

12^o Congresso de Inovação, Ciência e Tecnologia do IFSP - 2021

EXTRAÇÃO DE CARACTERÍSTICAS DE ELETROENCEFALOGRAMA BASEADAS NA TEORIA DE GRAFOS PARA SISTEMA DE PREVISÃO DE SURTOS EPILEPTICOS

Luís O. L. Amorim¹, Victor M. A. de Almeida², Miguel A. de A de Sousa³, Sara D. dos Santos⁴, Ricardo Pires⁵

¹Graduando em Engenharia Eletrônica, Bolsista PIBIFSP, IFSP, Câmpus São Paulo, luisotaviolamorim@gmail.com

²Graduando em Engenharia de Controle e Automação, Bolsista PIBIT, IFSP, Câmpus São Paulo, victormanuelalmeida98@gmail.com

³Doutor em Ciências, Universidade de São Paulo, angelo@ifsp.edu.br

⁴Doutora em Engenharia Elétrica, Universidade de São Paulo, sarad@ifsp.edu.br

⁵Doutor em Sistemas Automáticos e Microeletrônicos, Université Montpellier II, ricardo_pires@ifsp.edu.br

Área de conhecimento (Tabela CNPq): 3.13.01.01-0 Processamento de Sinais Biológicos.

RESUMO: A epilepsia é uma das doenças neurológicas mais comuns no mundo, caracterizada pela ocorrência de surtos. No seu estudo e diagnóstico, é utilizado eletroencefalograma (EEG), que é um registro da atividade elétrica cerebral. Há interesse em desenvolver sistemas capazes de prever a ocorrência daqueles surtos, sendo que vários trabalhos publicados que buscam encontrar a entrada do sinal de EEG no período pré-ictal, o qual precede imediatamente um surto. Dentro desses trabalhos, muitos não utilizam os dados de EEG diretamente mas sim características extraídas dessas informações. O trabalho encontrado na literatura em que foram apresentados os melhores resultados usou, dentre várias outras características do EEG, algumas delas obtidas por meio da Teoria de Grafos e do cálculo de correlação. Porém, essas características e a forma de extraí-las não foram apresentadas lá em detalhes, já que o foco daquele trabalho não era a extração. O presente trabalho teve como objetivo o domínio conceitual daquelas características e criação de programa que as extrai do EEG e que monta vetores a serem disponibilizados ao grupo de pesquisa em que se insere, objetivo este bem sucedido.

PALAVRAS-CHAVE: redes neurais; análise complexa de redes; epilepsia

EXTRACTION OF ELECTROENCEPHALOGRAM CHARACTERISTICS BASED ON GRAPH THEORY FOR EPILEPTICAL SEIZURES PREDICTION SYSTEM

ABSTRACT: Epilepsy is one of the most common neurological diseases in the world, characterized by the occurrence of seizures. In its study and diagnosis, the electroencephalogram (EEG), which is a record of the electrical activity of the brain, is used. There is interest in developing systems to predict the occurrence of those seizures, and there are several published works that seek to find the beginning of the EEG signal in the pre-ictal period, which immediately precedes a seizure. Within these works,

many do not use EEG data directly, but rather features extracted from this information. The previous studies with best results found in literature used, amongst other EEG features, some of them obtained by means of Correlation calculations and Graph Theory. Although, these features and the form of extraction was not shown in details since the work focus was not the extraction. This work aims the conceptual mastery of those features and creation of computer program for feature extraction of the EEG creating vectors that will be provided to the research group, this objective was fully completed.

KEYWORDS: neural networks; complex network analysis; epilepsy.

INTRODUÇÃO

A epilepsia é o distúrbio neurológico mais comum no planeta afetando em torno de 3% da população. Grande parte das vezes os surtos não tem causa definida e seus sintomas são: espasmos, alteração dos sentidos, emoções repentinas e anormais, e perda de consciência podendo variar em frequência de poucas vezes por ano até diversos surtos em um dia (KOHRMAN, 2007). Por isso, formas de previsão de surtos são estudadas visando garantir a autonomia e o bem estar dos pacientes com essa doença. No diagnóstico da epilepsia o Eletroencefalograma (EEG) é utilizado. Esse exame mede impulsos nervosos através de eletrodos posicionados na cabeça, para que o médico procure alterações nos sinais elétricos do paciente. O paciente passa por diversos estados: surto (estado ictal), normal (estado interictal) estado alterado que antecede o surto (estado pré-ictal) e após o surto, há um outro estado alterado (estado pós-ictal) antes de retornar ao estado normal. (SCARAMELLI et al., 2009)

Em Tsiouris et al. (2018) um sistema de previsão de surtos epilépticos é criado e apresenta resultados muito acima de todos antes obtidos utilizando uma rede neural do tipo *Long Short-Term Memory* (LSTM) (HOCHREITER; SCHMIDHUBER, 1997). Para isso foram utilizadas diversas características extraídas dos dados brutos de EEG como características do domínio do tempo, do domínio das frequências, de correlações e de Teoria de Grafos. Essas características são extraídas com 2 objetivos: o primeiro é diminuir o vetor de entrada da rede neural a ser treinada, já o segundo é selecionar os dados mais importantes para a classificação.

Nesse trabalho, fizemos a extração das características relacionadas à correlação e à teoria de grafos para que os vetores montados aqui pudessem ser utilizados como entradas para uma rede neural analisada em um outro estudo ainda em desenvolvimento.

MATERIAIS E MÉTODOS

Os programas desenvolvidos nesse estudo foram criados utilizando a linguagem de programação Python (FOUNDATION, 2021) e suas bibliotecas NetworkX (NETWORKX, 2021), Pandas (PANDAS, 2021) e Numpy (NUMPY, 2021).

Os dados utilizados foram obtidos do banco de dados *open source* PhysioNet especificamente da base CHB-MIT Scalp EEG Database (SHOEB, 2009) que foram baixados no formato *European Data Format* (EDF) e convertidos para *Comma Separated Values* (CSV) utilizando o software EDFbrowser (EDFBROWSER, 2021).

As características objetivadas em nosso estudo foram a máxima correlação cruzada absoluta entre canais de EEG dois a dois, tempo de decorrelação de um canal, coeficiente de agrupamento, centralidade da intermediação, excentricidade, eficiência local, eficiência global, diâmetro, raio e tamanho do caminho característico. (TSIOURIS et al., 2018)

Dessas a máxima correlação cruzada entre dois canais foi definida por Tsiouris et al. (2018) como:

$$C_{xy} = \max_{\tau} \left(\left| \frac{C_{xy}(\tau)}{\sqrt{C_x(0) \cdot C_y(0)}} \right| \right) \quad (1)$$

em que,

C_{xy} - Covariância entre x e y; $C_x(0)$ - Variância de x; τ - Defasagem entre os sinais variando de -5 a 5

A outra característica de correlação é o tempo de decorrelação e trata-se da constante n que torna o módulo da correlação entre $x[t]$ e $x[t+n]$ menor ou igual à 0,5. (JEON et al., 2019)

Cada um dos canais de EEG se tornou um nó do grafo a ser analisado e as ligações entre nós são ponderadas pelos valores obtidos nas correlações. A excentricidade de um nó do Grafo, como evidenciado por (URAKOV; TIMERYAEV, 2012) é a maior distância desse nó para qualquer outro nó do grafo, o raio trata-se da menor excentricidade e o diâmetro a maior excentricidade. (URAKOV; TIMERYAEV, 2012)

O coeficiente de agrupamento é a razão de triângulos ao redor de um nó definido pela equação 2, a centralidade da intermediação é a fração dos caminhos mais curtos que passam por um nó e está descrita na equação 3, a eficiência local, definida pela equação 4 é a média dos inversos dos caminhos mais curtos dos nós de um nó, já a versão global é a média dos inversos dos caminhos mais curtos como mostra a equação 5 e por fim, o tamanho do caminho característico, calculado com a equação 6 é a média dos caminhos mais curtos. (RUBINOV; SPORNS, 2010)

$$C = \frac{1}{n} \sum_{i \in \mathbb{N}} \frac{2t^i}{k_i(k_i - 1)} \quad (2)$$

em que,

n - Número de nós do grafo; t - Número de triângulos ao redor do nó i (média geométrica dos nós 3 a 3 incluindo o nó i); k_i - Grau do nó i (soma dos pesos das ligações entre um nó e o nó i).

$$b_i = \frac{1}{(n-1)(n-1)} \sum_{h, j \in \mathbb{N}, h \neq j, h \neq i, j \neq i} \frac{\rho_{hj}^{(i)}}{\rho_{hj}} \quad (3)$$

em que,

n - Número de nós no grafo; ρ_{hj} - Número de caminhos mais curtos entre h e j; $\rho_{hj}^{(i)}$ - Número de caminhos mais curtos entre h e j que passam por i

$$E_{loc}^w = \frac{1}{2k_i(k_i - 1)} \sum_{i \in \mathbb{N}} \sum_{j, h \in \mathbb{N}, j \neq i} (w_{ij} w_{ih} [d_{jh}^w(N_i)]^{-1})^{\frac{1}{3}} \quad (4)$$

em que,

k_i - Grau do nó i; w_{ij} - Peso da conexão entre os nós i e j; $d_{jh}^w(N_i)$ - Tamanho do caminho mais curto entre j e h que passa apenas por vizinhos de i

$$E^w = \frac{1}{n(n-1)} \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}, j \neq i} (d_{ij}^w)^{-1} \quad (5)$$

em que,

n - Número de nós no Grafo; d_{ij}^w = Tamanho do caminho mais curto entre os nós i e j

$$L^w = \frac{1}{n(n-1)} \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}, j \neq i} d_{ij}^w \quad (6)$$

em que,

n - Número de nós no Grafo; d_{ij}^w - Tamanho do caminho mais curto entre os nós i e j

RESULTADOS E DISCUSSÃO

Das características de Teoria de Grafos extraídas, há 4 locais (referentes a cada nó) e 4 globais (referentes ao grafo como um todo), dessa forma foram extraídas 76 no total. Com relação a correlações, há uma correlação cruzada para cada par de canais, ou seja 153 e um tempo de decorrelação para cada canal, ou seja 18. Dessa forma obtivemos 247 características no total.

As correlações cruzadas possuem sempre um valor muito baixo, a exemplo disso a Figura 1 é o mapa de calor das correlações de 5 segundos de EEG do paciente número 1, nele fica visível como esses valores não chegam nem a 0,5. Além disso os tempos de decorrelação também são pequenos, com valores máximos de 20 amostras o que equivale a aproximadamente 78 milissegundos.

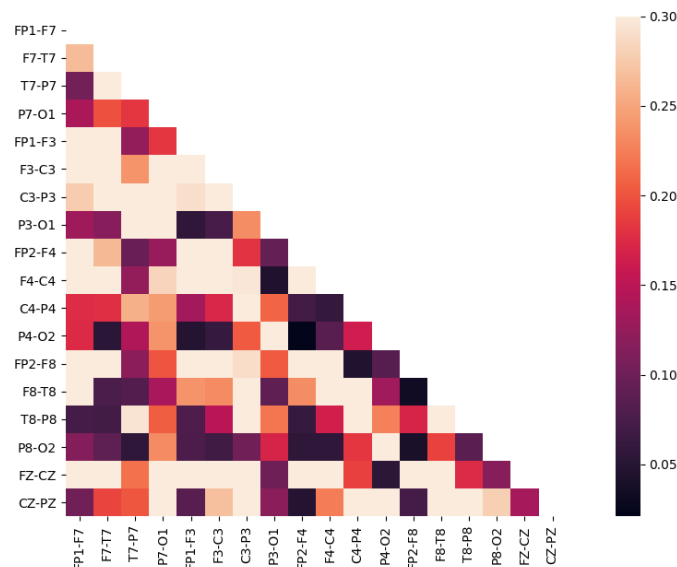


Figura 1: Mapas de calor das correlações cruzadas dos 5 primeiros segundos do primeiro paciente.

Devido aos baixos valores de correlações as conexões entre os nós ficam com baixos pesos e isso implica em um baixo tamanho de caminho característico, excentricidades, raio e diâmetro. Um problema comum em diversos exames de diversos pacientes é que, em certos momentos, os eletrodos acidentalmente soltam da pele do paciente fazendo com que o EEG fique constante tornando as variâncias dos canais nulas impossibilitando o cálculo das correlações.

CONCLUSÕES

O processo de extração de características com base na Teoria de Grafos aqui descrito foi bem sucedido, possibilitando a obtenção de vetores de características para sinais EEG de pacientes do banco de dados Physionet. O ponto negativo dos métodos utilizados neste trabalho é o tempo que demanda para a extração de tais características. Em alguns casos, cada hora de EEG levou 1 hora

para ser processado em um notebook, em uma aplicação real em que essa análise será feita em um dispositivo embarcado e em tempo real essa demora pode afetar e muito o uso dos dados aqui obtidos. Portanto, como trabalho futuro, propões-se a investigação de formas de redução desse tempo.

AGRADECIMENTOS

Os autores agradecem ao CNPq pelo apoio financeiro (Programa Institucional de Bolsas de Iniciação Científica e Tecnológica do Instituto Federal de Educação Ciência e Tecnologia de São Paulo PIBIFSP 2021)

REFERÊNCIAS

EDFBROWSER. *EDFbrowser*. 2021. Disponível em: <<https://www.teuniz.net/edfbrowser/>>. Acesso em: 20 de Agosto de 2021.

FOUNDATION, P. S. *3.9.6 Documentation*. 2021. Disponível em: <<https://www.python.org/>>. Acesso em: 20 de Agosto de 2021.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. *Neural Computation*, MIT Press Journals, v. 9, n. 8, p. 1735–1780, nov 1997. Acesso em: 20 de Agosto de 2021.

JEON, H.-J. et al. Laser speckle decorrelation time-based platelet function testing in microfluidic system. *Scientific Reports 2019 9:1*, Nature Publishing Group, v. 9, n. 1, p. 1–8, nov 2019. ISSN 2045-2322. Disponível em: <<https://www.nature.com/articles/s41598-019-52953-5>>. Acesso em: 20 de Agosto de 2021.

KOHRMAN, M. H. What is epilepsy? Clinical perspectives in the diagnosis and treatment. In: *Journal of Clinical Neurophysiology*. [s.n.], 2007. v. 24, n. 2, p. 87–95. ISSN 07360258. Disponível em: <https://journals.lww.com/clinicalneurophys/Fulltext/2007/04000/What__is__Epilepsy__Clinical__Perspectives__in__the.2.aspx>. Acesso em: 20 de Agosto de 2021.

NETWORKX. *NetworkX*. 2021. Disponível em: <<https://networkx.org/>>. Acesso em: 20 de Agosto de 2021.

NUMPY. *Num.Py*. 2021. Disponível em: <<https://numpy.org/>>. Acesso em: 20 de Agosto de 2021.

PANDAS. *Pandas - Python Data Analysis Library*. 2021. Disponível em: <<https://pandas.pydata.org/>>. Acesso em: 20 de Agosto de 2021.

RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, v. 52, n. 3, p. 1059–1069, sep 2010. ISSN 10538119. Acesso em: 20 de Agosto de 2021.

SCARAMELLI, A. et al. Prodromal symptoms in epileptic patients: Clinical characterization of the pre-ictal phase. *Seizure*, W.B. Saunders, v. 18, n. 4, p. 246–250, may 2009. ISSN 1059-1311. Acesso em: 20 de Agosto de 2021.

SHOEB, A. Application of machine learning to epileptic seizure onset detection and treatment. *Diss. Massachusetts Institute of Technology*, p. 157–162, 2009. Disponível em: <<http://dspace.mit.edu/handle/1721.1/54669>>. Acesso em: 20 de Agosto de 2021.

TSIOURIS, o. et al. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. *Computers in Biology and Medicine*, Elsevier Ltd, v. 99, p. 24–37, aug 2018. ISSN 18790534. Acesso em: 20 de Agosto de 2021.

URAKOV, A.; TIMERYAEV, T. Algorithms of Fast Search of Center, Radius and Diameter on Weighted Graphs. 2012. Acesso em: 20 de Agosto de 2021.