency activity in mesial temporal lobe structures. **Epilepsia**, v. 53, n. 5, p. 797–806, maio 2012.

MILLER, J. W.; GOTMAN, J. The meaning of interictal spikes in temporal lobe epilepsy: should we count them? **Neurology**, v. 71, n. 6, p. 392–3, 5 ago. 2008.

MONTENEGRO, M. A.; CENDES, F.; GUERREIRO, M. M.; GUERREIRO, C. A. M. **EEG na Prática Clínica**. 2. ed. Rio de Janeiro: Revinter, 2012.

MORMANN, F.; ANDRZEJAK, R. G.; ELGER, C. E.; LEHNERTZ, K. Seizure prediction: The long and winding road. **Brain**, v. 130, n. Pt 2, p. 314–333, fev. 2007.

MYATT, G. J.; JOHNSON, W. P. **Making Sense of Data II: A practical Guide to Data Visualization, Advanced Data Mining Methods, and Applications**. New Jersey: John Wiley & Sons, 2009.

MYLES, P. S.; CUI, J. Using the Bland-Altman method to measure agreement with repeated measures. **British Journal of Anaesthesia**, v. 99, n. 3, p. 309–11, set. 2007.

NAGHSH-NILCHI, A. R.; AGHASHAHI, M. Epilepsy seizure detection using eigen-system spectral estimation and Multiple Layer Perceptron neural network. **Biomedical Signal Processing and Control**, v. 5, n. 2, p. 147–157, abr. 2010.

NICOLAI, J.; KASTELEIJN-NOLST TRENITÉ, D. Interictal discharges and cognition. **Epilepsy & Behavior**, v. 22, n. 1, p. 134–6, set. 2011.

NONCLERCQ, A; FOULON, M.; VERHEULPEN, D.; DE COCK, C.; BUZATU, M.; MATHYS, P.; VAN BOGAERT, P. Spike detection algorithm automatically adapted to individual patients applied to spike-and-wave percentage quantification. **Clinical neurophysiology**, v. 39, n. 2, p. 123–31, abr. 2009.

NONCLERCQ, A.; FOULON, M.; VERHEULPEN, D.; DE COCK, C.; BUZATU, M.; MATHYS, P.; VAN BOGAERT, P. Cluster-based spike detection algorithm adapts to interpatient and inpatient variation in spike morphology. **Journal of Neuroscience Methods**, v. 210, n. 2, p. 259–65, 30 set. 2012.

NUWER, M. R.; COMI, G.; EMERSON, R.; FUGLSANG-FREDERIKSEN, A.; GUÉRIT, J.-M.; HINRICHS, H.; IKEDA, A.; LUCCAS, F. J. C.; RAPPERSBERGER, P. IFCN standards for digital recording of clinical EEG. In: DEUSCHL, G.; EISEN, A. (Eds.). **Recommendations for the Practice of Clinical Neurophysiology: Guidelines of the IFCN**. [s.l.]: Elsevier Science B.V., 1999. p. 11–14.

OIKONOMOU, V. P.; TZALLAS, A T.; FOTIADIS, D. I. A Kalman filter based methodology for EEG spike enhancement. **Computer Methods and Programs in Biomedicine**, v. 85, n. 2, p. 101–8, fev. 2007.

OSSADTCHI, A.; BAILLET, S.; MOSHER, J. C.; THYERLEI, D.; SUTHERLING, W.; LEAHY, R. M. Automated interictal spike detection and source localization in magnetoencephalography using independent components analysis and spatio-temporal clustering. **Clinical Neurophysiology**, v. 115, n. 3, p. 508–522, mar. 2004.

PALANIAPPAN, R. Electroencephalogram Signals from Imagined Activities: A Novel Biometric Identifier for a Small Population. In:

CORCHADO, E.; YIN, H.; BOTTI, V.; FYFE, C. (Eds.). **Intelligent Data Engineering and Automated Learning – IDEAL 2006**. Lecture Notes in Computer Science. Berlin: Springer Berlin Heidelberg, 2006. v. 4224, p. 604–611.

PANAYIOTOPOULOS, C. P. (ED.). **Atlas of Epilepsies**. London: Springer London, 2010.

PEREIRA, M. DO C. V. **Avaliação de Técnicas de Pré-Processamento de Sinais de EEG para Detecção de Eventos Epileptogênicos Utilizando Redes Neurais Artificiais**. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2003.

PILLAI, J.; SPERLING, M. R. Interictal EEG and the diagnosis of epilepsy. **Epilepsia**, v. 47 Suppl 1, n. Table 1, p. 14–22, jan. 2006.

RAYKOV, T.; MARCOULIDES, G. A. **An Introduction to Applied Multivariate Analysis**. New York: Routledge Academic, 2008.

RODIN, E.; CONSTANTINO, T.; RAMPP, S.; WONG, P. K. Spikes and epilepsy. **Clinical EEG and Neuroscience**, v. 40, n. 4, p. 288–299, 1 out. 2009.

ROMO VÁZQUEZ, R.; VÉLEZ-PÉREZ, H.; RANTA, R.; LOUIS DORR, V.; MAQUIN, D.; MAILLARD, L. Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling. **Biomedical Signal Processing and Control**, v. 7, n. 4, p. 389–400, jul. 2012.

SANEI, S.; CHAMBERS, J. A. **EEG Signal Processing**. 1. ed. West Sussex: John Wiley & Sons Ltd, 2007.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SCHAFFER, R. W. What is a savitzky-golay filter? **IEEE Signal Processing Magazine**, v. 28, n. July, p. 111–117, 2011.

SCHOMER, D. L.; SILVA, F. L. DA (EDS.). **Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields**. 6. ed. Philadelphia: Lippincott Williams & Wilkins, 2011.

SCOLARO, G. R. **Sistema computacional automatizado para a identificação e contagem de eventos epileptiformes em sinais de eletroencefalografia de longa duração**. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2014.

SCOLARO, G. R.; BOOS, C. F.; AZEVEDO, F. M. Neural Classifier for Detection and Classification of Spikes and Sharp Waves. In: International Conference on Bio-inspired Systems and Signal Processing, 2011, Roma, Itália, **Anais...**, [s.l.]: SciTePress - Science and Technology Publications, 2011, p. 504–509.

SELVITELLI, M. F.; WALKER, L. M.; SCHOMER, D. L.; CHANG, B. S. The relationship of interictal epileptiform discharges to clinical epilepsy severity: a study of routine electroencephalograms and review of the literature. **Journal of Clinical Neurophysiology**, v. 27, n. 2, p. 87–92, abr. 2010.

SEZER, E.; IŞIK, H.; SARACOĞLU, E. Employment and comparison of different Artificial Neural Networks for epilepsy diagnosis from EEG signals. **Journal of Medical Systems**, v. 36, n. 1, p. 347–62, fev. 2012.

SHIBASAKI, H.; NAKAMURA, M.; SUGI, T.; NISHIDA, S.; NAGAMINE, T.; IKEDA, A. Automatic interpretation and writing report of the adult waking electroencephalogram. **Clinical Neurophysiology**, v. 125, n. 6, p. 1081–1094, jun. 2014.

SO, E. L. Interictal epileptiform discharges in persons without a history of seizures: what do they mean? **Journal of Clinical Neurophysiology**, v. 27, n. 4, p. 229–238, ago. 2010.

SOVIERZOSKI, M. A. **Avaliação de descritores morfológicos na identificação de eventos epileptiformes**. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2009.

SOVIERZOSKI, M. A.; SCHWARZ, L.; DE AZEVEDO, F. M. Binary neural classifier of raw EEG data to separate spike and sharp wave of the eye blink artifact. In: 5th International Conference on Natural Computation, 2009, Tianjin, China, **Anais...**, New Jersey, 2009, p. 126–130.

SRINIVASAN, V.; ESWARAN, C.; SRIRAAM, N. Approximate entropy-based epileptic EEG detection using artificial neural networks. **IEEE Transactions on Information Technology in Biomedicine**, v. 11, n. 3, p. 288–95, maio 2007.

SUBASI, A.; GURSOY, M. I. EEG signal classification using PCA, ICA, LDA and support vector machines. **Expert Systems with Applications**, v. 37, n. 12, p. 8659–8666, dez. 2010.

TZALLAS, A T.; KARVELIS, P. S.; KATSIS, C. D.; FOTIADIS, D. I.; GIANNOPOULOS, S.; KONITSIOTIS, S. A method for classification of transient events in EEG recordings: application to epilepsy diagnosis. **Methods of Information in Medicine**, v. 45, n. 6, p. 610–21, jan. 2006.

TZALLAS, A. T.; TSIPOURAS, M. G.; TSALIKAKIS, D. G.; KARVOUNIS, E. C.; ASTRAKAS, L.; KONITSIOTIS, S.; TZAPHLIDOU, M. Automated Epileptic Seizure Detection Methods : A Review Study. In: STEVANOVIC, D. (Ed.). **Epilepsy - Histological, Electroencephalographic and Psychological Aspects**. Rijeka: InTech, 2012. p. 75–98.

ÜBEYLI, E. D. Statistics over features: EEG signals analysis. **Computers in Biology and Medicine**, v. 39, n. 8, p. 733–41, ago. 2009.

VAN BELLE, G.; FISHER, L. D.; HEAGERTY, P. J.; LUMLEY, T. **Biostatistics: A Methodology for the Health Sciences**. 2. ed. New Jersey: John Wiley & Sons, 2004.

WALCZAK, S.; NOWACK, W. J. An artificial neural network approach to diagnosing epilepsy using lateralized bursts of theta EEGs. **Journal of Medical Systems**, v. 25, n. 1, p. 9–20, fev. 2001.

WANG, C.; ZOU, J.; ZHANG, J.; WANG, M.; WANG, R. Feature extraction and recognition of epileptiform activity in EEG by combining

PCA with ApEn. **Cognitive neurodynamics**, v. 4, n. 3, p. 233–40, set. 2010.

WEBBER, W. R.; LITT, B.; WILSON, K.; LESSER, R. P. Practical detection of epileptiform discharges (EDs) in the EEG using an artificial neural network: a comparison of raw and parameterized EEG data. **Electroencephalography and Clinical Neurophysiology**, v. 91, n. 3, p. 194–204, set. 1994.

WILSON, S. B.; HARNER, R. N.; DUFFY, F. H.; THARP, B. R.; NUWER, M. R.; SPERLING, M. R. Spike detection. I. Correlation and reliability of human experts. **Electroencephalography and Clinical Neurophysiology**, v. 98, n. 3, p. 186–98, mar. 1996.

WILSON, S. B.; SCHEUER, M. L.; PLUMMER, C.; YOUNG, B.; PACIA, S. Seizure detection: correlation of human experts. **Clinical Neurophysiology**, v. 114, n. 11, p. 2156–2164, nov. 2003.

WILSON, S. B.; TURNER, C. A.; EMERSON, R. G.; SCHEUER, M. L. Spike detection II: automatic, perception-based detection and clustering. **Clinical Neurophysiology**, v. 110, n. 3, p. 404–11, mar. 1999.

WILSON, S.; EMERSON, R. Spike detection: A review and comparison of algorithms. **Clinical Neurophysiology**, v. 113, n. 12, p. 1873–1881, dez. 2002.

XU, G.; WANG, J.; ZHANG, Q.; ZHANG, S.; ZHU, J. A spike detection method in EEG based on improved morphological filter. **Computers in Biology and Medicine**, v. 37, n. 11, p. 1647–52, nov. 2007.

ZAKI, R.; BULGIBA, A.; ISMAIL, R.; ISMAIL, N. A. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: A systematic review. **PLoS ONE**, v. 7, n. 5, p. e37908, jan. 2012.

ZHOU, J.; SCHALKOFF, R. J.; DEAN, B. C.; HALFORD, J. J. A study of morphology-based wavelet features and multiple-wavelet strategy for EEG signal classification: Results and selected statistical analysis. In: 35th Annual International Conference of the IEEE Engineering in

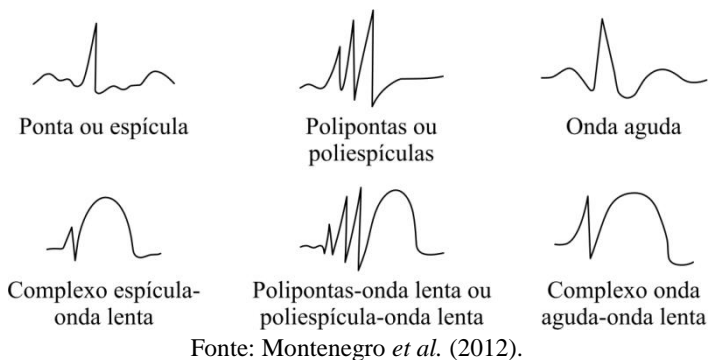
Medicine and Biology Society, 2013, Osaka, Japão, **Anais...**, New Jersey, 2013, p. 5998–6002.

ZIJLMANS, M.; JIRUSKA, P.; ZELMANN, R.; LEIJTEN, F. S. S.; JEFFERYS, J. G. R.; GOTMAN, J. High-frequency oscillations as a new biomarker in epilepsy. **Annals of Neurology**, v. 71, n. 2, p. 169–178, fev. 2012.

ZURADA, J. M. **Introduction to Artificial Neural Systems**. St. Paul: West Publishing Company, 1992.

Apêndice A – Protocolo para marcação das telas de EEG apresentado aos especialistas.

- Marcar em VERDE todos os paroxismos CERTOS (não há dúvida);



- Marcar em VERMELHO os paroxismos DUVIDOSOS;
- Realizar a marcação de acordo com a seguinte definição:

Quadro 12.1. Características da atividade epileptiforme

1. **Simetria versus assimetria da onda:** ondas agudas e espículas epileptiformes têm uma ascensão aguda e a segunda fase (descida) mais lenta, conferindo aspecto assimétrico entre dois lados da onda. Ondas agudas não epileptiformes são simétricas quanto à duração da fase de subida e descida.
2. **Ondas agudas e espículas epileptiformes são frequentemente seguidas por ondas lentas**, podendo ser da mesma polaridade ou de polaridade oposta. Ondas agudas e espículas não epileptiformes raramente são seguidas por atividade lenta.
3. **Ondas agudas e espículas epileptiformes geralmente são bifásicas ou trifásicas**, enquanto a atividade não epileptiforme é monofásica.
4. **Ondas agudas e espículas epileptiformes têm duração diferente da atividade de fundo normal do paciente**, sendo mais alta ou mais baixa e, conseqüentemente, chamam a atenção do eletrencefalografista.
5. A atividade de fundo próxima da atividade epileptiforme geralmente é perturbada, formando um “campo” ao redor das ondas agudas.

Fonte: Montenegro *et al.* (2012).

Apêndice B – Resumo dos elementos do conjunto de descritores

Identificação	Descrição
Amax, Bmin	amplitude máxima e mínima do sinal, respectivamente
Tdif	intervalo de tempo entre Amax e Bmin do sinal
DifAB	diferença entre Amax e Bmin do sinal
Amax_pts	amplitude máxima do segmento de 300ms
Bmin_pts	amplitude mínima do segmento de 300ms
Tdif_pts	intervalo de tempo entre Amax_pts e Bmin_pts
DifAB_pts	diferença entre Amax_pts e Bmax_pts
T	duração do padrão
dbase	base do padrão, relativa à sua máxima amplitude absoluta
dpos	lado oposto ao ângulo do pico positivo
dneg	lado oposto ao ângulo do pico negativo
trp	distância da amplitude máxima ao extremo do padrão
trn	distância da amplitude mínima ao extremo do padrão
tgp, tgn	tangente do ângulo positivo e negativo, respectivamente
ang_pos, ang_neg	ângulo do pico positivo e do pico negativo
ang_pico	ângulo do pico (vértice de maior amplitude absoluta)
Amax_i	amplitude máxima da região inicial do sinal
Bmin_i	amplitude mínima da região inicial do sinal
DifA_i	diferença em amplitude entre Amax e Amax_i
tA_i	diferença em distância entre Amax e Amax_i
DifB_i	diferença em amplitude entre Bmin e Bmin_i
tB_i	diferença em distância entre Amin e Amin_i
Amax_f	amplitude máxima da região final do sinal
Bmin_f	amplitude mínima da região final do sinal
DifA_f	diferença em amplitude entre Amax e Amax_f
tA_f	diferença em distância entre Amax e Amax_f
DifB_f	diferença em amplitude entre Bmin e Bmin_f
tB_f	diferença em distância entre Amin e Amin_f
desvio, desvioC	desvio padrão do sinal e do segmento de 300ms
media, mediaC	média do sinal e do segmento de 300ms
coef, coefC	coeficiente de variação do sinal e do segmento de 300ms
vari, variC	variância do sinal e do segmento de 300ms
CF, CFC	fator de crista do sinal e do segmento de 300ms
entrop_nor	entropia (normalizada) do sinal
entrop_norC	entropia (normalizada) do segmento de 300ms

Apêndice C – Interface gráfica utilizada nos testes dos filtros

Visualização dos filtros utilizados

1 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0

1 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0

1 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Legenda(s):

Controle da visualização

INICIAR Início ANL ATUALIZAR Próx. Final RESET

Informações sobre o conjunto de padrões

Tipo de padrão: Freq. amostragem:
 Número de padrão: Freq. original (Hz):
 Total de padrões: Sinal interpolado:

Pré-processamento utilizado (configuração e visualização)

Filtro notch (baseline): ver F.T. aplicar ver F.T. aplicar ver F.T. aplicar

fc = Hz BW = Hz

Filtro notch (power line): ver F.T. aplicar ver F.T. aplicar

Filtro passa-baixa: ver F.T. aplicar ver F.T. aplicar

fc = Hz BW = Hz

Filtro passa-alta: ver F.T. aplicar ver F.T. aplicar

fc = Hz BW = Hz

Filtro notch (power line): ver F.T. aplicar ver F.T. aplicar

fc = Hz BW = Hz

Filtro passa-baixa: ver F.T. aplicar ver F.T. aplicar

fc = Hz BW = Hz

Banco de filtros: ver F.T. aplicar ver F.T. aplicar

wavet = aplicar ver F.T. aplicar

fc = Hz aplicar ver F.T. aplicar

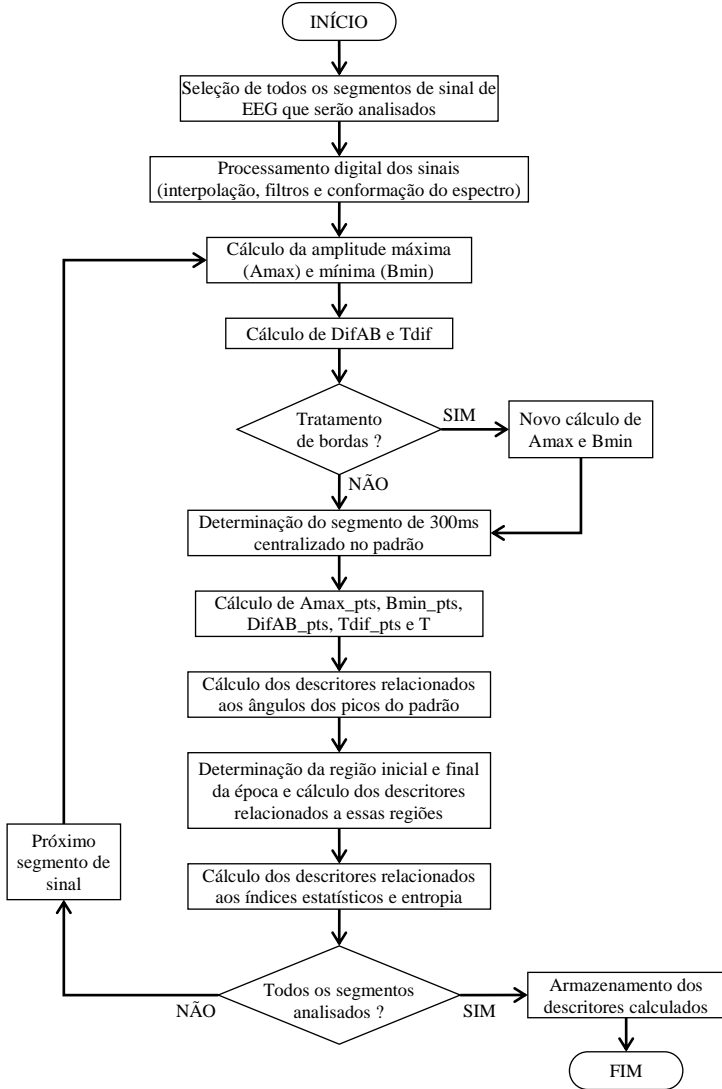
atenuação = aplicar ver F.T. aplicar

Conformação espectral: ver F.T. aplicar ver F.T. aplicar

tipo = aplicar ver F.T. aplicar

Zoom (frequência): de Hz a Hz

Apêndice D – Fluxograma do funcionamento do algoritmo desenvolvido para calcular o conjunto de descritores



Apêndice E – Interface gráfica utilizada nas simulações e testes para implementação das redes neurais

Redes Neurais - treinamento e testes

Parâmetros da rede

Nº de neurônios das camadas:
 entrada: 45 (no máximo 45)
 oculta: 8 (de 7 a 14)
 saída: 1 (valor padrão)

Função de ativação:
 Tangente hiperbólica k = 0.01
 camada de saída: Logística k = 0.13

Taxa de aprendizagem: 0.01
 Constante de momento: 0.9

Cálculos de pesos:
 Erro mínimo (em/époc)
 Erro mínimo (em/pod): 0.12
 Padrões li aprendidos: 0 % (máx)
 Nº máximo de épocas: 20000

Base da camada oculta: 1
 Base da camada de saída: 1

Pesos sinápticos INICIAIS

Camada oculta: pesos 1: [-0.1; 0.1]

1	2
1	0.0493 -0.1
2	-0.0801 0.1

 Camada de saída: pesos 1: [-0.1; 0.1]

1	2
1	0.0688
2	0.0349
3	-0.0229

Entradas da rede

Treinamento: 1
 Validação: 1
 Testes: 1

Erro atual (tm):
 Erro de época: Erro mín. tm (epc):
 Tempo decorrido: Erro mín. vó (epo):

Saídas da rede

treino: 1 2 3
 resultados (do) teste(s):
 dk: yk
 1 2 3

Valores de erro da rede

Erro por padrão: ENQ (tm,vid) por época

1	2	3
1	1	2

Análise da rede (desempenho)

Threshold:
 Qtd. padrões: Qtd. EEP:
 Resultado da análise do desempenho:
 NP - Sensibilid...
 VN - Especifici...
 FP - Eficiência
 FN - Vpp

Opções de 'tareja'

Treinamento concluído com sucesso
 Teste finalizado com sucesso
 tempo decorrido:

Rede nº

(Receba os parâmetros e salve os arquivos em uma pasta do Excel)

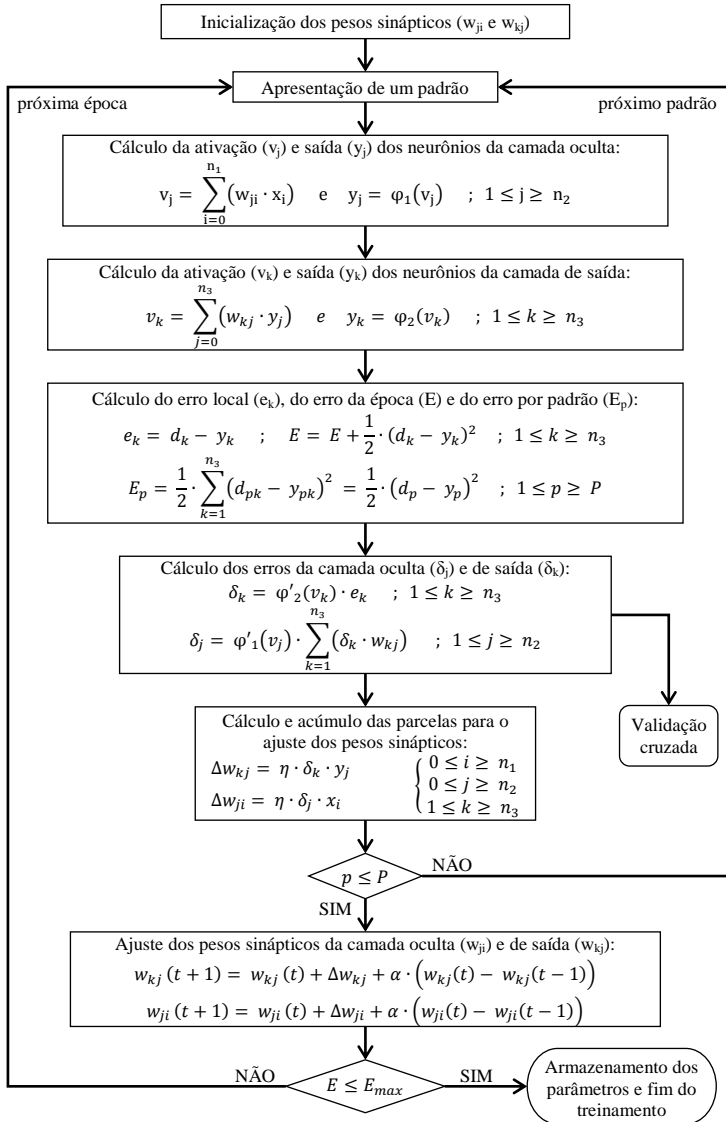
Gráfico

Erro mín. tm (epc):
 Erro mín. vó (epo):

Erro atual (tm):
 Erro de época:

Erro mín. tm (epc):
 Erro mín. vó (epo):

Apêndice F – Fluxograma do processo de treinamento realizado nas redes neurais artificiais desenvolvidas



Apêndice G – Interface gráfica utilizada nas simulações para teste e seleção dos classificadores

Seleção do conjunto dos conjuntos

» Selecione o conjunto que será analisado:

anlgo_sinais_anf_b.mat

» Selecione as opções para o processamento digital do sinal:

» frequência de amostragem: Hz

» Interpolar para a frequência de: Hz

Filtro notch (baseline)

Filtro notch (powerline)

Filtro passa-baixa

Filtro passa-alta

Banco de filtros

Conformação de espectro:

Seleção do classificador

REDE NEURAL ARTIFICIAL

» Carregar rede neural

» Neurônios por camada:

Entrada:

Oculto:

Saída:

» Função de ativação:

Camada oculta:

Camada de saída:

» Taxa de aprendizagem:

» Constante de momento:

» Bias da camada oculta:

» Bias da camada de saída:

» Pêso da camada oculta:

» Pêso da camada de saída:

Resultado da classificação

» Salda brute do classificador (para teste (normal artificial)):

1	1
2	2
3	3
4	4
5	5

» Classificação (threshold):

1 = EEP & 0 = NEP

» Threshold:

APLICAR threshold

ANALISAR threshold

Seleção dos descritores

Anmax

Bmin

DdB

Anmax_pts

Bmin_pts

Torq_pts

DdB_pts

T

dbase

dneq

dpos

IB_L

Anmax_f

Bmin_f

DdB_f

IA_f

DdB_f

Anmax_j

Bmin_j

DdB_j

T

dbase

dneq

dpos

coefC

Vref

VrefC

CF

CFC

entrop_norc

entrop_nor

desvio

desvioC

media

DdB_L

IA_L

coefJ

ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

» Pesos:

PCA 1	0.191	1
PCA 2	0.206	2

» Normalização das entradas:

» Pesos rotacionados:

PCA 1	0.157	1
PCA 2	0.219	2

ANÁLISE DE DISCRIMINANTES (LDA)

» Pesos das funções:

EEP	NEP
1	1
2	2
3	3

» Função discriminante:

função 1

CLASIFICAR conjunto

Status da análise:

ARMAZENAR classificação

» Indique o local e nome do arquivo no qual o resultado será armazenado:

(C:\teste\result_cont)

» Indique a extensão (tipo) do arquivo:

MATLAB data (.mat)

Planilha Excel (.xls)

ANALISAR classificação

VP	
VN	
FN	
Sensibilidade	
Especificidade	
Eficiência	
%FP	
FP/min	
FN/min	

Visualização do padrão - Amplitude [uV] vs. Tempo [s]

Apêndice H – Matriz de Correlação dos descritores da Base A

NEP \ EEP	Amax	Bmin	Tdif	DifAB	Amax_pts	Bmin_pts	Tdif_pts	DifAB_pts	T
Amax	-0,160	0,082	0,121	-0,144	-0,153	0,078	0,101	-0,146	0,178
Bmin	0,167	-0,129	-0,020	0,181	0,155	-0,115	0,002	0,177	-0,040
Tdif	0,173	-0,213	0,046	0,245	0,118	-0,209	-0,019	0,230	-0,015
DifAB	-0,195	0,126	0,083	-0,194	-0,184	0,115	0,058	-0,194	0,129
Amax_pts	-0,173	0,091	0,105	-0,157	-0,160	0,087	0,096	-0,158	0,178
Bmin_pts	0,165	-0,131	-0,015	0,182	0,152	-0,118	0,016	0,178	-0,014
Tdif_pts	0,147	-0,029	-0,030	0,098	0,165	-0,023	-0,135	0,110	-0,133
DifAB_pts	-0,210	0,138	0,074	-0,211	-0,193	0,128	0,048	-0,209	0,118
T	0,108	-0,063	-0,022	0,103	0,093	-0,055	-0,064	0,096	-0,056
dbase	0,001	-0,037	0,028	0,027	0,020	-0,035	-0,049	0,039	-0,025
dpos	-0,071	0,048	0,043	-0,073	-0,036	0,049	-0,013	-0,059	0,099
dneg	0,117	-0,092	-0,049	0,128	0,111	-0,093	-0,116	0,135	-0,153
trpl	-0,197	0,087	0,111	-0,167	-0,174	0,083	0,087	-0,163	0,207
trnl	-0,124	0,076	-0,004	-0,120	-0,122	0,064	-0,059	-0,118	-0,030
tgp	-0,072	0,072	-0,064	-0,090	-0,044	0,067	-0,024	-0,078	-0,009
tgn	-0,169	0,063	0,089	-0,134	-0,179	0,064	0,045	-0,151	0,041
ang_pos	0,138	-0,070	-0,040	0,124	0,142	-0,066	-0,086	0,131	-0,078
ang_neg	0,258	-0,180	-0,022	0,267	0,227	-0,171	-0,111	0,262	-0,122
ang_pico	0,207	-0,113	-0,058	0,190	0,205	-0,103	-0,101	0,196	-0,128
Amax_i	-0,131	0,059	0,001	-0,111	-0,101	0,058	0,097	-0,102	0,095
Bmin_i	0,090	0,057	-0,026	0,006	0,088	0,063	-0,029	0,000	-0,025
DifA_i	-0,089	0,054	0,113	-0,086	-0,096	0,051	0,031	-0,094	0,121
tA_i	-0,137	0,084	0,019	-0,133	-0,127	0,078	0,015	-0,132	0,037
DifB_i	0,095	-0,156	0,002	0,163	0,085	-0,148	0,033	0,163	0,002
tB_i	-0,023	0,011	-0,034	-0,020	-0,009	0,002	0,035	-0,007	0,051
Amax_f	-0,135	0,055	0,057	-0,111	-0,117	0,049	0,016	-0,104	0,159
Bmin_f	-0,006	0,087	0,068	-0,066	0,000	0,085	-0,062	-0,067	0,022
DifA_f	-0,069	0,050	0,064	-0,073	-0,071	0,052	0,091	-0,080	0,053
tA_f	0,115	-0,100	-0,068	0,133	0,070	-0,097	0,031	0,116	-0,095
DifB_f	0,178	-0,206	-0,070	0,243	0,159	-0,191	0,066	0,239	-0,032
tB_f	-0,024	-0,084	0,062	0,048	-0,058	-0,089	0,020	0,037	0,066
desvio	-0,166	0,097	0,052	-0,157	-0,144	0,089	0,016	-0,151	0,076
desvioC	-0,118	0,015	0,060	-0,073	-0,110	0,010	0,054	-0,069	0,107
media	-0,169	0,034	0,054	-0,114	-0,136	0,029	0,029	-0,099	0,091
mediaC	-0,094	-0,076	0,032	0,005	-0,085	-0,078	0,064	0,014	0,104
coef	-0,042	0,138	0,032	-0,122	-0,064	0,132	0,008	-0,140	-0,015
coefC	-0,122	0,248	0,054	-0,244	-0,109	0,237	-0,015	-0,247	0,013
vari	-0,129	0,115	0,108	-0,151	-0,126	0,106	0,037	-0,154	0,102
variC	-0,125	0,038	0,097	-0,093	-0,128	0,031	0,073	-0,097	0,107
CF	-0,181	0,140	0,055	-0,197	-0,192	0,128	0,137	-0,208	0,152
CFC	-0,215	0,206	0,058	-0,262	-0,210	0,190	0,066	-0,267	0,083
entrop_norC	-0,097	-0,074	0,035	0,003	-0,088	-0,076	0,065	0,011	0,105
entrop_nor	-0,169	0,038	0,058	-0,116	-0,137	0,032	0,031	-0,102	0,092

NEP EEP	dbase	dpos	dneg	trp1	trn1	tgp	tgn	ang_pos	ang_neg
Amax	0,085	0,032	0,055	-0,093	-0,060	-0,053	0,013	<i>0,181</i>	<i>0,158</i>
Bmin	0,011	-0,016	0,054	0,067	<i>0,121</i>	0,020	-0,043	<i>-0,164</i>	<i>-0,140</i>
Tdif	0,052	-0,061	0,072	0,047	<i>0,189</i>	-0,021	0,027	<i>-0,168</i>	<i>-0,143</i>
DifAB	0,043	0,028	0,000	-0,095	<i>-0,109</i>	-0,043	0,033	0,206	<i>0,178</i>
Amax_pts	0,080	0,043	0,043	-0,092	-0,068	-0,051	0,011	0,202	<i>0,163</i>
Bmin_pts	0,005	-0,021	0,050	0,064	<i>0,116</i>	0,016	-0,037	<i>-0,148</i>	<i>-0,139</i>
Tdif_pts	<i>-0,135</i>	-0,086	-0,096	0,077	-0,013	0,013	-0,001	-0,252	<i>-0,146</i>
DifAB_pts	0,046	0,039	-0,005	-0,096	<i>-0,115</i>	-0,041	0,030	0,217	<i>0,188</i>
T	<i>-0,107</i>	-0,088	-0,011	0,003	0,050	-0,007	0,019	-0,224	-0,082
dbase	-0,014	-0,021	-0,012	0,010	0,018	-0,027	0,032	-0,037	-0,069
dpos	-0,021	0,034	-0,050	-0,021	-0,062	0,009	0,036	<i>0,109</i>	0,009
dneg	-0,049	-0,077	-0,023	0,060	0,072	-0,013	-0,016	-0,206	<i>-0,131</i>
trp1	0,087	0,070	0,035	-0,090	-0,069	-0,034	0,040	0,266	<i>0,153</i>
trn1	-0,014	-0,016	-0,037	-0,059	-0,063	-0,027	0,018	0,081	0,082
tgp	0,006	0,002	0,009	-0,025	-0,055	-0,002	-0,003	0,057	0,049
tgn	<i>0,124</i>	0,095	0,032	<i>-0,114</i>	-0,032	0,006	<i>-0,177</i>	<i>0,141</i>	0,091
ang_pos	-0,051	-0,020	-0,032	0,082	0,052	0,077	0,028	-0,133	<i>-0,148</i>
ang_neg	-0,060	-0,058	-0,010	<i>0,142</i>	<i>0,144</i>	-0,009	-0,051	-0,266	-0,225
ang_pico	-0,065	-0,046	-0,036	<i>0,135</i>	0,069	-0,004	0,007	-0,245	-0,205
Amax_i	-0,005	-0,041	-0,008	<i>-0,121</i>	-0,084	-0,034	-0,028	0,061	0,060
Bmin_i	-0,021	<i>0,103</i>	-0,087	<i>0,101</i>	-0,071	0,015	-0,019	-0,029	-0,035
DifA_i	0,091	0,078	0,053	-0,008	-0,011	-0,029	0,033	<i>0,172</i>	<i>0,131</i>
tA_i	-0,055	-0,044	0,015	<i>-0,105</i>	-0,055	-0,038	0,038	<i>0,107</i>	<i>0,132</i>
DifB_i	0,018	-0,084	<i>0,100</i>	-0,004	<i>0,151</i>	0,005	-0,022	<i>-0,118</i>	<i>-0,107</i>
tB_i	-0,012	-0,010	0,029	-0,036	0,014	0,090	-0,030	0,037	-0,011
Amax_f	0,058	0,043	-0,006	-0,054	-0,041	-0,047	0,067	<i>0,154</i>	<i>0,124</i>
Bmin_f	-0,049	-0,009	-0,020	-0,012	-0,064	0,041	-0,013	0,040	0,034
DifA_f	0,036	0,008	0,053	-0,053	-0,038	-0,013	-0,049	0,083	0,068
tA_f	0,054	-0,017	<i>0,115</i>	-0,005	<i>0,122</i>	-0,039	0,015	<i>-0,111</i>	<i>-0,057</i>
DifB_f	0,044	-0,014	0,068	0,077	<i>0,171</i>	-0,016	-0,029	<i>-0,187</i>	<i>-0,173</i>
tB_f	0,061	0,002	0,020	-0,041	0,065	-0,018	0,085	0,004	-0,063
desvio	-0,006	-0,005	-0,023	-0,093	-0,090	-0,044	0,028	<i>0,120</i>	<i>0,137</i>
desvioC	0,036	-0,049	0,040	-0,114	-0,003	-0,061	0,038	0,046	0,092
media	-0,007	-0,039	-0,003	<i>-0,120</i>	-0,045	-0,051	0,040	0,083	0,087
mediaC	0,058	-0,059	0,080	<i>-0,116</i>	0,077	-0,059	0,035	0,017	0,036
coef	0,015	0,051	-0,015	0,012	-0,095	0,034	-0,035	<i>0,134</i>	<i>0,127</i>
coefC	-0,055	0,024	-0,085	-0,011	-0,209	0,003	0,010	<i>0,165</i>	<i>0,174</i>
vari	0,043	0,042	-0,017	-0,049	-0,097	-0,035	0,033	<i>0,154</i>	<i>0,126</i>
variC	0,058	-0,025	0,042	-0,099	-0,026	-0,045	0,041	0,092	<i>0,101</i>
CF	<i>0,151</i>	<i>0,127</i>	0,044	-0,062	-0,099	-0,006	0,012	0,317	<i>0,178</i>
CFC	0,045	<i>0,123</i>	-0,032	-0,045	<i>-0,183</i>	0,047	-0,001	0,348	0,201
entrop_norC	0,059	-0,060	0,081	<i>-0,117</i>	0,076	-0,058	0,036	0,020	0,038
entrop_nor	-0,005	-0,037	-0,002	<i>-0,119</i>	-0,048	-0,051	0,040	0,087	0,090

EEP \ NEP	ang_pico	Amax_i	Bmin_i	DifA_i	tA_i	DifB_i	tB_i	Amax_f	Bmin_f
Amax	0,188	-0,150	0,182	-0,062	-0,039	-0,034	0,044	-0,158	0,125
Bmin	-0,173	0,161	-0,158	0,051	-0,005	-0,025	0,035	0,148	-0,070
Tdif	-0,175	0,139	-0,074	0,013	-0,083	-0,189	0,054	0,166	-0,141
DifAB	0,216	-0,186	0,203	-0,067	-0,020	-0,005	0,005	-0,183	0,116
Amax_pts	0,196	-0,157	0,178	-0,065	-0,058	-0,021	0,025	-0,172	0,136
Bmin_pts	-0,164	0,155	-0,149	0,053	-0,024	-0,035	0,053	0,145	-0,054
Tdif_pts	-0,187	0,163	-0,064	0,067	0,017	0,017	-0,005	0,096	0,002
DifAB_pts	0,223	-0,194	0,203	-0,073	-0,020	0,009	-0,018	-0,197	0,117
T	-0,181	0,098	-0,102	0,029	-0,001	0,005	-0,014	0,128	-0,054
dbase	-0,049	-0,004	-0,054	0,043	0,056	-0,004	0,026	-0,015	-0,008
dpos	0,045	-0,067	-0,005	0,034	0,037	0,060	0,045	-0,096	0,064
dneg	-0,145	0,136	-0,027	0,006	0,028	-0,088	-0,001	0,102	-0,070
trp1	0,212	-0,185	0,131	-0,051	-0,014	0,006	0,034	-0,204	0,121
trn1	0,100	-0,113	0,129	-0,061	0,027	-0,014	-0,022	-0,097	0,005
tgp	0,053	-0,051	0,063	-0,007	0,103	0,035	-0,013	-0,083	0,016
tgn	0,132	-0,195	0,143	-0,044	0,108	-0,023	0,011	-0,139	0,135
ang_pos	-0,166	0,123	-0,139	0,082	0,022	0,018	0,031	0,109	-0,129
ang_neg	-0,261	0,248	-0,186	0,056	-0,037	-0,070	0,063	0,228	-0,127
ang_pico	-0,236	0,181	-0,177	0,114	0,021	0,001	0,022	0,175	-0,118
Amax_i	0,084	-0,095	0,160	-0,048	-0,004	-0,042	0,099	-0,149	0,086
Bmin_i	-0,029	0,072	-0,119	0,057	0,018	0,152	-0,065	0,096	0,059
DifA_i	0,149	-0,098	0,072	-0,034	-0,060	0,010	-0,048	-0,074	0,083
tA_i	0,085	-0,139	0,147	-0,031	0,000	-0,011	-0,012	-0,096	0,029
DifB_i	-0,133	0,098	-0,063	0,013	-0,034	-0,127	0,089	0,074	-0,086
tB_i	0,002	-0,061	-0,027	0,074	-0,002	0,021	-0,002	-0,009	0,051
Amax_f	0,160	-0,109	0,099	-0,057	0,026	-0,011	0,012	-0,124	0,127
Bmin_f	0,003	0,023	-0,071	-0,034	0,056	0,146	0,071	-0,021	0,074
DifA_f	0,072	-0,075	0,107	-0,020	-0,087	-0,013	0,016	-0,078	0,036
tA_f	-0,100	0,070	-0,103	0,027	-0,118	-0,041	0,046	0,113	-0,183
DifB_f	-0,174	0,144	-0,100	0,083	-0,070	-0,151	-0,001	0,168	-0,115
tB_f	-0,014	-0,014	-0,012	-0,088	-0,047	-0,094	0,014	-0,057	-0,007
desvio	0,144	-0,146	0,180	-0,052	-0,018	-0,019	-0,033	-0,146	0,078
desvioC	0,087	-0,097	0,173	-0,062	-0,032	-0,106	0,014	-0,103	0,037
media	0,090	-0,134	0,161	-0,055	-0,010	-0,076	0,022	-0,150	0,029
mediaC	0,023	-0,073	0,133	-0,050	-0,042	-0,179	0,035	-0,078	-0,038
coef	0,155	-0,065	0,046	-0,024	-0,077	0,120	-0,039	-0,046	0,110
coefC	0,207	-0,110	0,106	-0,040	-0,011	0,201	-0,013	-0,125	0,204
vari	0,163	-0,129	0,152	-0,044	-0,045	0,019	0,013	-0,121	0,067
variC	0,124	-0,110	0,146	-0,076	-0,070	-0,063	0,040	-0,126	0,012
CF	0,277	-0,184	0,113	-0,084	-0,048	0,071	0,036	-0,191	0,136
CFC	0,294	-0,221	0,103	-0,064	-0,003	0,149	0,002	-0,213	0,149
entrop_norC	0,027	-0,075	0,133	-0,053	-0,047	-0,177	0,039	-0,081	-0,040
entrop_nor	0,095	-0,135	0,162	-0,056	-0,016	-0,072	0,026	-0,151	0,028

NEP	DifA_f	tA_f	DifB_f	tB_f	desvio	desvioC	media	mediaC	coef
Amax	-0,026	0,085	-0,020	0,023	-0,102	-0,117	-0,103	-0,133	-0,004
Bmin	0,048	-0,046	-0,078	-0,005	0,152	0,152	0,174	0,174	-0,026
Tdif	-0,052	0,035	-0,129	0,017	0,242	0,246	0,254	0,178	-0,003
DifAB	-0,044	0,078	0,035	0,017	-0,152	-0,161	-0,166	-0,184	0,014
Amax_pts	-0,017	0,080	-0,018	0,026	-0,116	-0,134	-0,119	-0,146	-0,004
Bmin_pts	0,047	-0,044	-0,097	0,006	0,155	0,148	0,165	0,161	-0,004
Tdif_pts	0,155	-0,031	-0,030	-0,092	0,036	0,028	0,030	0,072	0,050
DifAB_pts	-0,040	0,076	0,050	0,012	-0,169	-0,175	-0,177	-0,191	0,000
T	-0,036	0,036	-0,019	-0,040	0,042	0,070	0,074	0,137	-0,023
dbase	0,063	-0,022	-0,036	0,043	0,032	0,010	0,018	-0,001	0,028
dpos	0,092	-0,061	0,002	0,079	-0,058	-0,075	-0,058	-0,069	0,048
dneg	0,040	0,023	-0,050	-0,078	0,107	0,091	0,065	0,064	0,022
trp1	0,009	0,042	-0,011	0,081	-0,109	-0,132	-0,107	-0,160	0,014
trn1	-0,070	0,054	0,075	-0,017	-0,096	-0,093	-0,124	-0,121	0,002
tgp	0,053	-0,080	0,069	-0,034	-0,080	-0,076	-0,031	-0,037	-0,055
tgn	-0,108	-0,114	-0,046	0,038	-0,103	-0,094	-0,064	-0,076	-0,027
ang_pos	0,087	-0,054	-0,039	-0,005	0,091	0,105	0,104	0,173	0,011
ang_neg	0,051	0,003	-0,094	-0,015	0,213	0,213	0,208	0,223	0,000
ang_pico	0,098	-0,051	-0,019	0,007	0,140	0,142	0,158	0,166	-0,001
Amax_i	0,055	-0,033	-0,008	-0,026	-0,073	-0,109	-0,113	-0,156	0,074
Bmin_i	0,007	-0,048	0,023	0,139	-0,049	-0,010	-0,025	0,055	-0,009
DifA_i	-0,060	0,112	-0,014	0,048	-0,071	-0,063	-0,043	-0,040	-0,061
tA_i	-0,081	0,024	0,069	0,092	-0,091	-0,106	-0,126	-0,124	0,073
DifB_i	0,038	-0,010	-0,104	-0,081	0,173	0,142	0,168	0,114	0,001
tB_i	-0,001	-0,047	-0,045	0,094	-0,008	-0,016	-0,026	-0,011	0,019
Amax_f	-0,016	0,024	-0,057	0,092	-0,063	-0,071	-0,070	-0,093	0,023
Bmin_f	0,035	0,034	0,037	0,089	-0,102	-0,086	-0,079	-0,016	-0,032
DifA_f	-0,004	0,066	0,031	-0,054	-0,071	-0,084	-0,068	-0,076	-0,025
tA_f	-0,054	0,021	0,050	-0,113	0,112	0,156	0,199	0,168	-0,106
DifB_f	0,021	-0,073	-0,130	-0,064	0,243	0,223	0,235	0,182	0,021
tB_f	-0,016	0,104	-0,105	0,013	0,055	0,049	0,057	0,022	-0,006
desvio	-0,029	0,064	0,038	-0,019	-0,117	-0,126	-0,139	-0,154	0,013
desvioC	-0,038	0,066	-0,023	-0,022	-0,022	-0,036	-0,033	-0,093	0,015
media	-0,008	0,021	0,009	-0,011	-0,051	-0,078	-0,087	-0,130	0,030
mediaC	-0,031	0,047	-0,061	-0,033	0,076	0,048	0,050	-0,041	0,025
coef	-0,045	0,104	0,062	0,002	-0,153	-0,127	-0,132	-0,084	-0,023
coefC	0,001	0,049	0,106	0,060	-0,281	-0,259	-0,261	-0,185	-0,017
vari	-0,037	0,052	0,069	0,026	-0,130	-0,120	-0,114	-0,124	-0,016
variC	-0,034	0,070	0,028	0,002	-0,056	-0,054	-0,044	-0,095	-0,005
CF	-0,045	0,040	0,033	0,093	-0,169	-0,163	-0,140	-0,153	-0,026
CFC	-0,043	0,061	0,098	0,067	-0,256	-0,243	-0,259	-0,211	-0,018
entrop_norC	-0,031	0,050	-0,058	-0,031	0,073	0,046	0,049	-0,043	0,024
entrop_nor	-0,010	0,024	0,014	-0,010	-0,055	-0,080	-0,089	-0,131	0,028

EEP \ NEP	coefC	vari	variC	CF	CFC	entrop_norC	entrop_nor
Amax	0,031	-0,073	-0,106	-0,125	-0,083	-0,132	-0,101
Bmin	0,020	0,092	0,122	0,009	0,049	0,170	0,170
Tdif	0,110	0,253	0,253	-0,079	0,009	0,178	0,255
DifAB	0,006	-0,099	-0,136	-0,080	-0,079	-0,181	-0,162
Amax_pts	0,022	-0,080	-0,117	-0,122	-0,081	-0,144	-0,117
Bmin_pts	0,037	0,097	0,120	-0,007	0,052	0,158	0,161
Tdif_pts	-0,026	-0,023	-0,029	0,157	0,206	0,067	0,024
DifAB_pts	-0,010	-0,110	-0,147	-0,069	-0,082	-0,188	-0,174
T	-0,035	-0,009	0,006	0,125	0,068	0,133	0,067
dbase	0,019	0,021	0,014	0,016	0,051	-0,003	0,017
dpos	-0,002	-0,055	-0,057	-0,005	0,012	-0,070	-0,058
dneg	0,031	0,077	0,053	0,063	0,111	0,061	0,062
trp1	0,032	-0,058	-0,081	-0,136	-0,101	-0,157	-0,103
trn1	-0,020	-0,051	-0,080	0,009	-0,029	-0,119	-0,121
tgp	-0,038	-0,086	-0,059	0,037	0,048	-0,035	-0,031
tgn	0,000	-0,043	-0,068	-0,060	-0,042	-0,074	-0,062
ang_pos	-0,009	0,050	0,082	0,109	0,066	0,111	0,101
ang_neg	0,014	0,144	0,163	0,062	0,120	0,218	0,203
ang_pico	-0,027	0,078	0,115	0,132	0,130	0,161	0,153
Amax_i	0,104	-0,087	-0,119	-0,080	-0,012	-0,157	-0,114
Bmin_i	-0,130	-0,046	-0,023	0,158	0,058	0,053	-0,026
DifA_i	-0,054	-0,020	-0,037	-0,072	-0,078	-0,037	-0,040
tA_i	0,062	-0,075	-0,103	-0,091	-0,098	-0,122	-0,124
DifB_i	0,115	0,118	0,125	-0,105	0,012	0,112	0,165
tB_i	-0,004	-0,073	-0,065	-0,017	-0,058	-0,014	-0,031
Amax_f	0,006	-0,022	-0,043	-0,087	-0,105	-0,091	-0,065
Bmin_f	-0,098	-0,107	-0,083	0,163	0,070	-0,018	-0,080
DifA_f	0,019	-0,067	-0,089	-0,055	0,005	-0,077	-0,070
tA_f	-0,002	0,115	0,186	-0,002	-0,010	0,171	0,200
DifB_f	0,116	0,186	0,192	-0,136	0,000	0,180	0,232
tB_f	0,062	0,070	0,060	-0,044	0,010	0,023	0,058
desvio	0,008	-0,074	-0,115	-0,061	-0,075	-0,152	-0,137
desvioC	0,094	0,005	-0,029	-0,139	-0,105	-0,091	-0,032
media	0,069	-0,011	-0,061	-0,114	-0,089	-0,129	-0,086
mediaC	0,167	0,102	0,060	-0,192	-0,122	-0,039	0,051
coef	-0,091	-0,123	-0,112	0,073	0,019	-0,081	-0,128
coefC	-0,161	-0,258	-0,254	0,125	0,038	-0,184	-0,260
vari	-0,014	-0,096	-0,110	-0,046	-0,073	-0,123	-0,112
variC	0,056	-0,027	-0,040	-0,111	-0,100	-0,093	-0,043
CF	-0,025	-0,098	-0,096	-0,055	-0,068	-0,148	-0,134
CFC	-0,130	-0,189	-0,199	0,055	-0,016	-0,208	-0,254
entrop_norC	0,167	0,100	0,059	-0,192	-0,123	-0,041	0,050
entrop_nor	0,068	-0,013	-0,062	-0,114	-0,090	-0,130	-0,087

**Apêndice I – Valores das constantes e pesos das Funções
Discriminantes obtidas nas simulações**

Variáveis	lda_1	lda_2	lda_3	lda_4	lda_5	lda_6	lda_7	lda_8	lda_9	lda_10
Amax							-1,78	-1,83	-2,53	-2,63
Bmin										0,56
Tdif	-1,62	-1,76	-1,57	-1,76	-2,05	-1,82	-1,71	-1,59	-1,74	-1,72
DifAB							7,86	8,35	12,41	12,75
Amax_pts							4,06	3,85	4,26	4,49
Bmin_pts									-1,87	-1,98
Tdif_pts						-0,77	-2,01	-2,23	-1,77	-1,78
T									-1,09	-1,23
dbase										0,66
dpos						0,81	0,96	1,25	0,93	0,74
dneg										-0,75
trp							-4,04	-4,03	-4,01	-4,07
trn							-1,55	-1,89	-0,95	-0,69
tgp										0,04
tgn							-1,17	-1,21	-1,20	-1,25
ang_pos							1,88	1,68	2,37	2,57
ang_neg				1,38	1,37	1,96	5,37	5,55	5,60	6,61
ang_pico									-0,43	-1,27
Amax_i									0,20	0,16
Bmin_i						0,61	0,41	0,42	0,38	0,22
DifA_i	1,10	1,38	1,54	1,35	1,38	1,40	1,33	1,43	1,61	1,63
tA_i							0,49	0,51	0,63	0,66
DifB_i		1,35	1,11	1,20	1,20	1,53	1,41	1,42	1,24	1,12
tB_i									0,36	0,43
Amax_f		1,20	-0,82		-0,80	-0,73	-0,67	-0,57	-0,63	-0,68
Bmin_f									0,40	0,34
DifA_f							-0,15	-0,31	-0,30	-0,32
tA_f								0,21	0,32	0,33
DifB_f								0,26	0,44	0,44
tB_f					0,33	0,29	0,29	0,11	0,19	0,19
desvio										0,65
mediaC		2,93	2,82	2,93	3,12	2,69	3,05	3,28	3,24	3,38
coef	3,98	5,40	4,67	5,40	4,98	4,76	5,02	4,86	4,49	4,10
vari		-1,83	-1,81	-1,83	-1,58	-1,63	-1,87	-1,87	-1,30	-1,49
variC							-1,60	-1,93	-3,92	-4,33
CF									-4,83	-4,88
Constantes:	1,82	-1,23	2,05	-1,23	-1,00	-3,22	-16,96	-17,26	-18,06	-19,35

Apêndice J – Desempenho dos classificadores nas simulações com os registros da base D e marcações do especialista 1.

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D1	17	3.583	rpca_04	4	3.474	109	13	23,53%	96,96%	96,61%
			rma_09	4	3.423	160	13	23,53%	95,53%	95,19%
			rlda_10	5	3.432	151	12	29,41%	95,79%	95,47%
			lda_06	14	1.944	1.639	3	82,35%	54,26%	54,39%
D2	11	3.588	rpca_04	4	3.500	88	7	36,36%	97,55%	97,36%
			rma_09	2	3.516	72	9	18,18%	97,99%	97,75%
			rlda_10	0	3.587	1	11	0,00%	99,97%	99,67%
			lda_06	10	1.991	1.597	1	90,91%	55,49%	55,60%
D3	0	3.600	rpca_04	0	3.363	237	0	100,00%*	93,42%	93,42%
			rma_09	0	3.378	222	0	100,00%*	93,83%	93,83%
			rlda_10	0	3.321	279	0	100,00%*	92,25%	92,25%
			lda_06	0	1.922	1.678	0	100,00%*	53,39%	53,39%
D4	0	5.659	rpca_04	0	5.488	171	0	100,00%*	96,98%	96,98%
			rma_09	0	5.339	320	0	100,00%*	94,35%	94,35%
			rlda_10	0	5.266	393	0	100,00%*	93,06%	93,06%
			lda_06	0	2.255	3.404	0	100,00%*	39,85%	39,85%
D5	1	5.685	rpca_04	0	5.128	557	1	0,00%	90,20%	90,19%
			rma_09	0	4.947	738	1	0,00%	87,02%	87,00%
			rlda_10	1	4.877	808	0	100,00%	85,79%	85,79%
			lda_06	1	1.927	3.758	0	100,00%	33,90%	33,91%
D6	2	6.264	rpca_04	0	5.922	342	2	0,00%	94,54%	94,51%
			rma_09	0	5.797	467	2	0,00%	92,54%	92,52%
			rlda_10	0	5.751	513	2	0,00%	91,81%	91,78%
			lda_06	1	2.361	3.903	1	50,00%	37,69%	37,70%
D7	16	3.584	rpca_04	2	2.970	614	14	12,50%	82,87%	82,56%
			rma_09	2	2.889	695	14	12,50%	80,61%	80,31%
			rlda_10	1	2.935	649	15	6,25%	81,89%	81,56%
			lda_06	15	1.378	2.206	1	93,75%	38,45%	38,69%
D8	45	3.583	rpca_04	9	3.472	111	36	20,00%	96,90%	95,95%
			rma_09	5	3.428	155	40	11,11%	95,67%	94,63%
			rlda_10	6	3.400	183	39	13,33%	94,89%	93,88%
			lda_06	34	1.591	1.992	11	75,56%	44,40%	44,79%
D9	19	3.581	rpca_04	4	3.298	283	15	21,05%	92,10%	91,72%
			rma_09	1	3.221	360	18	5,26%	89,95%	89,50%
			rlda_10	1	3.251	330	18	5,26%	90,78%	90,33%
			lda_06	15	1.499	2.082	4	78,95%	41,86%	42,06%
D10	0	3.600	rpca_04	0	3.481	119	0	100,00%*	96,69%	96,69%
			rma_09	0	3.399	201	0	100,00%*	94,42%	94,42%
			rlda_10	0	3.423	177	0	100,00%*	95,08%	95,08%
			lda_06	0	1.208	2.392	0	100,00%*	33,56%	33,56%

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D11	2	4.317	rpca_04	1	4.056	261	1	50,00%	93,95%	93,93%
			rna_09	0	3.927	390	2	0,00%	90,97%	90,92%
			rlda_10	1	3.998	319	1	50,00%	92,61%	92,59%
			lda_06	2	1.474	2.843	0	100,00%	34,14%	34,17%
D12	0	4.009	rpca_04	0	3.424	585	0	100,00%*	85,41%	85,41%
			rna_09	0	3.214	795	0	100,00%*	80,17%	80,17%
			rlda_10	0	3.473	536	0	100,00%*	86,63%	86,63%
			lda_06	0	1.258	2.751	0	100,00%*	31,38%	31,38%
D13	44	3.556	rpca_04	12	3.427	129	32	27,27%	96,37%	95,53%
			rna_09	11	3.373	183	33	25,00%	94,85%	94,00%
			rlda_10	10	3.368	188	34	22,73%	94,71%	93,83%
			lda_06	41	1.378	2.178	3	93,18%	38,75%	39,42%
D14	91	3.508	rpca_04	31	3.425	83	60	34,07%	97,63%	96,03%
			rna_09	20	3.420	88	71	21,98%	97,49%	95,58%
			rlda_10	26	3.394	114	65	28,57%	96,75%	95,03%
			lda_06	82	1.283	2.225	9	90,11%	36,57%	37,93%
D15	37	3.549	rpca_04	1	3.432	117	36	2,70%	96,70%	95,73%
			rna_09	7	3.294	255	30	18,92%	92,81%	92,05%
			rlda_10	7	3.340	209	30	18,92%	94,11%	93,34%
			lda_06	32	1.458	2.091	5	86,49%	41,08%	41,55%
D16	0	5.171	rpca_04	0	5.115	56	0	100,00%*	98,92%	98,92%
			rna_09	0	5.112	59	0	100,00%*	98,86%	98,86%
			rlda_10	0	5.100	71	0	100,00%*	98,63%	98,63%
			lda_06	0	3.382	1.789	0	100,00%*	65,40%	65,40%
D17	16	5.598	rpca_04	0	5.397	201	16	0,00%	96,41%	96,13%
			rna_09	0	5.398	200	16	0,00%	96,43%	96,15%
			rlda_10	0	5.346	252	16	0,00%	95,50%	95,23%
			lda_06	9	3.686	1.912	7	56,25%	65,84%	65,82%
D18	78	5.607	rpca_04	0	5.585	22	78	0,00%	99,61%	98,24%
			rna_09	2	5.581	26	76	2,56%	99,54%	98,21%
			rlda_10	2	5.572	35	76	2,56%	99,38%	98,05%
			lda_06	53	3.728	1.879	25	67,95%	66,49%	66,51%
D19	1	3.599	rpca_04	0	3.558	41	1	0,00%	98,86%	98,83%
			rna_09	0	3.529	70	1	0,00%	98,06%	98,03%
			rlda_10	0	3.529	70	1	0,00%	98,06%	98,03%
			lda_06	1	1.507	2.092	0	100,00%	41,87%	41,89%
D20	2	3.246	rpca_04	0	3.220	26	2	0,00%	99,20%	99,14%
			rna_09	0	3.186	60	2	0,00%	98,15%	98,09%
			rlda_10	0	3.208	38	2	0,00%	98,83%	98,77%
			lda_06	2	1.317	1.929	0	100,00%	40,57%	40,61%
D21	1	2.644	rpca_04	0	2.501	143	1	0,00%	94,59%	94,56%
			rna_09	0	2.460	184	1	0,00%	93,04%	93,01%
			rlda_10	0	2.467	177	1	0,00%	93,31%	93,27%
			lda_06	1	1.237	1.407	0	100,00%	46,79%	46,81%

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D22	17	3.742	rpca_04	9	3.590	152	8	52,94%	95,94%	95,74%
			rma_09	3	3.511	231	14	17,65%	93,83%	93,48%
			rlda_10	3	3.546	196	14	17,65%	94,76%	94,41%
			lda_06	16	2.265	1.477	1	94,12%	60,53%	60,68%
D23	40	3.560	rpca_04	9	3.365	195	31	22,50%	94,52%	93,72%
			rma_09	5	3.232	328	35	12,50%	90,79%	89,92%
			rlda_10	4	3.313	247	36	10,00%	93,06%	92,14%
			lda_06	32	1.599	1.961	8	80,00%	44,92%	45,31%
D24	30	3.570	rpca_04	3	3.385	185	27	10,00%	94,82%	94,11%
			rma_09	1	3.315	255	29	3,33%	92,86%	92,11%
			rlda_10	2	3.361	209	28	6,67%	94,15%	93,42%
			lda_06	25	1.960	1.610	5	83,33%	54,90%	55,14%
D25	19	3.661	rpca_04	2	3.347	314	17	10,53%	91,42%	91,01%
			rma_09	1	3.244	417	18	5,26%	88,61%	88,18%
			rlda_10	2	3.252	409	17	10,53%	88,83%	88,42%
			lda_06	12	2.108	1.553	7	63,16%	57,58%	57,61%
D26	250	3.350	rpca_04	46	3.207	143	204	18,40%	95,73%	90,36%
			rma_09	22	3.177	173	228	8,80%	94,84%	88,86%
			rlda_10	30	3.178	172	220	12,00%	94,87%	89,11%
			lda_06	186	2.085	1.263	66	73,81%	62,28%	63,08%
D27	16	3.532	rpca_04	0	3.449	83	16	0,00%	97,65%	97,21%
			rma_09	0	3.358	174	16	0,00%	95,07%	94,64%
			rlda_10	0	3.392	140	16	0,00%	96,04%	95,60%
			lda_06	4	1.093	2.323	1	80,00%	32,00%	32,07%
D28	5	3.416	rpca_04	2	3.241	175	3	40,00%	94,88%	94,80%
			rma_09	1	3.173	243	4	20,00%	92,89%	92,78%
			rlda_10	1	3.113	303	4	20,00%	91,13%	91,03%
			lda_06	4	1.093	2.323	1	80,00%	32,00%	32,07%
D29	6	3.736	rpca_04	3	3.525	211	3	50,00%	94,35%	94,28%
			rma_09	2	3.494	242	4	33,33%	93,52%	93,43%
			rlda_10	2	3.414	322	4	33,33%	91,38%	91,29%
			lda_06	5	1.382	2.354	1	83,33%	36,99%	37,07%
D30	29	3.571	rpca_04	6	3.420	151	23	20,69%	95,77%	95,17%
			rma_09	5	3.403	168	24	17,24%	95,30%	94,67%
			rlda_10	4	3.371	200	25	13,79%	94,40%	93,75%
			lda_06	27	1.231	2.340	2	93,10%	34,47%	34,94%

[†] Identificação do registro de EEG da base de dados.

[§] Marcação refere-se à identificação dos paroxismos (EEP) e ao valor assumido para a quantidade de padrões não-epileptiformes (NEP) calculada subtraindo-se a quantidade de épocas contendo paroxismos da quantidade de segundos do registro.

^{*} Assume-se um valor de 100% de sensibilidade tanto o especialista quanto o classificador não identificaram paroxismos no registro de EEG analisado.

Apêndice K – Desempenho dos classificadores nas simulações com os registros da base D e marcações do especialista 2.

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D1	0	3.600	rpca_04	0	3.487	113	0	100,00%*	96,86%	96,86%
			rma_09	0	3.436	164	0	100,00%*	95,44%	95,44%
			rlda_10	0	3.444	156	0	100,00%*	95,67%	95,67%
			lda_06	0	1.947	1.653	0	100,00%*	54,08%	54,08%
D2	4	3.595	rpca_04	3	3.506	89	1	75,00%	97,52%	97,50%
			rma_09	1	3.522	73	3	25,00%	97,97%	97,89%
			rlda_10	0	3.594	1	4	0,00%	99,97%	99,86%
			lda_06	4	1.992	1.603	0	100,00%	55,41%	55,46%
D3	1	3.599	rpca_04	1	3.363	236	0	100,00%	93,44%	93,44%
			rma_09	0	3.377	222	1	0,00%	93,83%	93,81%
			rlda_10	0	3.320	279	1	0,00%	92,25%	92,22%
			lda_06	1	1.922	1.677	0	100,00%	53,40%	53,42%
D4	2	5.657	rpca_04	0	5.486	171	2	0,00%	96,98%	96,94%
			rma_09	0	5.337	320	2	0,00%	94,34%	94,31%
			rlda_10	0	5.264	393	2	0,00%	93,05%	93,02%
			lda_06	2	2.255	3.402	0	100,00%	39,86%	39,88%
D5	0	5.686	rpca_04	0	5.129	557	0	100,00%*	90,20%	90,20%
			rma_09	0	4.948	738	0	100,00%*	87,02%	87,02%
			rlda_10	0	4.877	809	0	100,00%*	85,77%	85,77%
			lda_06	0	1.927	3.759	0	100,00%*	33,89%	33,89%
D6	13	6.253	rpca_04	0	5.911	342	13	0,00%	94,53%	94,33%
			rma_09	1	5.787	466	12	7,69%	92,55%	92,37%
			rlda_10	1	5.741	512	12	7,69%	91,81%	91,64%
			lda_06	7	2.356	3.897	6	53,85%	37,68%	37,71%
D7	48	3.552	rpca_04	8	2.944	608	40	16,67%	82,88%	82,00%
			rma_09	6	2.861	691	42	12,50%	80,55%	79,64%
			rlda_10	3	2.905	647	45	6,25%	81,78%	80,78%
			lda_06	43	1.374	2.178	5	89,58%	38,68%	39,36%
D8	81	3.547	rpca_04	12	3.439	108	69	14,81%	96,96%	95,12%
			rma_09	8	3.395	152	73	9,88%	95,71%	93,80%
			rlda_10	10	3.368	179	71	12,35%	94,95%	93,11%
			lda_06	64	1.585	1.962	17	79,01%	44,69%	45,45%
D9	5	3.595	rpca_04	2	3.310	285	3	40,00%	92,07%	92,00%
			rma_09	2	3.236	359	3	40,00%	90,01%	89,94%
			rlda_10	1	3.265	330	4	20,00%	90,82%	90,72%
			lda_06	4	1.502	2.093	1	80,00%	41,78%	41,83%
D10	1	3.599	rpca_04	0	3.480	119	1	0,00%	96,69%	96,67%
			rma_09	0	3.398	201	1	0,00%	94,42%	94,39%
			rlda_10	0	3.422	177	1	0,00%	95,08%	95,06%
			lda_06	0	1.207	2.392	1	0,00%	33,54%	33,53%

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D11	6	4.313	rpca_04	1	4.052	261	5	16,67%	93,95%	93,84%
			rna_09	0	3.923	390	6	0,00%	90,96%	90,83%
			rlda_10	1	3.994	319	5	16,67%	92,60%	92,50%
			lda_06	5	1.473	2.840	1	83,33%	34,15%	34,22%
D12	0	4.009	rpca_04	0	3.424	585	0	100,00%*	85,41%	85,41%
			rna_09	0	3.214	795	0	100,00%*	80,17%	80,17%
			rlda_10	0	3.473	536	0	100,00%*	86,63%	86,63%
			lda_06	0	1.258	2.751	0	100,00%*	31,38%	31,38%
D13	55	3.545	rpca_04	12	3.416	129	43	21,82%	96,36%	95,22%
			rna_09	11	3.362	183	44	20,00%	94,84%	93,69%
			rlda_10	9	3.356	189	46	16,36%	94,67%	93,47%
			lda_06	47	1.373	2.172	8	85,45%	38,73%	39,44%
D14	138	3.461	rpca_04	33	3.380	81	105	23,91%	97,66%	94,83%
			rna_09	22	3.375	86	116	15,94%	97,52%	94,39%
			rlda_10	30	3.351	110	108	21,74%	96,82%	93,94%
			lda_06	122	1.276	2.185	16	88,41%	36,87%	38,84%
D15	78	3.508	rpca_04	3	3.393	115	75	3,85%	96,72%	94,70%
			rna_09	10	3.256	252	68	12,82%	92,82%	91,08%
			rlda_10	13	3.305	203	65	16,67%	94,21%	92,53%
			lda_06	63	1.448	2.060	15	80,77%	41,28%	42,14%
D16	0	5.171	rpca_04	0	5.115	56	0	100,00%*	98,92%	98,92%
			rna_09	0	5.112	59	0	100,00%*	98,86%	98,86%
			rlda_10	0	5.100	71	0	100,00%*	98,63%	98,63%
			lda_06	0	3.382	1.789	0	100,00%*	65,40%	65,40%
D17	0	5.614	rpca_04	0	5.413	201	0	100,00%*	96,42%	96,42%
			rna_09	0	5.414	200	0	100,00%*	96,44%	96,44%
			rlda_10	0	5.362	252	0	100,00%*	95,51%	95,51%
			lda_06	0	3.693	1.921	0	100,00%*	65,78%	65,78%
D18	46	5.639	rpca_04	0	5.617	22	46	0,00%	99,61%	98,80%
			rna_09	1	5.612	27	45	2,17%	99,52%	98,73%
			rlda_10	1	5.603	36	45	2,17%	99,36%	98,58%
			lda_06	30	3.737	1.902	16	65,22%	66,27%	66,26%
D19	3	3.597	rpca_04	0	3.556	41	3	0,00%	98,86%	98,78%
			rna_09	1	3.528	69	2	33,33%	98,08%	98,03%
			rlda_10	1	3.528	69	2	33,33%	98,08%	98,03%
			lda_06	3	1.507	2.090	0	100,00%	41,90%	41,94%
D20	4	3.244	rpca_04	0	3.218	26	4	0,00%	99,20%	99,08%
			rna_09	0	3.184	60	4	0,00%	98,15%	98,03%
			rlda_10	0	3.206	38	4	0,00%	98,83%	98,71%
			lda_06	4	1.317	1.927	0	100,00%	40,60%	40,67%
D21	0	2.645	rpca_04	0	2.502	143	0	100,00%*	94,59%	94,59%
			rna_09	0	2.461	184	0	100,00%*	93,04%	93,04%
			rlda_10	0	2.468	177	0	100,00%*	93,31%	93,31%
			lda_06	0	1.237	1.408	0	100,00%*	46,77%	46,77%

EEG [†]	EEP [§]	NEP [§]	Classificador	VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
D22	45	3.714	rpca_04	13	3.566	148	32	28,89%	96,02%	95,21%
			rna_09	5	3.485	229	40	11,11%	93,83%	92,84%
			rlda_10	9	3.524	190	36	20,00%	94,88%	93,99%
			lda_06	30	2.251	1.463	15	66,67%	60,61%	60,68%
D23	37	3.563	rpca_04	6	3.365	198	31	16,22%	94,44%	93,64%
			rna_09	5	3.235	328	32	13,51%	90,79%	90,00%
			rlda_10	3	3.315	248	34	8,11%	93,04%	92,17%
			lda_06	30	1.600	1.963	7	81,08%	44,91%	45,28%
D24	59	3.541	rpca_04	4	3.357	184	55	6,78%	94,80%	93,36%
			rna_09	2	3.287	254	57	3,39%	92,83%	91,36%
			rlda_10	3	3.333	208	56	5,08%	94,13%	92,67%
			lda_06	37	1.943	1.598	22	62,71%	54,87%	55,00%
D25	11	3.669	rpca_04	2	3.355	314	9	18,18%	91,44%	91,22%
			rna_09	1	3.252	417	10	9,09%	88,63%	88,40%
			rlda_10	2	3.260	409	9	18,18%	88,85%	88,64%
			lda_06	9	2.113	1.556	2	81,82%	57,59%	57,66%
D26	307	3.293	rpca_04	52	3.156	137	255	16,94%	95,84%	89,11%
			rna_09	27	3.125	168	280	8,79%	94,90%	87,56%
			rlda_10	37	3.128	165	270	12,05%	94,99%	87,92%
			lda_06	211	2.056	1.238	95	68,95%	62,42%	62,97%
D27	12	3.536	rpca_04	0	3.453	83	12	0,00%	97,65%	97,32%
			rna_09	0	3.362	174	12	0,00%	95,08%	94,76%
			rlda_10	0	3.396	140	12	0,00%	96,04%	95,72%
			lda_06	14	1.092	2.313	2	87,50%	32,07%	32,33%
D28	16	3.405	rpca_04	6	3.234	171	10	37,50%	94,98%	94,71%
			rna_09	6	3.167	238	10	37,50%	93,01%	92,75%
			rlda_10	5	3.106	299	11	31,25%	91,22%	90,94%
			lda_06	14	1.092	2.313	2	87,50%	32,07%	32,33%
D29	29	3.713	rpca_04	8	3.507	206	21	27,59%	94,45%	93,93%
			rna_09	6	3.475	238	23	20,69%	93,59%	93,03%
			rlda_10	7	3.396	317	22	24,14%	91,46%	90,94%
			lda_06	26	1.380	2.333	3	89,66%	37,17%	37,57%
D30	117	3.483	rpca_04	16	3.342	141	101	13,68%	95,95%	93,28%
			rna_09	18	3.328	155	99	15,38%	95,55%	92,94%
			rlda_10	19	3.298	185	98	16,24%	94,69%	92,14%
			lda_06	108	1.224	2.259	9	92,31%	35,14%	37,00%

[†] Identificação do registro de EEG da base de dados.

[§] Marcação refere-se à identificação dos paroxismos (EEP) e ao valor assumido para a quantidade de padrões não-epileptiformes (NEP), calculada subtraindo-se a quantidade de épocas contendo paroxismos da quantidade de segundos do registro.

^{*} Assume-se um valor de 100% de sensibilidade tanto o especialista quanto o classificador não identificaram paroxismos no registro de EEG analisado.

Apêndice L – Análise do desempenho de classificação dos registros de EEG da Base D utilizando o especialista 1 como referência e o especialista 2 no papel de classificador.

EEG	Marcação [§]		VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
	eep	nep							
D1	17	3.583	0	3.583	0	17	0,00%	100,00%	99,53%
D2	11	3.588	2	3.586	2	9	18,18%	99,94%	99,69%
D3	0	3.600	0	3.599	1	0	100,00%*	99,97%	99,97%
D4	0	5.659	0	5.657	2	0	100,00%*	99,96%	99,96%
D5	1	5.685	0	5.685	0	1	0,00%	100,00%	99,98%
D6	2	6.264	1	6.252	12	1	50,00%	99,81%	99,79%
D7	16	3.584	6	3.542	42	10	37,50%	98,83%	98,56%
D8	45	3.583	23	3.525	58	22	51,11%	98,38%	97,79%
D9	19	3.581	1	3.577	4	18	5,26%	99,89%	99,39%
D10	0	3.600	0	3.599	1	0	100,00%*	99,97%	99,97%
D11	2	4.317	0	4.311	6	2	0,00%	99,86%	99,81%
D12	0	4.009	0	4.009	0	0	100,00%	100,00%	100,00%
D13	44	3.556	23	3.524	32	21	52,27%	99,10%	98,53%
D14	91	3.508	78	3.448	60	13	85,71%	98,29%	97,97%
D15	37	3.549	30	3.501	48	7	81,08%	98,65%	98,47%
D16	0	5.171	0	5.171	0	0	100,00%*	100,00%	100,00%
D17	16	5.598	0	5.598	0	16	0,00%	100,00%	99,71%
D18	78	5.607	41	5.602	5	37	52,56%	99,91%	99,26%
D19	1	3.599	0	3.596	3	1	0,00%	99,92%	99,89%
D20	2	3.246	1	3.243	3	1	50,00%	99,91%	99,88%
D21	1	2.644	0	2.644	0	1	0,00%	100,00%	99,96%
D22	17	3.742	13	3.710	32	4	76,47%	99,14%	99,04%
D23	40	3.560	14	3.537	23	26	35,00%	99,35%	98,64%
D24	30	3.570	23	3.534	36	7	76,67%	98,99%	98,81%
D25	19	3.661	10	3.660	1	9	52,63%	99,97%	99,73%
D26	250	3.350	223	3.266	84	27	89,20%	97,49%	96,92%
D27	16	3.532	12	3.532	0	4	75,00%	100,00%	99,89%
D28	5	3.416	4	3.404	12	1	80,00%	99,65%	99,62%
D29	6	3.736	6	3.713	23	0	100,00%	99,38%	99,39%
D30	29	3.571	16	3.470	101	13	55,17%	97,17%	96,83%

[§] Marcação refere-se à identificação dos paroxismos (EEP) e ao valor assumido para a quantidade de padrões não-epileptiformes (NEP) calculada subtraindo-se a quantidade de épocas contendo paroxismos da quantidade de segundos do registro.

* Assume-se um valor de 100% de sensibilidade tanto o especialista quanto o classificador não identificaram paroxismos no registro de EEG analisado.

Apêndice M – Análise do desempenho de classificação dos registros de EEG da Base D utilizando o especialista 2 como referência e o especialista 1 no papel de classificador.

EEG	Marcação [§]		VP	VN	FP	FN	Sensibilidade	Especificidade	Eficiência
	EEP	NEP							
D1	0	3.600	0	3.583	17	0	100,00%*	99,53%	99,53%
D2	4	3.595	2	3.586	9	2	50,00%	99,75%	99,69%
D3	1	3.599	0	3.599	0	1	0,00%	100,00%	99,97%
D4	2	5.657	0	5.657	0	2	0,00%	100,00%	99,96%
D5	0	5.686	0	5.685	1	0	100,00%*	99,98%	99,98%
D6	13	6.253	1	6.252	1	12	7,69%	99,98%	99,79%
D7	48	3.552	6	3.542	10	42	12,50%	99,72%	98,56%
D8	81	3.547	23	3.525	22	58	28,40%	99,38%	97,79%
D9	5	3.595	1	3.577	18	4	20,00%	99,50%	99,39%
D10	1	3.599	0	3.599	0	1	0,00%	100,00%	99,97%
D11	6	4.313	0	4.311	2	6	0,00%	99,95%	99,81%
D12	0	4.009	0	4.009	0	0	100,00%*	100,00%	100,00%
D13	55	3.545	23	3.524	21	32	41,82%	99,41%	98,53%
D14	138	3.461	78	3.448	13	60	56,52%	99,62%	97,97%
D15	78	3.508	30	3.501	7	48	38,46%	99,80%	98,47%
D16	0	5.171	0	5.171	0	0	100,00%*	100,00%	100,00%
D17	0	5.614	0	5.598	16	0	100,00%*	99,71%	99,71%
D18	46	5.639	41	5.602	37	5	89,13%	99,34%	99,26%
D19	3	3.597	0	3.596	1	3	0,00%	99,97%	99,89%
D20	4	3.244	1	3.243	1	3	25,00%	99,97%	99,88%
D21	0	2.645	0	2.644	1	0	100,00%*	99,96%	99,96%
D22	45	3.714	13	3.710	4	32	28,89%	99,89%	99,04%
D23	37	3.563	14	3.537	26	23	37,84%	99,27%	98,64%
D24	59	3.541	23	3.534	7	36	38,98%	99,80%	98,81%
D25	11	3.669	10	3.660	9	1	90,91%	99,75%	99,73%
D26	307	3.293	223	3.266	27	84	72,64%	99,18%	96,92%
D27	12	3.536	12	3.532	4	0	100,00%	99,89%	99,89%
D28	16	3.405	4	3.404	1	12	25,00%	99,97%	99,62%
D29	29	3.713	6	3.713	0	23	20,69%	100,00%	99,39%
D30	117	3.483	16	3.470	13	101	13,68%	99,63%	96,83%

[§] Marcação refere-se à identificação dos paroxismos (EEP) e ao valor assumido para a quantidade de padrões não-epileptiformes (NEP) calculada subtraindo-se a quantidade de épocas contendo paroxismos da quantidade de segundos do registro.

* Assume-se um valor de 100% de sensibilidade porque tanto o especialista (referência) quanto o classificador não identificaram paroxismos no registro de EEG analisado.

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica, Departamento de Engenharia Elétrica e Eletrônica, do Centro Tecnológico da Universidade Federal de Santa Catarina, como requisito para obtenção do Título de Doutor em Engenharia Elétrica

Orientador: Fernando Mendes de Azevedo, Dr.Sc.

Florianópolis, 2015