

Universidade Federal do ABC



Trabalho de Graduação

Análise de sinais de EEG para identificação de emoções

**Professora Dra. Aline Neves
Professor Dr. Denis Fantinato**

Vinícius Nasser Bernaldo - RA: 11201722677

**Santo André
2022**

VINICIUS NASSER BERNALDO

Análise de sinais de EEG para identificação de emoções

Trabalho de Graduação apresentado como requisito para a conclusão do curso de Engenharia de Informação pela Universidade Federal do ABC (UFABC).

Orientador: Aline de Oliveira Neves Panazio.

Coorientador: Denis Gustavo Fantinato.

SANTO ANDRÉ

2022

Resumo

Este trabalho tem como objetivo estudar o problema de separação cega de fontes e a análise de componentes independentes, aplicando-os em um caso prático, que é a remoção de artefatos em sinais de Eletroencefalograma (EEG). Como pretende-se classificar sinais relativos como referentes a emoções positivas ou negativas, é necessário encontrar a melhor maneira de tratá-los para que seja obtida a maior acurácia possível. Após a remoção de artefatos, foi obtido um sinal de EEG limpo, cujas features foram extraídas utilizando os parâmetros de Hjorth e *Higher Order Crossings* dos dados do sinal sem artefatos e como classificador *Support Vector Machine*. Este classificador, por sua vez, foi utilizado em conjunto com o algoritmo mRMR para redução da dimensionalidade dos dados, resultando em uma acurácia melhor quando utilizado o subconjunto ótimo que equivale a 10% do conjunto de todas as características calculadas.

Sumário

1	Introdução	4
2	Apresentação do projeto e objetivos.....	6
3	Análise de Componentes Independentes	7
3.1	Problema de separação cega de fontes	7
3.2	Branqueamento	8
3.3	FastICA	8
3.4	Infomax	10
3.5	Ambiguidades	11
3.6	Simulação de um problema de separação cega de fontes	11
4	Remoção de artefatos	15
4.1	<i>Multiple Artifact Rejection Algorithm (MARA)</i>	16
5	Extração, seleção de características e classificação	17
5.1	Parâmetros de Hjorth	17
5.2	Higher Order Crossings (HOC)	17
5.3	Minimum redundancy maximum relevance (mRMR)	20
5.4	Support Vector Machine (SVM)	21
6	Resultados e discussão	23
6.1	Descrição da base	23
6.2	Remoção de artefatos	24
6.3	Acurácia do classificador	26
7	Conclusão	37
8	Referências Bibliográficas	38
9	Anexos	40

1. Introdução

Os sinais de Eletroencefalograma (EEG) são aqueles resultantes da variação de potencial elétrico decorrente do acionamento dos neurônios, sendo emitidos entre a parte interna do neurônio e seu espaço externo e capturados por eletrodos do equipamento de nome homônimo aos sinais. Dessa forma, são representados os registros da variação das cargas elétricas sob a área de um determinado eletrodo de forma transiente. A intensidade do potencial elétrico captados no escalpo variam entre 0 e 200 microvolts, sua frequência oscila entre 0,3 Hz até pouco mais de 50 Hz. Além disso, é possível identificar em pessoas saudáveis padrões rítmicos conhecidos como ondas Alfa, Beta, Teta ou Delta [14].

As ondas Alfa são aquelas cuja frequência está entre 8 e 13Hz e são identificadas em pessoas adultas em estado de vigília submetidas a poucos estímulos, possuindo uma intensidade de 50 microvolts e desaparecem quando uma pessoa se encontra em sono profundo.

As ondas Beta, cuja frequência varia entre 14 e 30 Hz, são substitutas das ondas Alfa quando a atenção de um indivíduo está direcionada para a execução de tarefas mentais, sendo que estas ondas caracterizam-se como um sinal de maior frequência e menor amplitude elétrica.

As ondas Teta possuem frequência entre 4 e 7Hz e geralmente são detectadas em dois tipos de indivíduos: crianças e alguns adultos submetidos à estímulos de estresse emocional e frustrações.

As ondas Delta possuem uma frequência está entre 3,5Hz e 0,33 Hz, sendo características de um sono muito profundo.

Durante a elaboração do Trabalho de Graduação I, foi realizado um estudo sobre a teoria que envolve o problema de separação cega de fontes e sua resolução por meio da Análise de Componentes Independentes (ICA) e dois algoritmos clássicos relativos à ela: FastICA e Infomax.

Neste mesmo projeto, foi realizada uma simulação do problema de separação cega de fontes utilizando a linguagem de programação Python [15]. Nela, foram gerados dois vetores com valores aleatórios seguindo uma distribuição uniforme e misturados em seguida, sendo recuperados a partir da utilização dos algoritmos FastICA e Infomax.

Durante o Trabalho de Graduação II, foram aplicados os algoritmos de ICA no conceito de remoção de artefatos com o objetivo de melhorar a análise dos sinais de EEG contidos em uma base de dados. Para fazer isto, utilizamos o *software* MATLAB [16] e sua ferramenta EEGLab [17], onde foram aplicados os algoritmos de ICA para a separação de fontes e o algoritmo *Multiple Artifact Rejection Algorithm* (MARA) [6] para a remoção de artefatos.

Finalmente, o Trabalho de Graduação III procura consolidar todos os conceitos teóricos aprendidos durante os dois quadrimestres anteriores. Trabalhamos na extração de *features* através dos parâmetros de Hjorth [7] e *Higher Order Crossings* [9] para os sinais presentes na base de dados e os utilizamos para treinar um classificador do tipo Máquina de Vetores de Suporte (do inglês *Support Vector Machine* – SVM). Dessa forma, a fim de aumentar o desempenho deste classificador, foi empregado o conceito de *Minimum Redundancy - Maximum Relevance* (mRMR) [12] para selecionar as características que têm maior correlação com a classe desejada e menor correlação com as demais.

O trabalho está organizado da seguinte forma: a seção 2 irá definir os objetivos que procuram ser alcançados durante a realização do Trabalho de Graduação e como foram

coletados os dados que alimentam a base de dados que foi utilizada, bem como uma descrição de sua estruturação. As seções 3 e 4 apresentam uma explicação dos principais conceitos e algoritmos utilizados para este projeto em relação à Análise de Componentes Independentes e remoção de artefatos, respectivamente. A seção 5 relata a definição dos tipos de extração de características dos tipos parâmetros de Hjorth e *Higher Order Crossings*, além do modelo classificador *Support Vector Machine* e do método de seleção de características *Minimum Redundancy - Maximum Relevance*. Finalmente, a seção 6 exibe os resultados do projeto e uma análise dos mesmos.

2. Apresentação do projeto e objetivos

O enfoque deste trabalho é estudar e discutir algoritmos de Análise de Componentes Independentes e aplicá-los em um problema real de remoção de artefatos em sinais de EEG. Procuramos inicialmente simular, com o auxílio da linguagem de programação Python, um problema de separação cega de fontes com apenas duas fontes com distribuição uniforme, que serão misturadas e depois recuperadas a partir da utilização dos algoritmos de ICA FastICA e Infomax.

Em seguida, com o auxílio do *software* MATLAB e seu *plug-in* EEGLab, devemos manipular os sinais de EEG coletados e disponibilizados por [1] e selecionar aqueles que são decorrentes de emoções mais intensas por parte dos voluntários, realizando a remoção de artefatos de todos eles com o auxílio do algoritmo MARA.

Além disso, outro objetivo é estudar duas técnicas clássicas de extração de características para sinais de EEG: os parâmetros de Hjorth e *Higher Order Crossings*. Estas características, por sua vez, devem servir como entrada para um classificador que utiliza a técnica *Support Vector Machine* para realizar a classificação binária de um sinal como sendo referente a uma emoção positiva ou negativa.

3. Análise de Componentes Independentes

Nesta seção serão vistos alguns pontos relativos à Análise de Componentes Independentes: inicialmente vamos definir o que é um problema de separação cega de fontes, depois iremos discutir sobre uma técnica de tratamento dos dados chamada branqueamento. Finalmente, vamos abordar as motivações e características dos algoritmos FastICA e Infomax, além das ambiguidades que podem surgir após a sua aplicação. Finalmente, iremos mostrar os resultados de uma simulação de um problema de separação cega de fontes utilizando a linguagem de programação Python.

3.1. Problema de Separação Cega de Fontes

Em um cenário onde existe um conjunto de N fontes, que compõem um vetor $\mathbf{s} = [s_1, s_2, \dots, s_N]$ e M sinais captados pelos sensores, as amostras observadas formam um vetor $\mathbf{x} = [x_1, x_2, \dots, x_M]$, conhecido como mistura, que pode ser definido conforme a equação 1.

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

onde \mathbf{A} é uma matriz desconhecida de M linhas e N colunas, que é chamada de matriz de mistura.

Assim, um problema de separação cega de fontes consiste em recuperar um conjunto de sinais, denominados fontes, a partir de observações que são misturas desconhecidas das mesmas. O termo “cego”, por sua vez, é utilizado para explicitar que não é possível observar as fontes diretamente e que o sistema de mistura é desconhecido [2].

Uma abordagem possível para resolver este tipo de problema é a partir da Análise de Componentes Independentes (ICA - do inglês *Independent Component Analysis*), um conjunto de técnicas que procuram solucionar um problema de separação cega de fontes tentando obter o vetor \mathbf{y} , que é a melhor estimativa possível de \mathbf{s} . Esta aproximação é feita a partir da utilização de algoritmos que aproximem a matriz \mathbf{W} , que idealmente corresponde à matriz inversa da matriz de mistura, conforme mostrado na equação 2 [3].

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (2)$$

No entanto, como não existe nenhum conhecimento explícito sobre a matriz de mistura, deve-se criar suposições sobre as fontes. A hipótese central, que norteia a técnica ICA, é que todas as fontes possuem independência estatística entre si, isto é, a função densidade de probabilidade de \mathbf{s} pode ser escrita conforme a equação 3 [2].

$$\mathbf{p}(\mathbf{s}) = \mathbf{p}(s_1) \mathbf{p}(s_2) \dots \mathbf{p}(s_N) \quad (3)$$

onde $\mathbf{p}(s_i)$ é a função densidade de probabilidade da i -ésima fonte.

Além disso, outras suposições relevantes para ICA são que a matriz de mistura seja quadrada e de posto completo, além da não interferência de ruído, presença de no máximo uma variável com distribuição gaussiana e centralização dos dados, isto é, tenham média nula [3].

Dessa forma, existem diversos algoritmos capazes de realizar a separação de fontes, porém neste relatório serão tratados apenas dois deles: FastICA e Infomax, que foram os escolhidos para o desenvolvimento do projeto por serem algoritmos que são amplamente aceitos pela comunidade.

Finalmente, podemos resumir todo o processo que envolve um problema de separação cega de fontes a partir do diagrama de blocos representado pela Figura 1.



Figura 1: Diagrama de blocos

3.2. Branqueamento

Conforme mencionado em [4], uma maneira de simplificar a resolução de um problema de separação cega de fontes é trabalhar com vetores de mistura branqueados, isto é, aqueles cujos termos estão descorrelacionados e possuem variâncias unitárias. O processo de branqueamento, por sua vez, consiste em aplicar uma transformação linear V nos vetores de mistura originais, conforme mostra a equação 4.

$$\mathbf{z} = \mathbf{V}\mathbf{x} \quad (4)$$

onde \mathbf{z} é o vetor resultante do processo de branqueamento.

A transformação linear V pode ser definida a partir da equação 5.

$$\mathbf{V} = \mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T \quad (5)$$

onde D é a matriz diagonal dos autovalores da matriz de covariância de \mathbf{x} e E , a matriz dos autovetores.

A relevância do processo de branqueamento consiste no fato de que a matriz de mistura do vetor branqueado é ortogonal, isto é, sua inversa é igual à sua transposta. Dessa forma, a busca pela matriz desejada que resolva um problema de separação cega de fontes fica limitada ao universo das matrizes deste tipo, conforme demonstrado por [4].

3.3. FastICA

Uma das principais abordagens para a resolução de um problema de separação cega de fontes é a maximização da não gaussianidade dos sinais estimados. O porquê disso pode ser explicado a partir da equação 6 e sua descrição subsequente.

$$y_i = \mathbf{b}^T \mathbf{x} = \mathbf{b}^T \mathbf{A} \mathbf{s} = \mathbf{q}^T \mathbf{s} = \sum_{i=1}^N q_i s_i \quad (6)$$

onde \mathbf{b}^T é uma das linhas da inversa da matriz de mistura [4].

Pelo Teorema Central do Limite, o resultado da soma de duas variáveis aleatórias é, sob algumas condições, mais gaussiano que elas. Assim, pela análise da equação 6, nota-se que cada um dos sinais recuperados é uma combinação linear dos sinais originais, o que indica que y_i tem a sua não gaussianidade maximizada conforme se aproxima de uma das fontes, isto é, apenas um dos termos do somatório é não nulo [4]. Ainda que o teorema pressuponha que as variáveis aleatórias sendo somadas sejam independentes e identicamente distribuídas, o que não é o caso dos sinais de mistura usados aqui, a aplicação do teorema se mostra como bastante eficiente na resolução do problema de separação de fontes.

Isto torna a não gaussianidade interessante no contexto de resolução de um problema de separação cega de fontes. Assim, é preciso definir uma maneira de mensurar o quão próximo \mathbf{x} está de uma distribuição gaussiana, sendo que uma maneira para a tal é a partir do conceito de negentropia.

A negentropia é definida como quão afastado um conjunto de valores aleatórios está de uma distribuição normal, conforme mostrado na equação 7. Como a distribuição gaussiana caracteriza-se por ser aquela que possui a maior entropia, a negentropia torna-se um parâmetro que possibilita medir a não gaussianidade de uma variável aleatória.

$$J(\mathbf{x}) = H(\mathbf{x}_{Gauss}) - H(\mathbf{x}) \quad (7)$$

onde $J(\mathbf{x})$ é a negentropia de \mathbf{x} , $H(\mathbf{x}_{Gauss})$ é a entropia do vetor com distribuição gaussiana cuja matriz de covariância seja igual à de \mathbf{x} e $H(\mathbf{x})$ é a entropia de \mathbf{x} .

Note que, como a entropia da distribuição gaussiana é a maior possível, a negentropia nunca pode ser negativa. Dessa forma, quanto mais a negentropia se aproximar de zero, mais gaussiano é o sinal [4].

No entanto, a negentropia do sistema não é uma grandeza fácil de ser determinada, pois muitas vezes é necessário utilizar cumulantes de ordem superior para o seu cálculo, o que torna o método de determinação muito sofisticado e custoso, sendo então necessário buscar maneiras para realizar aproximações desta grandeza.

Assim, o algoritmo FastICA busca maximizar a aproximação da negentropia baseando-se no uso de funções não quadráticas, fornecendo uma alternativa aos cumulantes de ordem superior. O pseudocódigo referente a este algoritmo é mostrado a seguir.

Algoritmo FastICA

Enquanto a convergência não for atingida faça

para $i = 1; i \leq M; i \leftarrow i + 1$

$$w_i \leftarrow E\{z g(w_i^T z)\} - E\{g'(w_i^T z)\} w \quad (8a)$$

$$W \leftarrow (W W^T)^{-0.5} W \quad (8b)$$

onde w_i é o i -ésimo vetor-coluna de W e g é uma função de ativação (geralmente definida como a tangente hiperbólica) e z é o vetor \mathbf{x} após a etapa de branqueamento.

Uma maneira de verificar a convergência é analisar os valores de w_i antes e depois da nova iteração, calculando o valor absoluto do erro entre os dois. Se este resultado for menor ou igual a um limiar previamente definido, é possível considerar que a convergência foi atingida.

3.4. Infomax

Uma abordagem alternativa para a determinação de \mathbf{y} seria pensar em alguma maneira para maximizar a verossimilhança entre os seus valores e o de \mathbf{s} , ou seja, o conjunto de valores de \mathbf{y} que tenham a maior probabilidade possível de pertencerem a \mathbf{s} . No entanto, como estes valores podem assumir qualquer valor real, este procedimento envolve analisar uma função de distribuição de probabilidade (*pdf*) contínua para os valores observados, o que dificulta qualquer tipo de análise [4].

Dessa forma, o algoritmo Infomax retoma o princípio do FastICA de rumar iterativamente em uma direção interessante, porém desta vez procurando maximizar a informação mútua entre a entrada e a saída. Esta grandeza, por sua vez, está definida na equação 9.

$$I(x, y) = H(y) - H(y | x) \quad (9)$$

onde $I(x, y)$ é a informação mútua \mathbf{x} e \mathbf{y} , $H(y)$ é a entropia de \mathbf{y} e $H(y | x)$ é a entropia condicional de \mathbf{y} dado \mathbf{x} .

Assim, como a entropia pode ser enxergada como uma medida da incerteza de uma variável aleatória, a maximização da informação mútua pode ser vista como a redução da incerteza da entrada após a observação da saída [3], que corresponde ao único conjunto de valores disponíveis em um problema de separação cega de fontes.

Para este algoritmo, [4] descreve a sequência de etapas que caracterizam o algoritmo Infomax e estas são reproduzidas abaixo.

Algoritmo Infomax

Inicialize a matriz de \mathbf{W} com valores aleatórios

Enquanto a convergência não for atingida faça

$$\mathbf{y} \leftarrow \mathbf{W}\mathbf{x} \quad (10a)$$

$$\beta_i \leftarrow -E\{y_i g(y_i)\} \quad (10b)$$

$$\alpha_i \leftarrow \frac{-1}{\beta_i + E\{g'(y_i)\}} \quad (10c)$$

$$\mathbf{W} \leftarrow \mathbf{W} + \text{diag}(\alpha_i)[\text{diag}(\beta_i) - E\{g(y_i)\mathbf{y}^T\}]\mathbf{W} \quad (10d)$$

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{C}\mathbf{W}^T)^{-0,5}\mathbf{W} \quad (10e)$$

sendo que g é uma função de ativação (geralmente definida como a tangente hiperbólica) e y_i é a i -ésima linha de \mathbf{y} .

3.5. Ambiguidades

Após a determinação de \mathbf{W} e cálculo de \mathbf{y} , é possível que o sinal reconstruído seja diferente do que é desejado. Isto ocorre pelo fenômeno de ambiguidade, que pode ser de ganho ou ordem [4].

A ambiguidade de ganho ocorre quando o sinal recuperado se comporta da mesma maneira que o sinal original, porém apresentando uma amplitude diferente. Isto ocorre pelo fato das fontes e da matriz de mistura serem desconhecidas, o que possibilita que haja a divisão de uma das colunas da matriz de mistura por um escalar qualquer que multiplica o vetor com as amostras de uma das fontes. Uma forma de contornar este problema é assumir que cada fonte possui variância unitária, o que força uma adaptação da matriz para se ajustar a esta nova restrição.

A ambiguidade de ordem, por sua vez, é definida como um fenômeno onde os sinais não são recuperados na mesma ordem. Isto ocorre pelo mesmo motivo da outra ambiguidade, isto é, como há o desconhecimento tanto das fontes como da matriz de mistura, podemos atribuir qualquer ordem para as fontes.

3.6. Simulação de um problema de separação cega de fontes

A partir dos conceitos abordados nesta seção, simulamos um problema de separação cega de fontes utilizando a linguagem de programação Python.

O primeiro passo para a simulação é a geração de dois vetores cujos dados seguem uma distribuição uniforme. Os vetores foram centralizados e normalizados, misturados e branqueados, sendo que três gráficos a seguir mostram as distribuições dos dados no plano cartesiano em cada uma destas etapas.

A Figura 2 mostra os vetores aleatórios s_1 e s_2 .

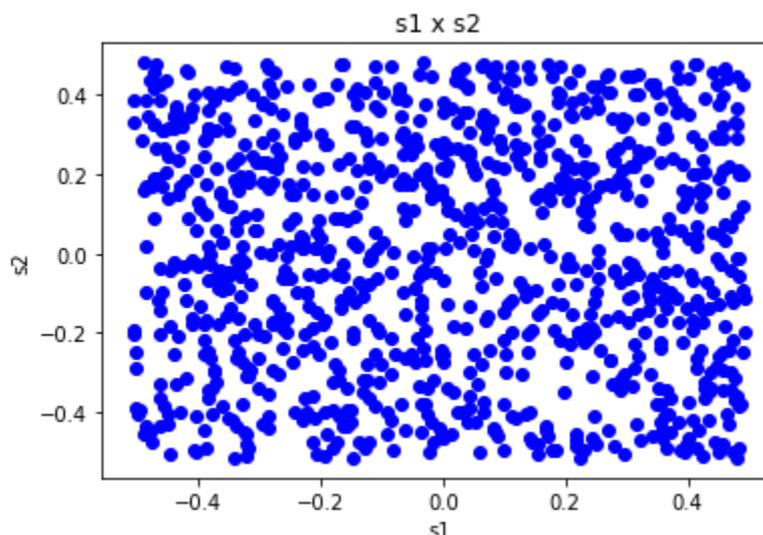


Figura 2: $s_1 \times s_2$

A partir da análise da distribuição conjunta entre os dois vetores aleatórios, percebe-se que os dados estão dispostos de maneira uniforme no plano e são independentes entre si, não sendo possível estimar o valor de uma variável a partir da observação da outra.

A Figura 3, por sua vez, mostra a distribuição dos dados relativos aos vetores x_1 e x_2 , que são resultantes da multiplicação entre as fontes e a matriz de mistura, que foi inicializada com valores aleatórios.

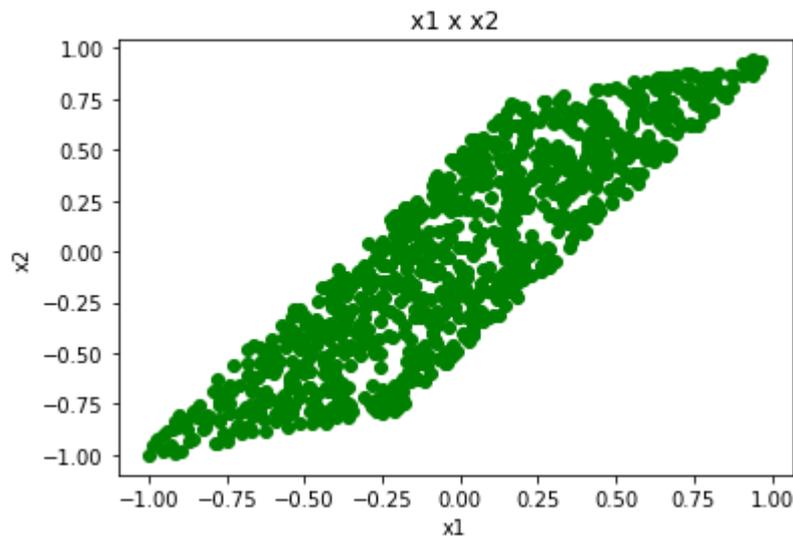


Figura 3: $x_1 \times x_2$

Podemos observar que os dados estão organizados em torno de uma reta cujo crescimento se dá na direção do primeiro quadrante. Assim, continua existindo uma correlação entre x_1 e x_2 , porém estes não são independentes: note que é possível estimar o valor de x_1 a partir do que é observado de x_2 , principalmente quando estes valores estão nas bordas do paralelogramo, cujas arestas correspondem as colunas da matriz de mistura [4].

Finalmente, a Figura 4 mostra a distribuição dos dados dos sinais resultantes do branqueamento dos vetores x_1 e x_2 , z_1 e z_2 .

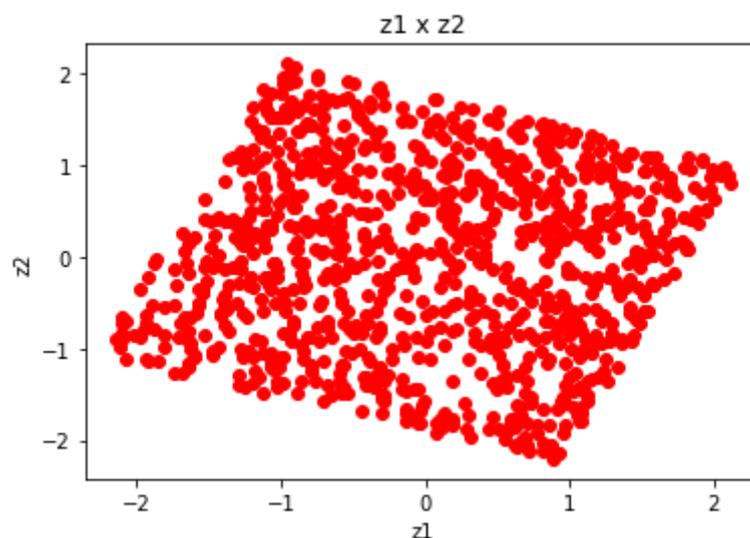


Figura 4: $z_1 \times z_2$

Este gráfico, por sua vez, mostra a característica do gráfico para sinais branqueados, onde a distribuição conjunta volta a se assemelhar a um quadrado, porém este com um ângulo de rotação, o que mostra que os sinais estão descorrelacionados e são independentes, restando apenas a necessidade de estimar o ângulo que causa a rotação [4].

Isto mostra como os dados são descorrelacionados mas as fontes ainda não foram separadas, o que só será possível com a aplicação de um dos métodos de ICA.

Na sequência, foram aplicados tanto o algoritmo FastICA como Infomax sobre os vetores de mistura, a fim de recuperar os sinais originais. Os sinais resultantes, por sua vez, foram normalizados e plotados em conjunto com os sinais originais, conforme mostrado nas Figuras abaixo, que demonstram o valor da n-ésima posição de cada uma das cinquenta primeiras posições dos vetores.

As Figuras 5 e 6 mostram os sinais originais plotados em conjunto com os sinais recuperados utilizando o algoritmo FastICA para a separação de fontes.

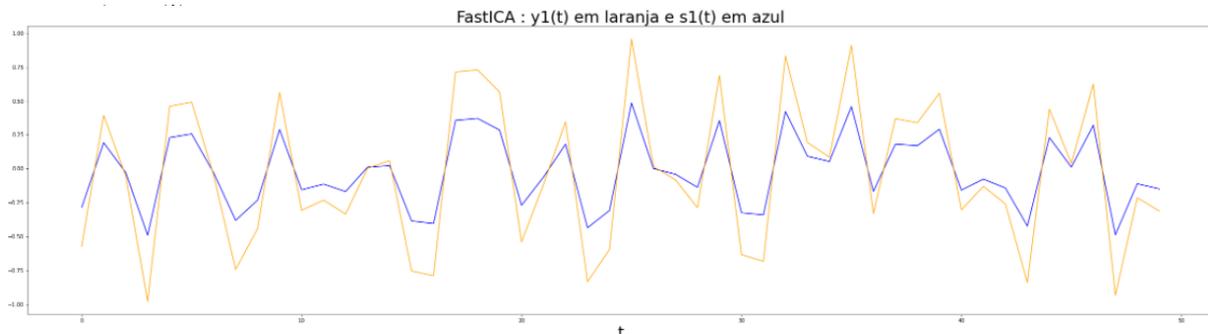


Figura 5: Vetor original s_1 e recuperado y_1 pelo algoritmo FastICA

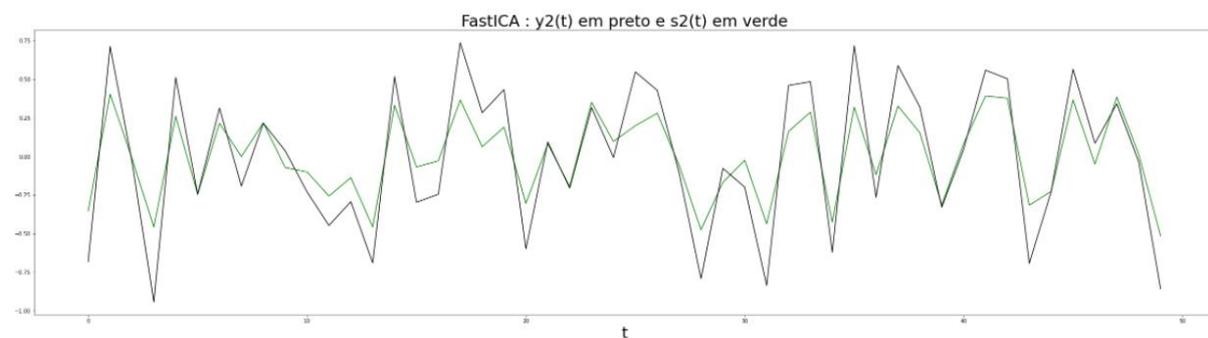


Figura 6: Vetor original s_2 e recuperado y_2 pelo algoritmo FastICA

Como os sinais originais e recuperados estão normalizados, inexistente a ambiguidade de ganho neste caso. Além disso, também não houve ambiguidade de ordem para a recuperação dos dois vetores originais.

Assim como as duas Figuras anteriores, as Figuras 7 e 8 também mostram os sinais originais plotados em conjunto com os sinais recuperados, porém desta vez utilizando o algoritmo Infomax ao invés do FastICA para a separação de fontes.



Figura 7: Vetor original s_1 e recuperado y_2 pelo algoritmo Infomax

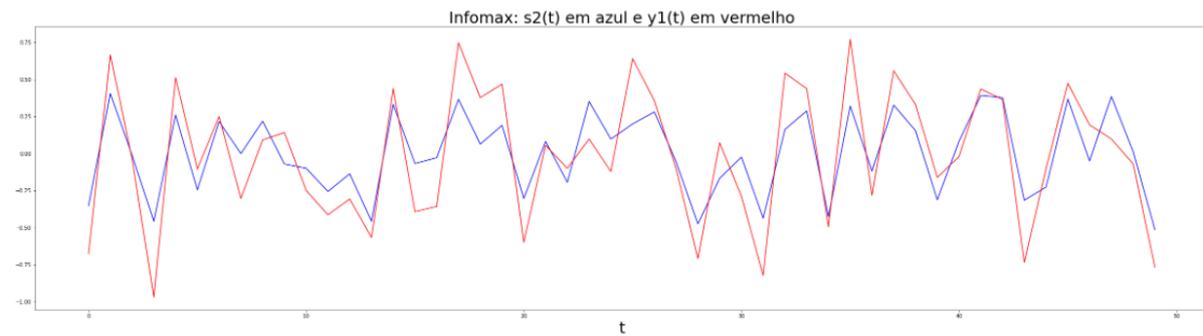


Figura 8: Vetor original s_2 e recuperado $-y_1$ pelo algoritmo Infomax

Diferentemente da recuperação dos vetores originais a partir do algoritmo FastICA, a utilização do algoritmo Infomax resultou em uma ambiguidade de ordem, ou seja, o segundo vetor recuperado refere-se ao primeiro vetor original e o primeiro vetor recuperado refere-se ao segundo vetor original. Além disso, o primeiro vetor recuperado apresenta um caráter espelhado em relação ao original, sendo necessário então multiplicá-lo por -1.

4. Remoção de Artefatos

Os sinais de EEG podem ser vistos como um problema de separação cega de fontes, porque, após a utilização de um eletroencefalograma para a coleta de sinais, o usuário tem em mãos apenas uma mistura dos sinais captados por cada um dos eletrodos, sem ter informação alguma sobre os mesmos. Dessa forma, não há inicialmente informação sobre quais foram os canais que obtiveram comportamento adequado, se houve algum defeito no equipamento durante a coleta ou se algum eletrodo captou apenas atividades decorrentes de fenômenos fisiológicos como espirros, tosse ou movimentos oculares, o que contamina o sinal colhido.

Neste contexto, um artefato pode ser definido como um sinal indesejado que contamina a qualidade dos dados coletados, podendo ser do tipo intrínseco ou extrínseco, conforme definido por [5].

Os artefatos intrínsecos são compostos por todos os fenômenos fisiológicos que podem ocorrer no voluntário durante a coleta, como movimentos oculares e piscadas, que foram propagados e acabam interferindo na coleta dos dados de EEG. Também é possível ressaltar os movimentos musculares em regiões próximas onde os eletrodos estão conectados, que podem ocorrer quando o voluntário espirrar, tossir, falar ou engolir saliva.

Os artefatos extrínsecos, por sua vez, correspondem às fontes de interferência que podem surgir dentro do ambiente de coleta, como o posicionamento equivocado de um dos eletrodos para coleta ou movimento dos cabos.

A Figura 9 mostra um exemplo da presença de artefatos em sinais de EEG: um conjunto de impulsos que se repetem periodicamente, que podem ser referentes a um conjunto de piscadas, batimentos cardíacos ou defeitos no equipamento.

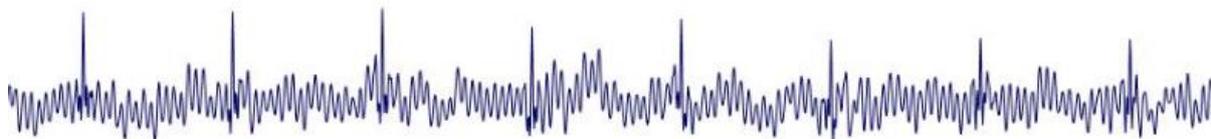


Figura 9: Exemplo da presença de artefatos em um sinal de EEG

A Figura 10, por sua vez, mostra um exemplo de um sinal de EEG sem a presença evidente de artefatos.



Figura 10: Exemplo de um sinal de EEG sem a presença de artefatos

4.1. *Multiple Artifact Rejection Algorithm (MARA)*

A identificação e remoção de um artefato nem sempre é uma tarefa fácil de ser concluída manualmente, o que pode tornar a análise de um sinal de EEG demorada e ineficiente.

Além disso, não podem ser descartados os equívocos por parte do analista que podem ocorrer durante este processo, o que pode ser prejudicial para a base de dados, como a remoção de um sinal que *a priori* se assemelhe a um artefato, mas não seja, o que acarreta em perda de informação ou a não detecção de dados ruins e manutenção de artefatos, mantendo-a contaminada. Dessa forma, a existência de um algoritmo capaz de detectar artefatos de forma automática durante a análise de um sinal de EEG seria capaz de otimizar o tempo e a qualidade dos estudos desta área.

Neste contexto, é interessante utilizar o algoritmo MARA (do inglês - *Multiple Artifact Rejection Algorithm*), que é um *plug-in open-source* do EEGLAB, que automatiza o processo de análise das componentes independentes que correspondem a artefatos intrínsecos e extrínsecos, como movimentos oculares, artefatos musculares ou eletrodos soltos. Esta ferramenta é resultado de um algoritmo de *Machine Learning* de aprendizado do tipo supervisionado que foi treinado utilizando-se seis *features* dos domínios espacial, espectral e temporal a partir de uma base de dados contendo 1290 tipos de artefatos diferentes [6]. Este algoritmo recebe como entrada o conjunto das componentes independentes de um sinal de EEG e retorna uma lista de quais delas correspondem a artefatos.

Dessa forma, a relevância deste algoritmo para este projeto é sua utilização dentro do EEGLAB para a remoção de artefatos para cada um dos sinais de EEG que foram selecionados, sendo que, após a detecção de um artefato em um canal, o sinal presente no mesmo é reconstruído sem a presença dos dados indesejados [6].

5. Extração, seleção de características e classificação

Nesta seção iremos discutir dois tipos de características muito utilizadas na literatura para a análise dos sinais de EEG: os parâmetros de Hjorth e *Higher Order Crossings*. Na sequência, iremos abordar as principais características do tipo de classificador SVM e como ele foi aplicado neste trabalho.

5.1. Parâmetros de Hjorth

Os parâmetros de Hjorth são maneiras de indicar propriedades estatísticas de um sinal qualquer $h(t)$ em sua forma temporal, sendo de extrema utilidade devido ao seu baixo custo computacional. [7]

O primeiro parâmetro é conhecido como Atividade, que é definido pela equação 11.

$$\text{Atividade} = \text{var}(h(t)) \quad (11)$$

sendo que $\text{var}(h(t))$ é a variância do sinal $h(t)$.

O segundo parâmetro é chamado de Mobilidade, que é definido pela equação 12.

$$\text{Mobilidade} = \sqrt{\frac{\text{var}(h'(t))}{\text{var}(h(t))}} \quad (12)$$

onde $h'(t)$ é a primeira derivada do sinal $h(t)$.

Finalmente, o último parâmetro é a Complexidade, que é calculado a partir da equação 13.

$$\text{Complexidade} = \frac{\text{Mobilidade}(h'(t))}{\text{Mobilidade}(h(t))} \quad (13)$$

Durante a aplicação deste projeto, foi criado um código na linguagem de programação Python que calcula os parâmetros de Atividade, Mobilidade e Complexidade para cada uma das dezenove componentes independentes de cada um dos sinais de EEG selecionados. Dessa forma, para cada sinal é calculado um total de 57 características.

5.2. Higher Order Crossings (HOC)

Conforme mencionado anteriormente, os sinais de EEG são transientes, ocorrendo sucessivas oscilações conforme o progresso temporal e, após a amostragem, podem ser escritos como uma série temporal finita $[F_1, F_2, \dots, F_n]$, onde F_i corresponde à i -ésima amostra do sinal e n , ao número total de amostras. Dessa forma, segundo [9], uma maneira de caracterizar os sinais seria contando a quantidade de vezes que cada sinal cruza o zero no eixo temporal utilizando diferentes filtros que, quando aplicados para um determinado sinal, alteram suas oscilações e, conseqüentemente, a quantidade de cruzamentos. Assim, o processo de *Higher Order Crossings* (HOC) consiste em aplicar iterativamente a sequência de etapas descrita abaixo.

1. Aplicar um filtro sobre o sinal original
2. Contar a quantidade de vezes que o sinal filtrado cruza o zero
3. Escolher um filtro que seja uma ordem superior ao atual e voltar à primeira etapa

Assim, conforme explicado por [9], podemos definir as características a partir do filtro passa-altas ∇ , que é um operador diferença descrito pela equação 14.

$$\nabla F_T = F_T - F_{T-1} \quad (14)$$

Dessa forma, podemos generalizar a aplicação de um filtro de ordem k sobre o sinal discretizado conforme a equação 15.

$$\nabla^{k-1} F_t = \sum_{j=1}^k \frac{(k-1)!}{(j-1)!(k-j)!} (-1)^{j-1} F_{n-j+1} \quad (15)$$

Após a aplicação da equação anterior sobre o sinal original, podemos transformá-lo em uma série binária seguindo as equações abaixo.

$$B_t(k) = 1, \text{ se } \nabla^{k-1} F_t > 0 \quad (16a)$$

$$B_t(k) = 0, \text{ caso contrário} \quad (16b)$$

Esta transformação, por sua vez, torna simples a tarefa de calcular a quantidade de cruzamentos, pois é necessário somente aplicar a equação 17 à sequência binária.

$$D_k = \sum_{t=2}^N |B_t(k) - B_{t-1}(k)| \quad (17)$$

O valor de D_k , por sua vez, pode ser armazenado dentro de um vetor de *features*, que armazenará todos os valores de cruzamentos para cada um dos filtros de diferentes ordens utilizados. Portanto, cada elemento deste vetor corresponde a uma característica diferente do sinal de EEG.

Para ilustrar a aplicação do HOC, suponha um sinal qualquer dado por uma série temporal finita como mostrado na Figura 11. A este sinal, aplicamos o filtro dado pela equação (15) com k=10 e o resultado é mostrado na Figura 12. Como utilizamos aqui um filtro passa alta, podemos notar pela Figura 12 que o sinal resultante apresenta as maiores frequências presentes no sinal original, e portanto temos uma grande quantidade de cruzamentos com o zero.

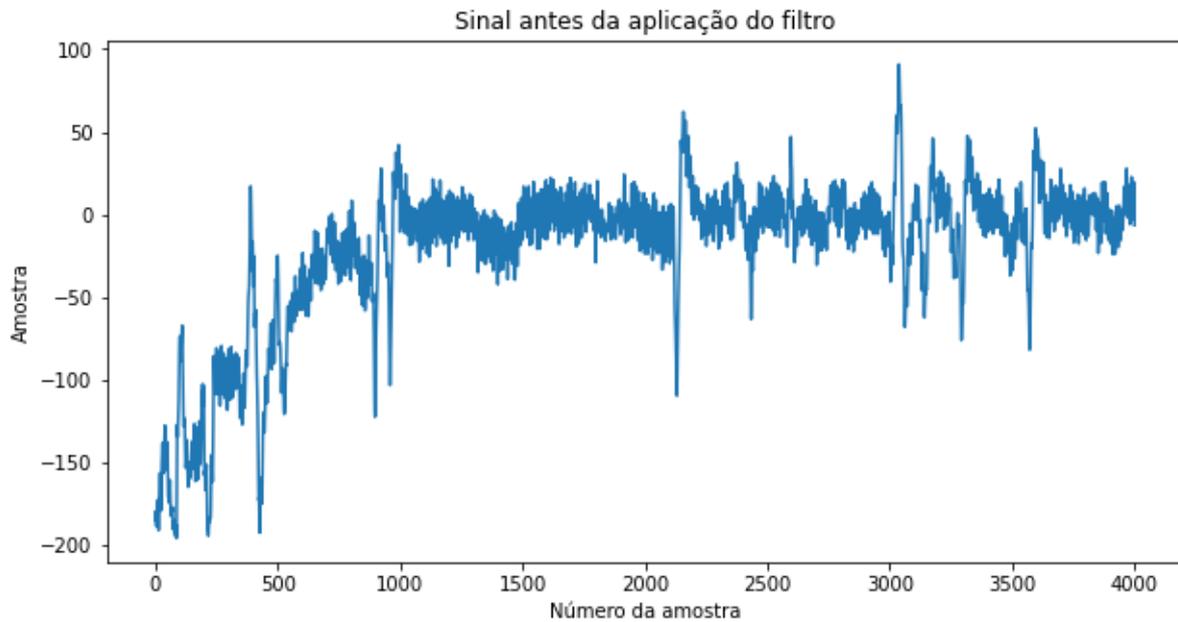


Figura 11: Sinal antes da aplicação do filtro

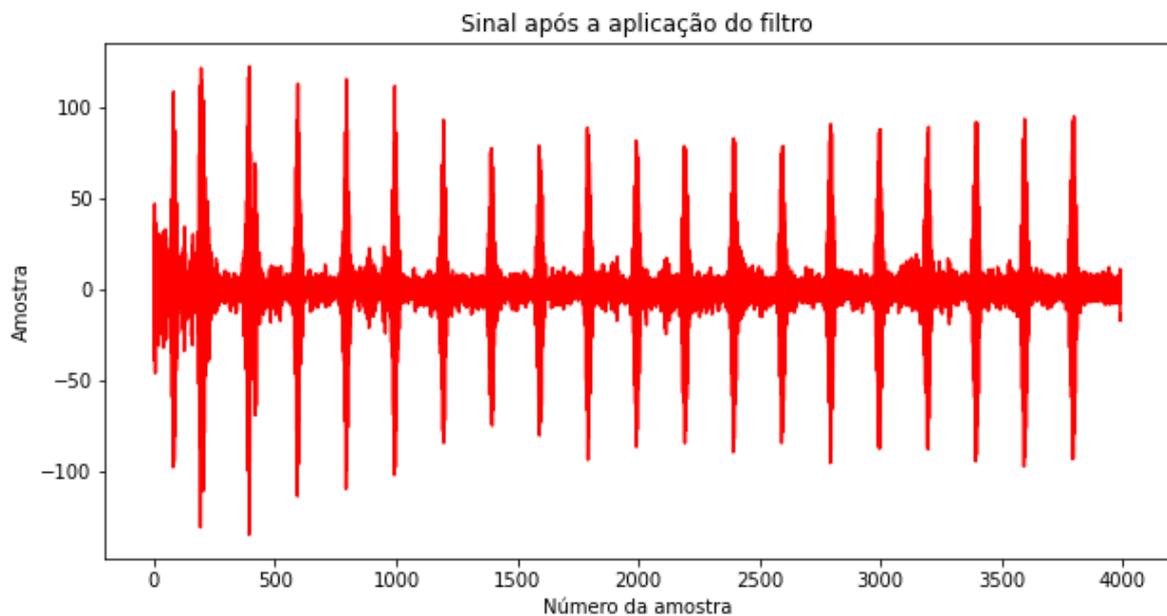


Figura 12: Sinal após a aplicação do filtro

Durante a aplicação deste projeto, foi criado um código na linguagem de programação Python que aplica as equações descritas nesta seção para cada componente independente de cada um dos sinais selecionados. Dessa forma, para cada sinal é calculado um total de $19 \cdot k$ características, onde $k \in \{8,9,10,11,12\}$, sendo que foi realizada uma simulação diferente para cada um dos possíveis valores de k .

5.3. Minimum redundancy maximum relevance (mRMR)

Durante o desenvolvimento de um modelo de classificação, é importante selecionar as características que sejam mais vantajosas para que a maior acurácia possível seja atingida. No entanto, é uma tarefa árdua para o desenvolvedor deste modelo encontrá-las, já que muitas delas, embora apresentem um alto grau de correlação com a classificação, podem ser informações redundantes, pois outras características também podem transmitir a mesma informação, conforme é ilustrado pela Figura abaixo. [12]

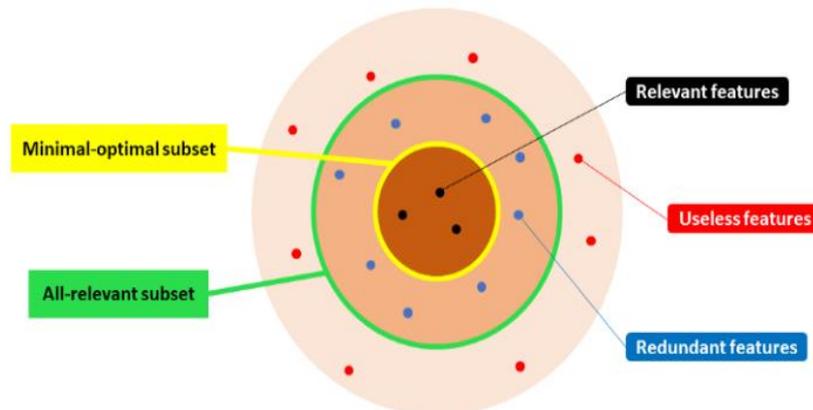


Figura 13: Representação dos tipos de subconjuntos de características [12]

A Figura 13 esquematiza, em um diagrama de Venn, como as características podem ser agrupadas: o conjunto universo contempla todas as características obtidas para o modelo, onde algumas são irrelevantes, sendo apresentadas apenas dentro do conjunto mais externo. Dentro do conjunto das características relevantes, o conjunto mais interno representa um subconjunto ótimo com características que possuem informações de maior relevância para a classificação e que não estão relacionadas entre si, isto é, todas transmitem informações diferentes e extremamente relevantes para a classificação. As características relevantes que não estão neste subconjunto ótimo, por sua vez, podem ser descartadas, pois as informações que são transmitidas por elas já estão dentro do subconjunto ótimo, o que as caracteriza como redundantes.

Para encontrar este subconjunto ótimo de características e resolver este problema, foi criado o algoritmo mRMR (*Minimum Redundancy Maximum Relevance*), que é um método de seleção de características que escolhe aquelas que possuem a maior correlação com a classificação e, simultaneamente, a menor correlação com as demais. Esta escolha, por sua vez, é feita de maneira gulosa, onde as características são escolhidas uma a uma para maximizar uma função objetivo de relevância e redundância. A relevância pode ser obtida a partir da maximização da informação mútua entre as *features* e uma classe definida previamente, enquanto a redundância pode ser determinada pela minimização da informação mútua entre as *features* já escolhidas. Assim, os dois tipos mais populares para a função de ativação são MID (*Mutual Information Difference criterion*), que corresponde ao resultado da subtração entre estas grandezas ou MIQ (*Mutual Information Quotient criterion*), que representa o quociente entre elas. [11]

Segundo [18], as funções de ativação MID e MIQ estão representadas, respectivamente, pelas equações 18 e 19.

$$\max_{x_i \in \mathcal{D}} \left[I(x_i, c) - \frac{1}{|S|} \sum_{x_j \in S} I(x_i, x_j) \right] \quad (18)$$

$$\max_{x_i \in \mathcal{D}} \left[\frac{I(x_i, c)}{\frac{1}{|S|} \sum_{x_j \in S} I(x_i, x_j)} \right] \quad (19)$$

onde \mathcal{D} é o conjunto de todas as *features* disponíveis, c é o conjunto de todas as classes e S é o subconjunto de *features* selecionadas.

O mRMR será utilizado neste projeto de três maneiras diferentes: realizando a seleção de características apenas dentro do universo dos parâmetros de Hjorth de cada sinal, depois para universo dos parâmetros HOC e finalmente, para o universo das características que contempla aquelas produzidas pelos dois métodos de extração de *features*.

5.4. Support Vector Machine (SVM)

Support Vector Machine linear é uma técnica de aprendizado de máquina supervisionado, onde é traçado um hiperplano que representa uma fronteira linear para um conjunto de dados, separando-os em classes, cuja equação é mostrada na equação 20, que define um classificador do tipo SVM linear com margens rígidas.

$$f(l) = v \cdot l + b \quad (20)$$

onde v é o vetor ortogonal ao hiperplano, l é uma variável qualquer, b uma constante e $v \cdot l$ o produto escalar entre os vetores.

Assim, a partir da equação 20, podemos realizar uma classificação binária das variáveis dividindo o plano em duas regiões distintas, como $f(l) > 0$ e $f(l) < 0$. Dessa forma, é possível definir uma função sinal $g(l)$, que é responsável por atribuir uma classificação para uma variável a partir da região em que o valor resultante da aplicação da equação 18 sobre esta está localizado, conforme exemplificado nas equações 21a e 21b e ilustrado pela Figura 14.

$$g(l) = 1, \text{ se } f(l) > 0 \quad (21a)$$

$$g(l) = -1, \text{ se } f(l) < 0 \quad (21b)$$

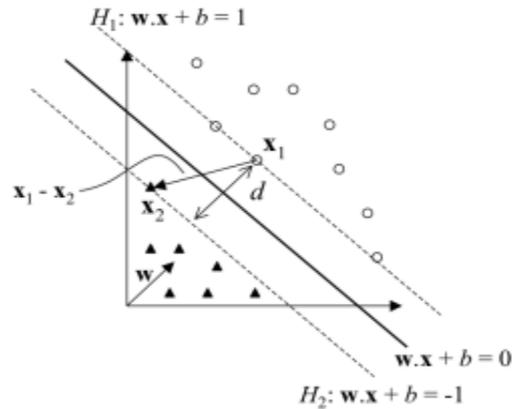


Figura 14: Exemplo de classificação para um modelo com duas variáveis [10]

A equação que define o hiperplano, por sua vez, é obtida a partir dos vetores de suporte, que são os pontos no plano que estão mais próximos do hiperplano em cada região, influenciando assim a sua orientação. Estes pontos, por sua vez, caracterizam uma outra grandeza relativa a esta técnica: a margem, que é definida como a distância entre os vetores de suporte de duas classes distintas. Assim, os algoritmos de SVM mais eficientes são aqueles que conseguem maximizar a margem. [13]

Na Figura 14, os vetores de suporte são os pontos x_1 e x_2 e a margem é a região delimitada entre os hiperplanos H_1 e H_2 .

Para este projeto, foi desenvolvido um algoritmo na linguagem de programação Python que utiliza a técnica SVM, presente dentro da biblioteca Scikit-Learn. Neste algoritmo, é realizada a classificação binária dos sinais de EEG, separando-os em representantes de emoções positivas ou negativas, utilizando os características definidas nas seções anteriores, isto é, os parâmetros de Hjorth e *Higher Order Crossings* em conjunto com a técnica mRMR para a seleção das melhores características para treinar o modelo.

6. Resultados e discussão

Nesta seção iremos exibir e apresentar uma discussão sobre os resultados obtidos durante o projeto. Inicialmente, vamos mostrar os resultados da simulação de um problema de separação cega de fontes e a recuperação dos sinais originais a partir da aplicação dos algoritmos de ICA. Na sequência, iremos exibir alguns conjuntos de sinais antes e depois da remoção de artefatos, discutindo a importância desta etapa. Finalmente, iremos discorrer sobre o desempenho do classificador para os diferentes cenários propostos.

6.1. Descrição da base

Para a realização deste projeto, foi utilizado um conjunto de sinais de EEG, que foram coletados na Universidade Federal do Vale do São Francisco (UNIVASF), aprovado pelo comitê de ética pelo código 79457917.0.0000.5196. Este estudo consistiu em selecionar vinte e três participantes, sendo nove homens e catorze mulheres com uma idade média de $24,43 \pm 4,38$ anos e para os quais foram exibidos um total de seis vídeos diferentes com duração aproximada de vinte segundos cada, sendo que o primeiro, terceiro e sexto procuravam causar emoções positivas nos voluntários e os demais, sensações negativas. A coleta dos dados, por sua vez, foi realizada utilizando o equipamento Brain Wave II - Neurovirtual®, com 20 eletrodos posicionados de acordo com o sistema 10 - 20, conforme ilustrado na Figura 15. Neste caso, foi utilizada uma frequência de amostragem de 200 Hz [1], que aplicada durante os vinte segundos de duração de cada vídeo resulta em um sinal com 4000 amostras. No entanto, um dos canais foi desconsiderado, pois apresentava mal funcionamento, restando 19 canais para análise.

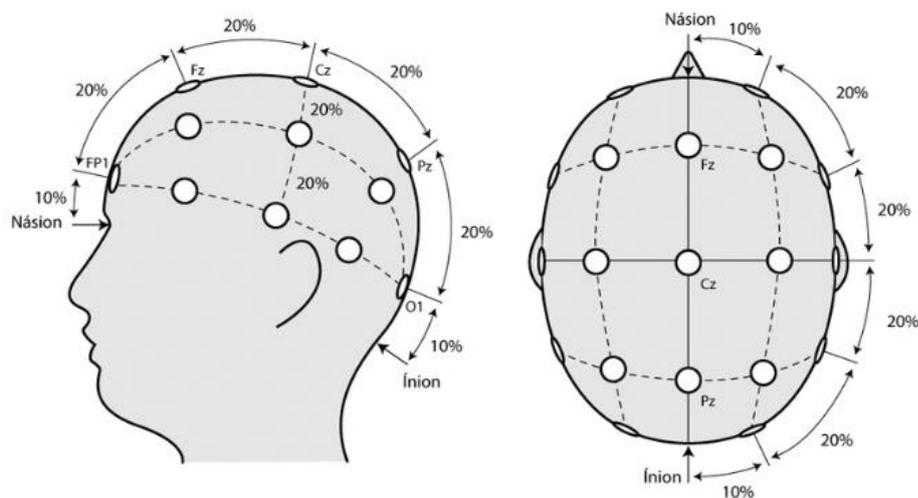


Figura 15: Posição dos eletrodos seguindo o sistema 10 - 20

Em um próximo passo, foi aplicado um questionário para identificar qual foi o impacto que cada um dos vídeos provocou nos voluntários, com o objetivo de avaliar se a emoção gerada era realmente a que se desejava, sendo que as respostas deles estão de acordo com as emoções que cada um dos vídeos procurava causar [1]. Tais respostas são mostradas graficamente na Figura 16. Pela análise dos gráficos, nota-se que o vídeo que provocou emoções positivas mais intensas foi o 1 e os vídeos que causaram emoções negativas mais profundas foram o 4 e 5.

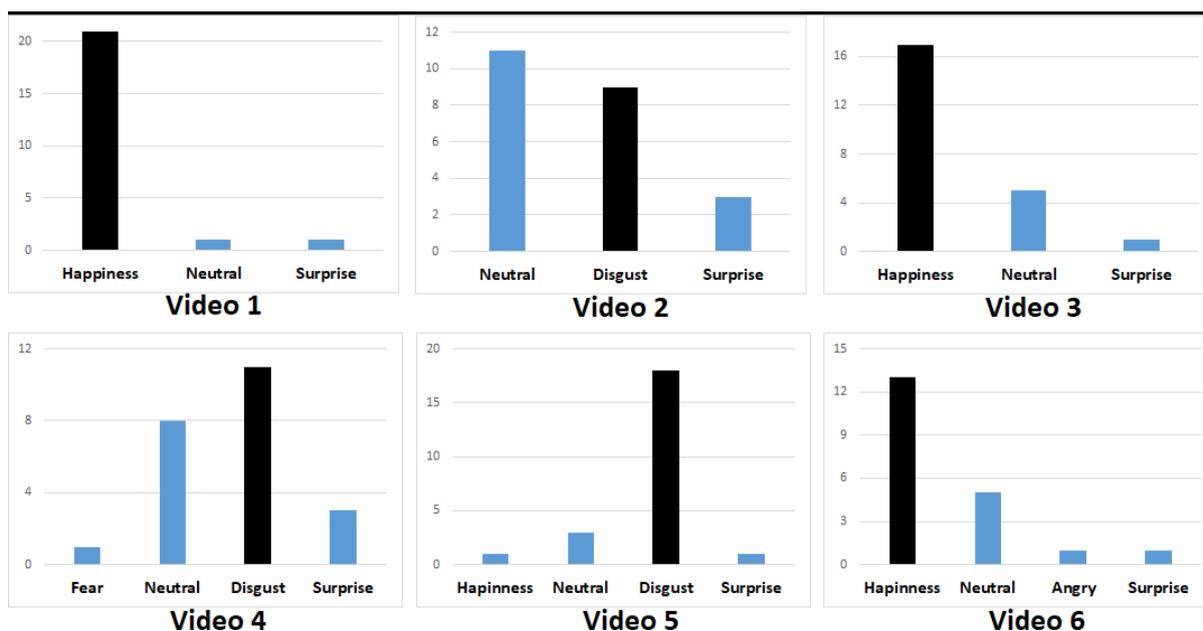


Figura 16: Impacto dos vídeos nas emoções dos voluntários [1]

6.2. Remoção de Artefatos

Para ilustrar a atuação do MARA sobre os sinais de EEG selecionados, foram escolhidos, dentro da base de dados utilizada neste projeto, dois sinais cujas componentes independentes são mostradas antes e depois da remoção de artefatos.

O primeiro conjunto de imagens corresponde ao sinal de EEG decorrente da exibição do primeiro vídeo para o primeiro voluntário.

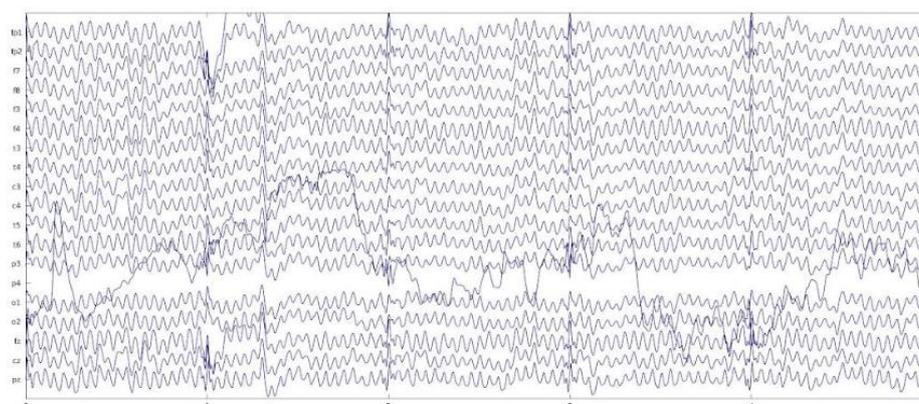


Figura 17: Componentes independentes antes da remoção de artefatos

Note que na Figura 17 há um impulso que se repete periodicamente em quase todos os canais, que pode corresponder a um artefato, como uma piscada ou algum defeito no equipamento, o que resulta em um sinal cuja presença prejudica a análise, já que não

corresponde à atividade cerebral decorrente da exibição do vídeo. A Figura 18, por sua vez, mostra este mesmo conjunto de componentes após a utilização do MARA, resultando em um conjunto de sinais sem a presença deste artefato, que traduz toda a atividade produzida durante o decorrer do experimento.

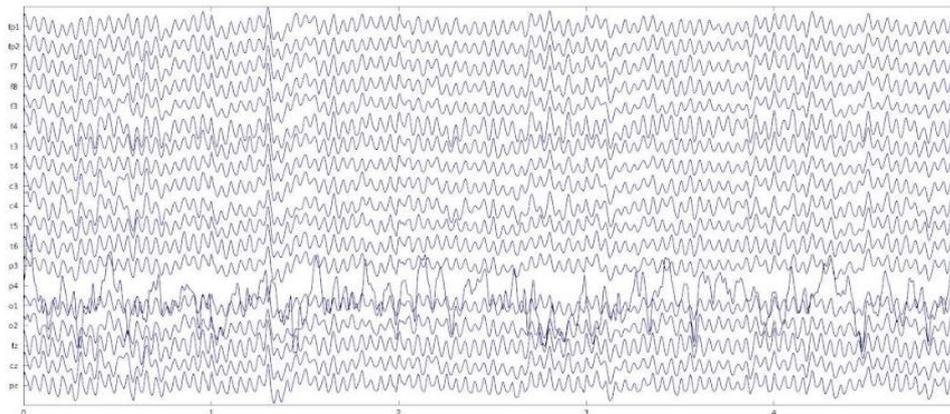


Figura 18: Componentes independentes após a remoção de artefatos

O segundo conjunto de Figuras corresponde as componentes independentes do sinal de EEG decorrente da exibição do primeiro vídeo para o quarto voluntário.

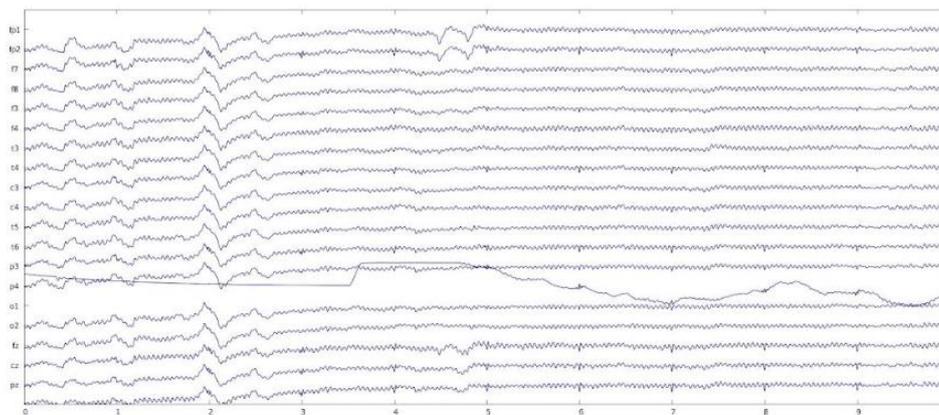


Figura 19: Componentes independentes antes da remoção de artefatos

A Figura 19 mostra um conjunto de sinais que oscilam muito pouco, o que não transmite muita informação para quem o está analisando, já que o cálculo das características dos sinais será muito afetado. Um exemplo deste problema seria o cálculo dos parâmetros de Hjorth: como estas características são calculadas a partir das variâncias dos sinais e de suas derivadas, a presença de oscilações de amplitude baixa fará com que estes parâmetros estatísticos sejam muito inferiores àqueles relativos aos sinais desejados.

Assim, a Figura 20 mostra este mesmo conjunto de componentes independentes após a remoção de artefatos, o que resultou em um conjunto de sinais que demonstram toda a atividade cerebral que foi produzida durante o experimento. Diferentemente do conjunto de sinais presentes na Figura 19, os sinais da Figura 20 oscilam em amplitudes superiores, o que resolve o problema descrito no parágrafo anterior.

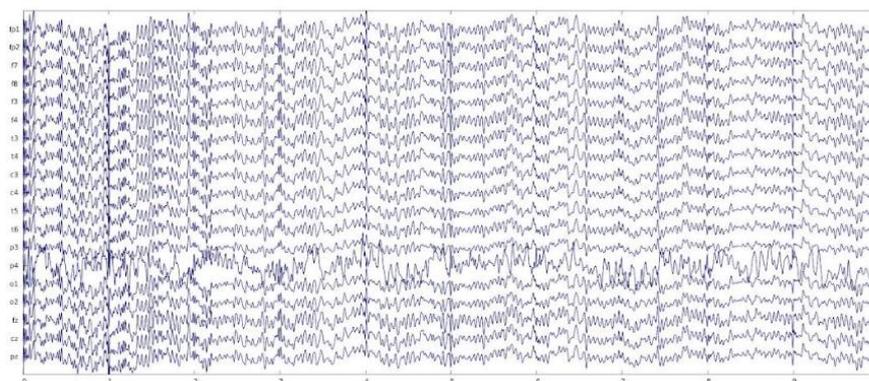


Figura 20: Componentes independentes após a remoção de artefatos

6.3. Acurácia do classificador

Conforme mencionado anteriormente, o classificador foi treinado com as características de dois conjuntos de sinais diferentes: um deles é derivado dos sinais obtidos das exibições dos vídeos 1 e 4 e o outro, dos vídeos 1 e 5, resultando em dois conjuntos de dados com 46 linhas, pois temos 23 voluntários e um sinal de cada classe para cada, além de um número de colunas que varia de acordo com o conjunto de características a ser utilizado. Este procedimento foi adotado para não desbalancear a entrada do classificador.

Foram escolhidos como objeto de estudo apenas os vídeos 1, 4 e 5 porque eles foram os que desencadearam as emoções mais intensas nos voluntários, como eles mesmos responderam no questionário cujos resultados estão presentes na Figura 2.

Foram simulados três casos diferentes, onde cada um deles engloba dez cenários distintos, sendo que para cada cenário foi realizado um total de vinte simulações. O primeiro caso corresponde à utilização de apenas os parâmetros de Hjorth como entrada, o segundo, apenas os parâmetros HOC e o terceiro, os dois em conjunto.

O primeiro cenário ocorre com a seleção do conjunto de *features* sem a utilização do algoritmo mRMR, isto é, utilizando todas as características disponíveis. Os demais cenários consistem em aplicar este algoritmo de seleção de *features*, selecionando cada vez um percentual menor de características, ou seja, na segunda iteração seria selecionado o subconjunto ótimo cuja cardinalidade corresponde a 90% do total de características, na terceira 80% do total e assim por diante.

A avaliação da acurácia do classificador seguiu um método iterativo, sendo que a cada simulação os dados eram escolhidos aleatoriamente e divididos em treino e teste numa proporção de 70% a 30%, respectivamente. Como o número de linhas é pequeno, um acerto a mais ou a menos é capaz de provocar uma variação considerável no percentual de acerto, o que faz com que seja interessante utilizar as acurácias médias obtidas em cada cenário, isto é, o valor médio do percentual de acerto do classificador a cada simulação dentro de cada cenário, para comparar os resultados entre diferentes casos e cenários.

Como neste contexto a acurácia do classificador é uma variável aleatória e o experimento de checagem de sua acurácia foi realizado várias vezes, podemos avaliar o rendimento do processo de classificação observando como cada um dos resultados de cada simulação variou em relação à média, isto é, sua variância.

- Parâmetros de Hjorth

Para a análise dos resultados decorrentes da acurácia do classificador com a utilização apenas dos parâmetros de Hjorth, foi gerada a tabela 1 (presente nos anexos), e seus dados são exibidos no gráfico de barras presente na Figura 21.

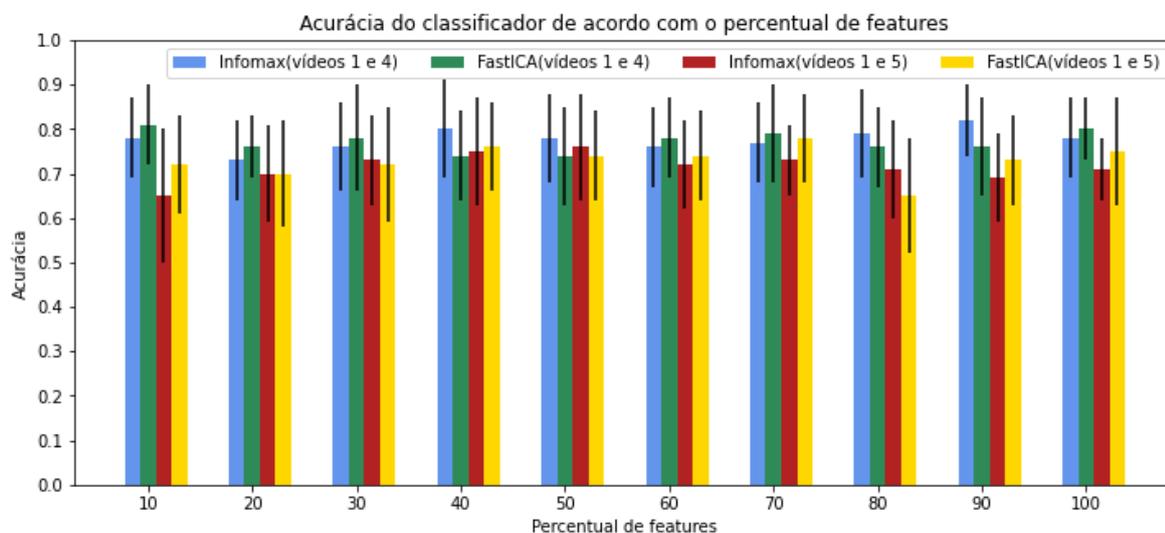


Figura 21: Análise da acurácia do classificador utilizando apenas os parâmetros de Hjorth

Na Figura 21, cada barra é uma combinação entre os dados de um conjunto de sinais e um algoritmo de ICA, o eixo horizontal representa o percentual de *features* que foi selecionado com o auxílio do algoritmo mRMR e o vertical, a acurácia média que foi atingida em cada cenário. Além disso, cada uma das linhas pretas indicam o desvio padrão da acurácia obtida.

Pela análise do gráfico, percebe-se que, em cada cenário, as maiores acurácias e os menores desvios padrões tendem a aparecer quando são utilizados como entrada os dados decorrentes da extração de características dos sinais coletados durante as exibições dos vídeos 1 e 4 do que quando são analisados os dados provenientes das exibições dos vídeos 1 e 5, independentemente do algoritmo de separação de fontes que foi utilizado, sendo que estes resultados também costumam apresentar as menores variâncias, o que mostra que o vídeo 4 foi o que provocou a emoção desejada com maior intensidade.

Uma maneira de demonstrar este fenômeno é comparando os maiores resultados para cada combinação entre um conjunto de dados e o algoritmo de separação de fontes: a acurácia da classificação dos dados decorrentes das exibições dos vídeos 1 e 4 atingiu um percentual $0,82 \pm 0,08$ (90% do total de *features*) com a utilização do Infomax e $0,81 \pm 0,09$ (10% do total de *features*) com a utilização do FastICA, enquanto que para os vídeos 1 e 5 atingiu $0,75 \pm 0,12$ (40% do total de *features*) com a utilização do Infomax e $0,78 \pm 0,1$ (70% do total de *features*) com a utilização do FastICA, sendo que a quantidade total de *features* utilizada neste caso foi de 57.

Além disso, é possível perceber pela análise do gráfico que em todos os casos a acurácia não costuma apresentar variações bruscas conforme o número de *features* aumenta. Também podemos ressaltar, a partir da análise do gráfico, que nenhum dos algoritmos de separação foi um fator determinante para o aumento da acurácia para um

mesmo conjunto de dados, pois em alguns conjuntos a utilização do FastICA produz resultados melhores e em outros, o Infomax é melhor.

- *Higher Order Crossings*

Para a análise dos resultados decorrentes da utilização de apenas o conjunto de características Higher Order Crossings, foram geradas as tabelas 2 a 5 (em anexo), que mostram as acurácias do classificador para diferentes valores de k. Os valores presentes nestas tabelas são mostrados nos gráficos abaixo, sendo que as Figuras 22 e 23 mostram quando foi utilizado o algoritmo FastICA para a separação e as Figuras 24 e 25, quando é utilizado o algoritmo Infomax.

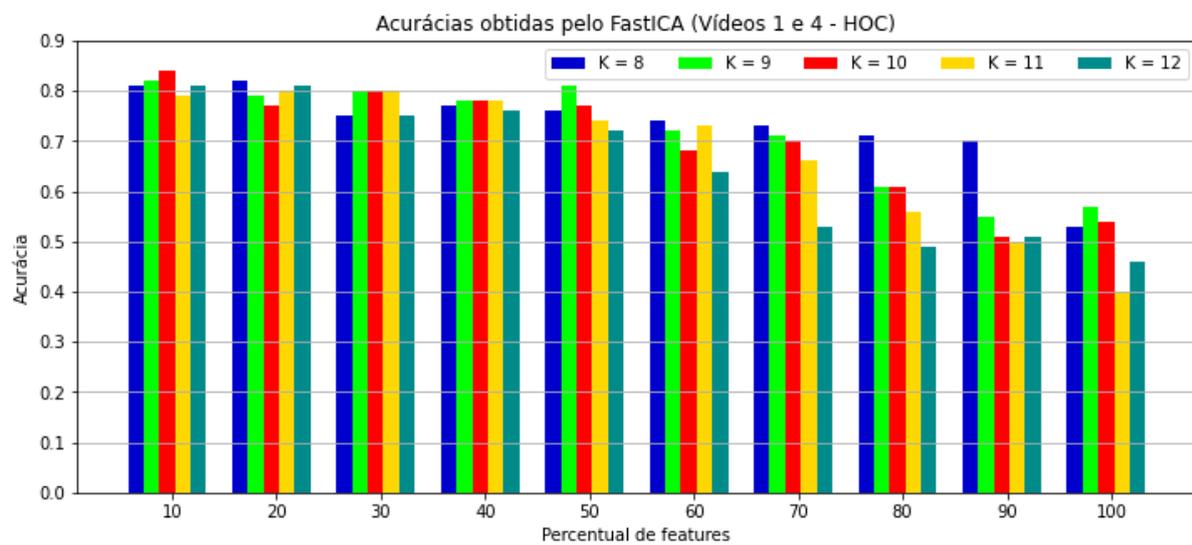


Figura 22: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 4 (FastICA - HOC)

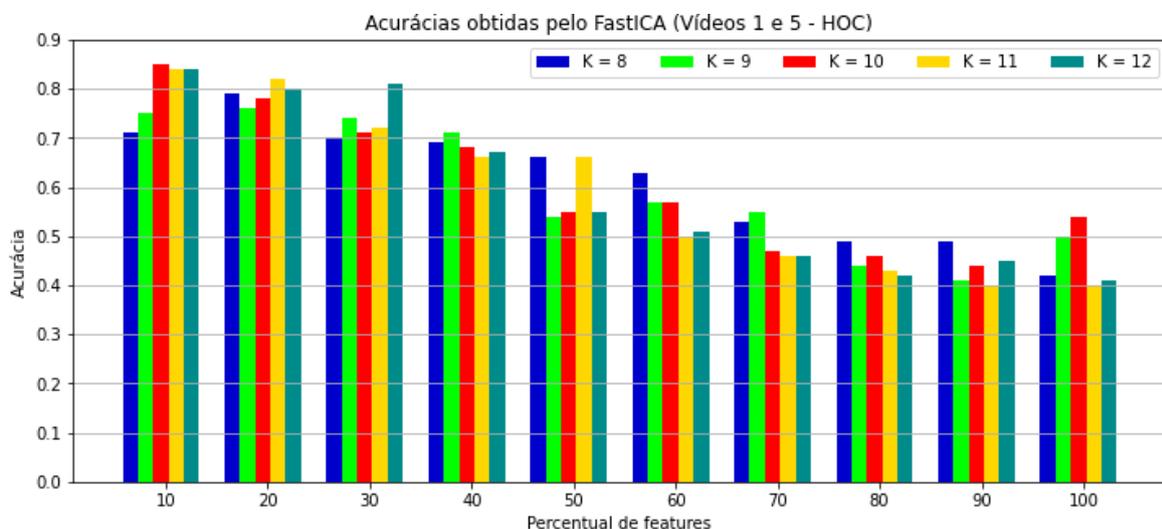


Figura 23: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 5 (FastICA - HOC)

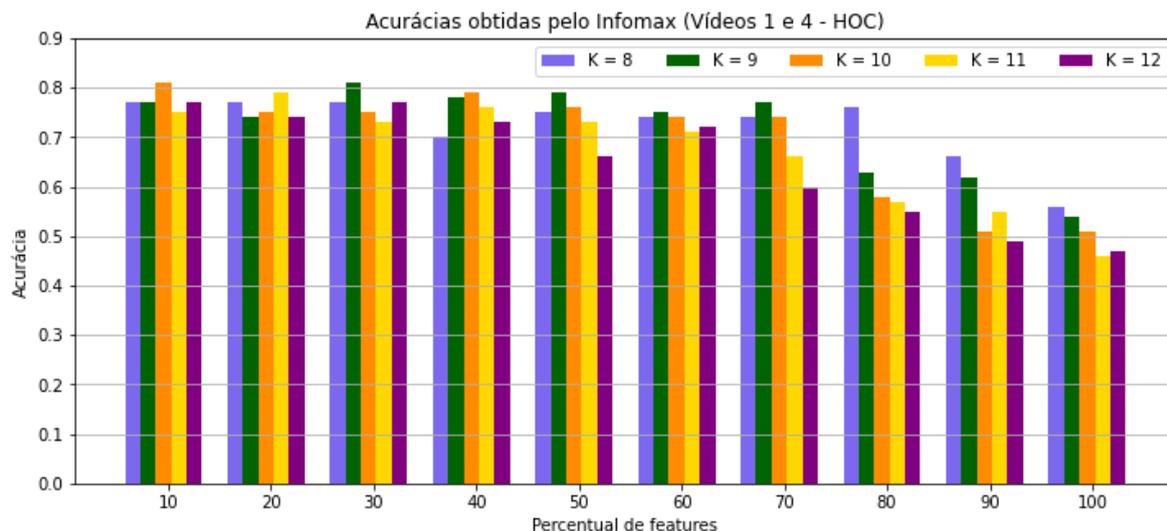


Figura 24: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 4 (Infomax - HOC)

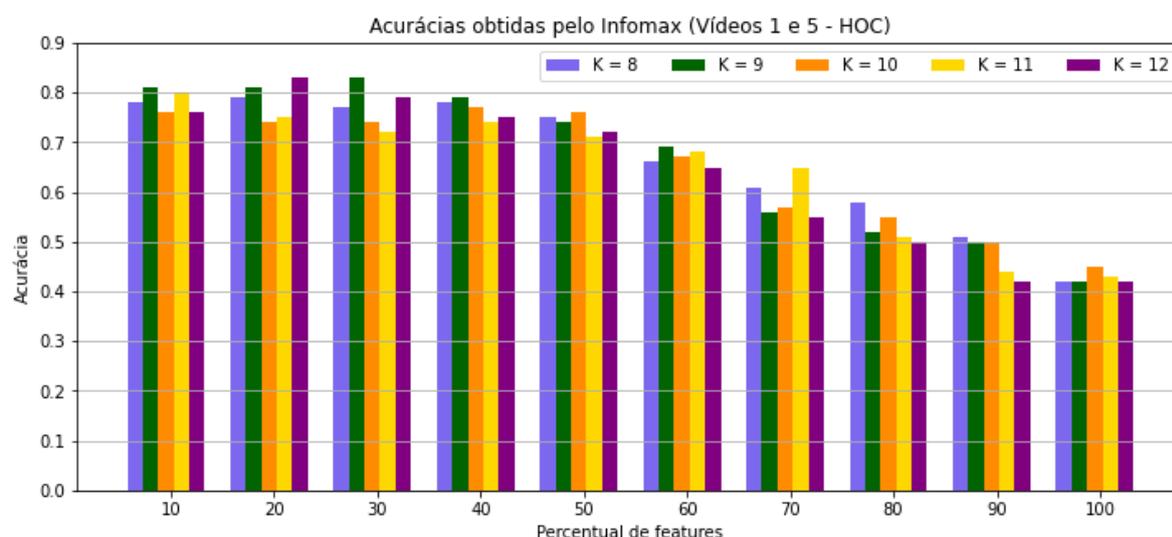


Figura 25: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 5 (Infomax - HOC)

Analisando os gráficos das Figuras 22 até 25, é possível perceber que sempre há uma tendência de queda da acurácia muito acentuada conforme a porcentagem de *features* selecionadas aumenta, pois, diferentemente do caso onde apenas os parâmetros de Hjorth foram utilizados, o conjunto de *features* é grande demais para que um bom desempenho do classificador seja obtido, pois há a utilização em excesso de características tanto redundantes quanto irrelevantes.

Além disso, a análise dos resultados permite constatar que, assim como para o item anterior, existe uma ligeira diferença de acurácia a favor dos sinais resultantes das exibições dos vídeos 1 e 4 em relação àqueles resultantes das exibições dos vídeos 1 e 5, independentemente se o algoritmo utilizado foi o FastICA ou Infomax. Isto pode ser confirmado ao comparar a Figura 22 com a 23 e a Figura 24 com a 25, onde a queda da

acurácia ocorre de forma mais suave nas Figuras 22 e 24, o que é um indício que os resultados decorrentes da exibição dos vídeos 1 e 4 tendem a ser superiores àqueles resultantes das exibições dos vídeos 1 e 5 para um mesmo cenário.

Assim, é possível obter um subconjunto ótimo que possui um tamanho que permite uma classificação mais robusta quando selecionado um percentual pequeno do total, sendo que a classificação do conjunto de dados decorrentes da exibição dos vídeos 1 e 4 atingiram uma acurácia média máxima de $0,84 \pm 0,08$ (FastICA, $k=10$ e um cenário onde o subconjunto ótimo equivale a 10% do total) e $0,81 \pm 0,07$ (Infomax, $k=10$ e um cenário onde o subconjunto ótimo equivale a 10% do total), enquanto que os dados decorrentes da exibição dos vídeos 1 e 5 atingiram uma acurácia média máxima de $0,85 \pm 0,11$ (FastICA, $k=10$ e um cenário onde o subconjunto ótimo equivale a 10% do total) e $0,83 \pm 0,09$ (Infomax, $k=9$ e um cenário onde o subconjunto ótimo equivale a 30% do total). Note que em ambos os casos os resultados quando utilizado o FastICA foram melhores do que quando foi empregado o Infomax e que estes valores são superiores àqueles obtidos na seção anterior, o que comprova que o aumento da dimensionalidade foi positivo para a classificação.

A análise dos resultados também mostra que, se forem mantidos os dados e algoritmo de separação de fontes, variando apenas o valor de k , os melhores resultados tendem a aparecer quando o seu valor é 10, mostrando que a escolha deste parâmetro também é relevante para aumentar a acurácia do algoritmo. Este fenômeno pode ser observado na Figura 26, que mostra a maior acurácia obtida quando comparados os resultados das simulações de todos os cenários, mantendo um mesmo conjunto de dados e algoritmo de separação de fontes.

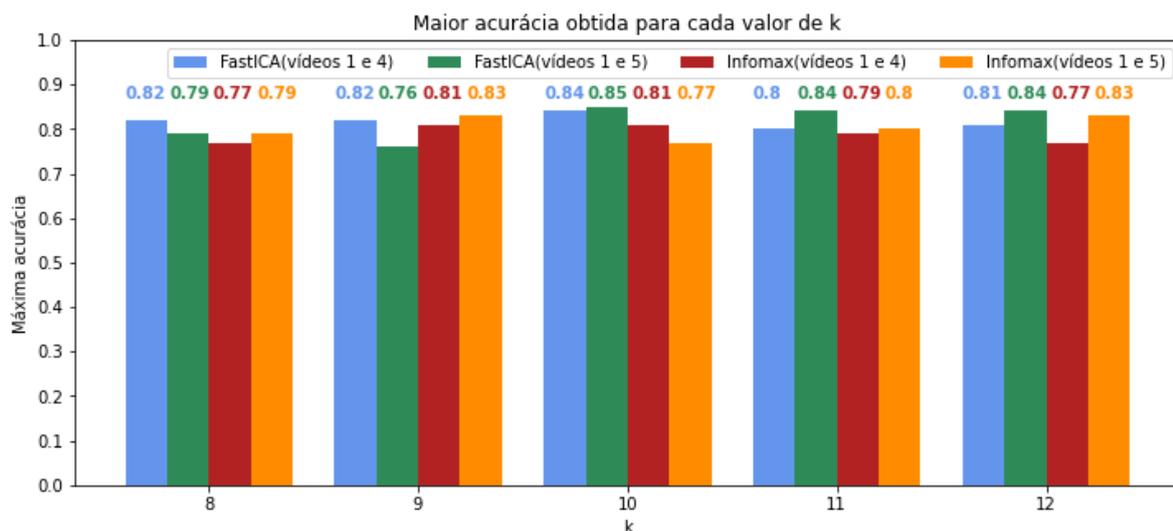


Figura 26: Máxima acurácia obtida para cada valor de k

Assim como no item anterior, podemos analisar as variâncias das acurácias calculadas para cada cenário, conforme mostrado na Figura 27.

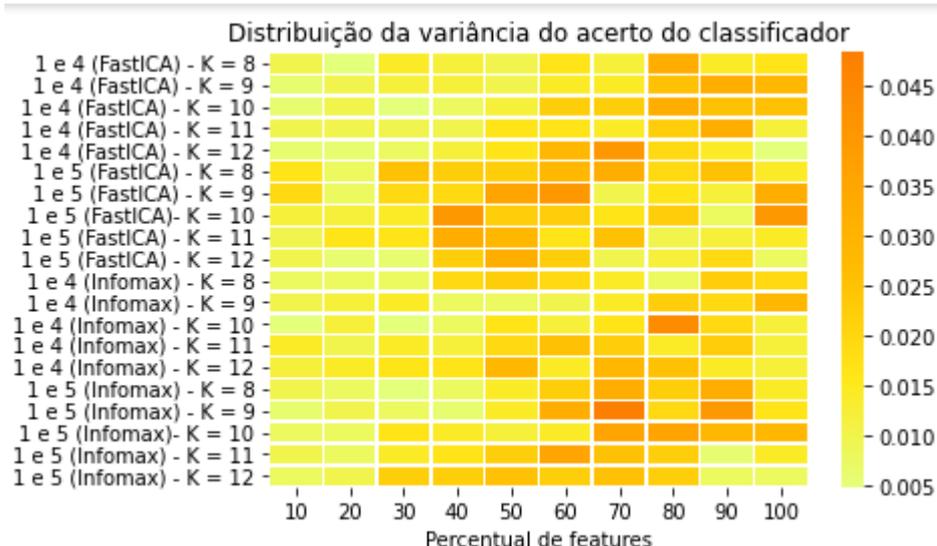


Figura 27: Análise da variância da acurácia do classificador utilizando os parâmetros HOC

Na Figura 27, o eixo horizontal corresponde ao percentual do total de *features* utilizado na simulação, o eixo vertical indica a combinação entre a divisão dos dados empregada, o algoritmo de separação de fontes utilizado e o valor de *k* em cada simulação. Finalmente, cada retângulo do gráfico corresponde à variância obtida em cada tipo de simulação.

Este gráfico permite a percepção de que há uma tendência de aumento da variância da acurácia conforme o tamanho do subconjunto ótimo de *features* aumenta. Além disso, outro fator de influência para o aumento da variância do resultado é o conjunto de dados utilizado: quando utilizados os dados dos sinais coletados após a exibição dos vídeos 1 e 5, as variâncias costumam ser mais elevadas que as relativas aos dados dos sinais colhidos após a exibição dos vídeos 1 e 4 para um mesmo algoritmo de separação de fontes e valor de *k*, independentemente do algoritmo de separação de fontes empregado ou percentual de *features* selecionado.

Podemos também ressaltar que neste caso as variâncias aumentaram se comparadas com o item anterior, enquanto a utilização de apenas os parâmetros de Hjorth resultou em um conjunto que atingia no máximo 0,022, a utilização de apenas o HOC resultou em um conjunto de resultados cujas variâncias podem atingir até 0,045.

- Hjorth e *Higher Order Crossings*

A análise da acurácia do classificador para a união dos dois conjuntos de *features* segue o mesmo padrão dos itens anteriores, sendo geradas as tabelas 6 a 9 em anexo e os gráficos apresentados nas Figuras 28 a 31.

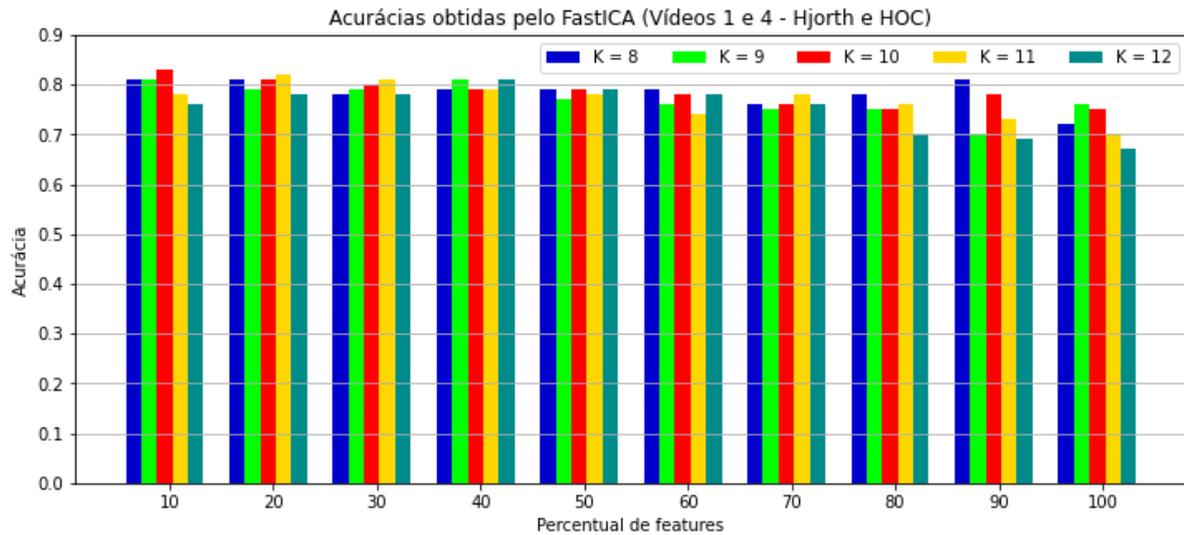


Figura 28: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 4 (FastICA - Hjorth e HOC)

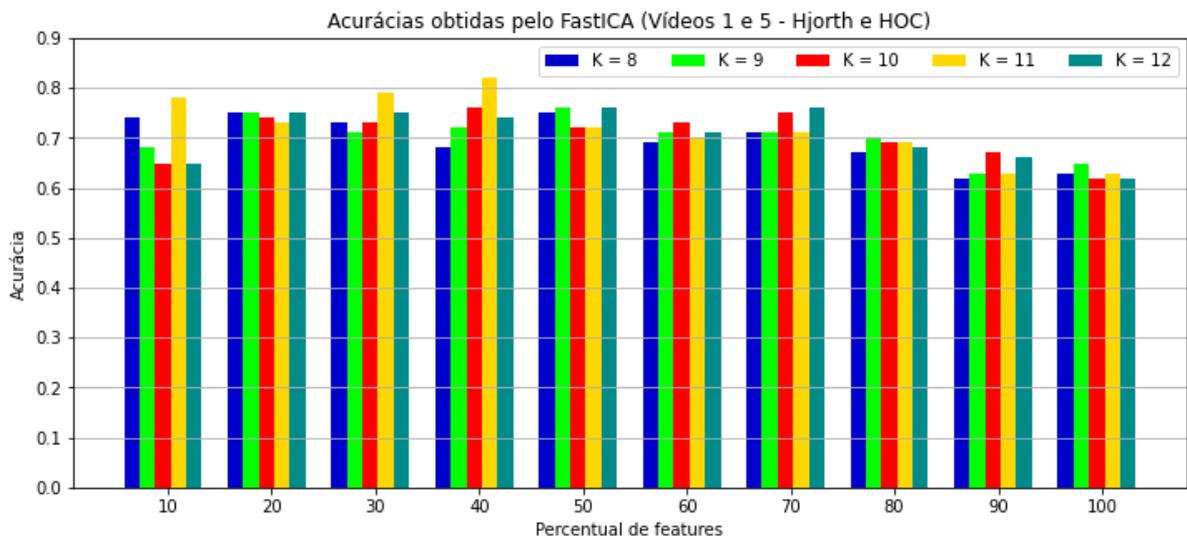


Figura 29: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 5 (FastICA - Hjorth e HOC)

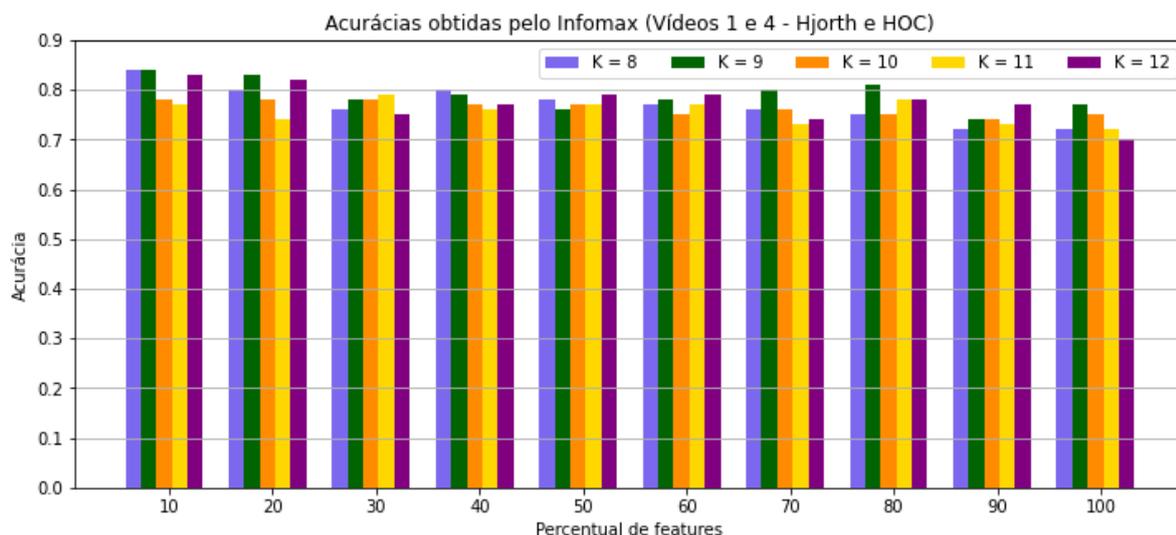


Figura 30: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 4 (Infomax - Hjorth e HOC)

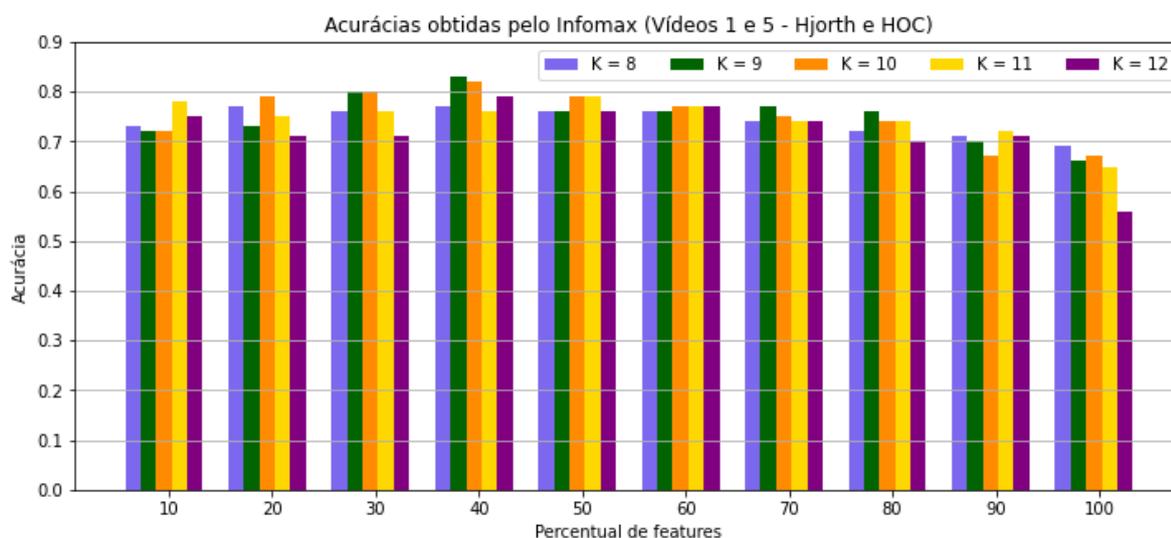


Figura 31: Análise da acurácia do classificador para os sinais decorrentes das exibições dos vídeos 1 e 5 (Infomax - Hjorth e HOC)

Analisando os resultados obtidos, é possível perceber que a união dos dois conjuntos de *features* fez com que cada um deles exercesse alguma influência sobre o comportamento da acurácia, resultando em uma tendência de queda da mesma conforme o número de *features* aumenta, o que torna necessária uma redução de dimensionalidade. No entanto, é possível ressaltar que há uma influência positiva dos parâmetros de Hjorth, já que neste caso o índice mínimo de acerto é de 56%, enquanto quando é utilizado apenas o HOC, este valor mínimo atinge 40%.

A classificação do conjunto de dados decorrentes da exibição dos vídeos 1 e 4 atingiram uma acurácia média máxima de $0,83 \pm 0,09$ (FastICA, $k = 10$ e um cenário onde o subconjunto ótimo equivale a 10% do total) e $0,84 \pm 0,1$ (Infomax, $k = 8$ e um cenário onde o subconjunto ótimo equivale a 10% do total), enquanto os dados decorrentes da exibição dos vídeos 1 e 5 atingiram uma acurácia média máxima de $0,78 \pm 0,09$ (FastICA, $k = 10$ e um

cenário onde o subconjunto ótimo equivale a 10% do total) e $0,8 \pm 0,09$ (Infomax, $k=9$ e um cenário onde o subconjunto ótimo equivale a 30% do total).

Além disso, foi mantida a acurácia maior quando analisados os dados referentes à exibição dos vídeos 1 e 4 em relação aos referentes à exibição dos vídeos 1 e 5, independentemente se o algoritmo para a separação de fontes utilizado foi o FastICA ou Infomax, o que já era esperado, devido aos resultados obtidos nas seções anteriores.

Da mesma maneira que foi feito na seção anterior, isto pode ser confirmado a partir da análise dos gráficos: note que na Figura 26 as acurácias costumam permanecer no intervalo $[0,8 ; 0,7]$ durante quase todo gráfico, com exceção de alguns poucos *outliers*, enquanto que na Figura 27, as acurácias tendem a decair no intervalo $[0,75 ; 0,6]$, o que indica que os resultados do primeiro conjunto de dados costumam ser melhores que os do segundo.

Podemos também comparar como a acurácia variou conforme os diferentes valores de k , o que é abordado na Figura 32, que exhibe os índices de acerto da mesma maneira que o gráfico presente na Figura 26.

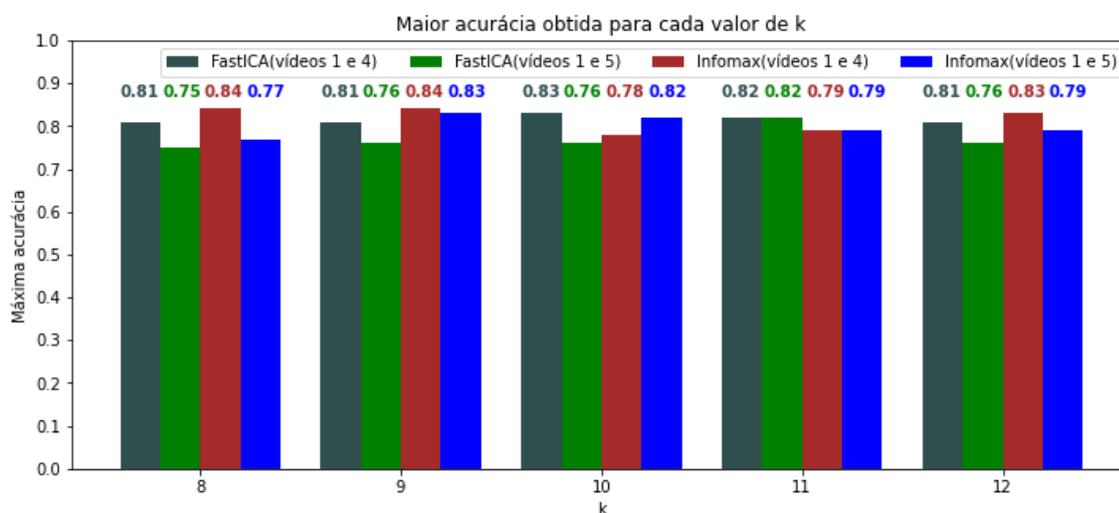


Figura 32: Máxima acurácia obtida para cada valor de k

Note que, diferentemente do caso onde apenas o HOC é utilizado, os maiores índices de acerto para um mesmo conjunto de dados e algoritmo de separação de fontes não necessariamente estão presentes no caso onde o valor de k é 10, sendo que isto ocorre apenas para o conjunto de dados decorrente da exibição dos vídeos 1 e 4 com a utilização do FastICA, onde a acurácia média obtida é de 0,83. Quando foi utilizado o algoritmo Infomax, o classificador alcançou a acurácia máxima para os dois conjuntos de dados quando o valor de k é 9, isto é, 0,84 e 0,83 para os dados dos vídeos 1 e 4 e 1 e 5, respectivamente. Finalmente, a classificação do conjunto de dados restante atingiu sua acurácia máxima quando o valor de k é 11, isto é, 82%.

Assim como nos dois itens anteriores, é interessante analisar como a variância das acurácias se comportou em diferentes cenários, como mostra a Figura 33.

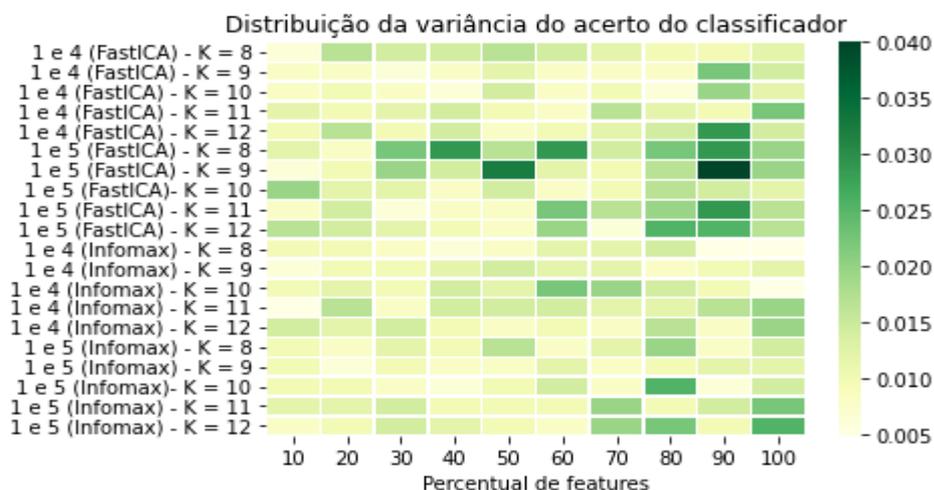


Figura 33: Análise da variância da acurácia do classificador utilizando os parâmetros de Hjorth e HOC

Pela análise do mapa, nota-se que, embora os maiores valores das variância tenham se mantido concentrados nos casos onde foram utilizadas as maiores quantidade de *features* e tenham um limitante superior similar ao HOC, o mapa apresenta uma coloração muito mais homogênea que aquele representado na Figura 27, o que indica que os valores das variâncias para diferentes cenários ficaram mais próximos entre si, o que é outra evidência da influência positiva dos parâmetros de Hjorth.

Somado a isto, é possível ressaltar que manteve-se o padrão apresentado pelos itens anteriores onde as maiores variâncias estão presentes nos dados dos sinais decorrentes das exhibições dos vídeos 1 e 5, independentemente do algoritmo de separação utilizado.

- Discussão dos resultados

A partir das análises em cada um dos itens, percebe-se que a origem dos dados foi o fator de maior influência para o rendimento do algoritmo quando utilizados os parâmetros de Hjorth, HOC ou ambos em conjunto, já que o classificador obteve a maior acurácia quando recebeu como entrada os dados decorrentes das exhibições dos vídeos 1 e 4, pois foram estes que obtiveram as acurácias mais elevadas e as menores variâncias quando comparados aos dados obtidos a partir da exibição dos vídeos 1 e 5 na maioria das situações, com exceção de alguns *outliers*. Este resultado é inesperado, se for analisado em conjunto com o gráfico da Figura 16, pois ele mostra que o vídeo 5 ocasionou reações mais intensas que o 4, tendo causado uma emoção negativa numa quantidade maior de voluntários.

Podemos ressaltar que quando utilizamos unicamente os parâmetros de Hjorth não há uma tendência de queda da acurácia conforme o tamanho do conjunto de *features* aumenta. Com este conjunto de características, o classificador pode atingir acurácias de até 82% e apresenta a menor variância dentre todos os casos, que pertence a intervalo [0,006 ; 0,022].

Quando utilizados unicamente os parâmetros HOC, os melhores resultados costumam aparecer quando o valor de k é 10, não havendo uma diferença significativa em relação aos demais valores. Outro fenômeno relevante é o fato da acurácia do classificador mostrar uma relação inversa com o tamanho da entrada, já que ocorre uma queda acentuada da acurácia de um valor próximo a 80% para 40% para as curvas relativas a cada algoritmo

de separação de fontes e o valor de k conforme o percentual de características utilizadas se aproxima de 100%.

Esta relação se deve ao fato de que o número de características criadas por este método é muito superior àquele produzido pelos parâmetros de Hjorth, o que possibilita uma classificação mais robusta quando selecionado um subconjunto ótimo do total de *features*. Esta técnica produziu uma acurácia máxima de 85% (conjunto de dados decorrentes da exibição dos vídeos 1 e 5, com a utilização do FastICA, $k = 10$ e um cenário onde o subconjunto ótimo equivale a 10% do total), porém os seus resultados possuem variâncias que se concentraram nas regiões onde houve a inclusão dos maiores percentuais de características e pertencem ao intervalo $[0; 0,045]$, sendo mais elevadas que as decorrentes da utilização dos parâmetros de Hjorth. Estes resultados são esperados, já que, conforme o tamanho do conjunto de características aumenta, cada vez mais características redundantes e irrelevantes são utilizadas como entrada, o que diminui a eficiência do classificador, resultando em índices de acertos cada vez mais baixos e que variam muito a cada simulação.

Quando os dois conjuntos de *features* foram utilizados em conjunto, o comportamento dos resultados mostra uma influência dos dois tipos de conjuntos de *features*: há uma queda da acurácia do classificador conforme o número de características aumenta, mas ela não é tão acentuada como para quando apenas o HOC é utilizado, sendo que ela decai de um valor em torno de 80% para outro que está em torno de 60% para todas as curvas analisadas, que é uma influência positiva da utilização dos parâmetros de Hjorth. No entanto, as maiores variâncias dos resultados, que podem atingir valores tão altos como quando apenas o HOC é utilizado, estão concentradas em regiões onde o percentual de *features* utilizadas se aproxima de 100%, enquanto na seção anterior as maiores variâncias apareciam tanto quando o percentual de *features* utilizado era de 50% como quando era 100%.

Além disso, ao contrário do HOC, não existe a tendência dos melhores resultados obtidos estarem concentrados nos casos onde $k = 10$, o que é outra evidência da influência dos parâmetros de Hjorth. A acurácia média máxima, por sua vez, foi de 84% (conjunto de dados decorrentes da exibição dos vídeos 1 e 4, com a utilização do Infomax, $k=8$ e um cenário onde o subconjunto ótimo equivale a 10% do total).

Podemos ressaltar também a influência do algoritmo de separação de fontes utilizado em cada um dos conjuntos de dados quando utilizamos apenas os parâmetros HOC ou a união dos dois tipos de técnicas de extração de *features*: na maioria dos casos, a aplicação do algoritmo Infomax resulta em acurácias melhores do que quando é utilizado o FastICA. No entanto, o segundo algoritmo está associado a resultados melhores quando o tamanho do subconjunto ótimo equivale a 10%, que é o caso onde as maiores acurácias são atingidas.

Finalmente, percebe-se que a classificação pode ser otimizada quando os parâmetros de Hjorth forem utilizados em conjunto com o HOC (atingindo uma acurácia máxima de $0,84 \pm 0,08$), pois embora a maior acurácia tenha sido atingida quando o HOC foi utilizado sozinho ($0,85 \pm 0,11$), a utilização em conjunto das duas técnicas não provoca quedas bruscas de acurácia e estas possuem variâncias mais baixas na maioria dos casos, principalmente se o subconjunto ótimo possuir um tamanho equivalente a 10% do conjunto do total de características, que é o mais indicado para uma boa classificação, além de economizar processamento por reduzir a dimensionalidade.

7. Conclusão

Durante o desenvolvimento deste projeto, conseguimos estudar os principais conceitos que envolvem um problema de separação cega de fontes, Análise de Componentes Independentes, remoção de artefatos e caracterização dos sinais de EEG a partir dos parâmetros de Hjorth e *Higher Order Crossings*.

Assim, conseguimos simular com sucesso um problema de separação cega de fontes utilizando tanto o algoritmo FastICA como Infomax em um contexto simples, com duas fontes, duas misturas e sem adição de ruído, sendo que o erro absoluto médio foi inferior para o algoritmo Infomax em relação ao FastICA.

Em seguida, conseguimos realizar a remoção de artefatos dos sinais de EEG selecionados dentro da base dados utilizada neste projeto e os caracterizamos conforme os dois tipos de *features* mencionados, que foram responsáveis por serem a entrada de um classificador *Support Vector Machine*, que foi testado para diversos tipos de entradas diferentes, sendo que o melhor cenário é quando ocorre a utilização dos dois tipos de características em conjunto e é selecionado um subconjunto ótimo pelo algoritmo mRMR de cardinalidade igual a 10% do total de características disponíveis, o que resultou em uma acurácia que pode atingir até 83% e apresenta uma variância muito próxima de zero.

Além disso, o algoritmo de separação de fontes que está relacionado com as melhores acurácias quando é selecionado um subconjunto ótimo de tamanho equivalente a 10% do total é o FastICA. Embora este fator seja importante, a principal influência para a melhora dos resultados da classificação é a escolha dos dados a serem utilizados como entrada, pois a acurácia do classificador quando utiliza como entrada as características dos sinais de EEG decorrentes da exibição dos vídeos 1 e 4 é superior à quando este utiliza os sinais colhidos durante a exibição dos vídeos 1 e 5, o que é inesperado, já que o vídeo 4 gerou muito mais reações neutras que o 5.

8. Referências Bibliográficas

- [1] GRILO JR, M. RIBEIRO, L. MORAES, C. MELO, C. FANTINATO, D. SAMPAIO, L. NEVES, A. RAMOS, R. **Artifact removal in EEG based emotional signals through linear and nonlinear methods**. The 7th IEEE International Conference on E-Health and Bioengineering - EHB 2019. Grigore T. Popa University of Medicine and Pharmacy, Iasi, Romania. 21 a 23 de Novembro de 2019.
- [2] SUYAMA, R. **Proposta de métodos de separação cega de fontes para misturas convolutivas e não-lineares**. 2007. Tese (Doutorado) - Engenharia Elétrica - Universidade Estadual de Campinas.
- [3] LANGLOIS, D. CHARTIER, S. GOSELIN, D. **An Introduction to Independent Component Analysis: InfoMax and FastICA algorithms**. Tutorials in Quantitative Methods for Psychology 2010, Vol. 6(1), p. 31-38.
- [4] HYVARINEN, A. KARHUNEN, J. OJA, E. **Independent Component Analysis**. New York: John Wiley & Sons.
- [5] JIANG, X. BIAN, G. TIAN, Z. **Removal of Artifacts from EEG Signals: A Review**. Institute of Automation, Chinese Academy of Science, Beijing 2018.
- [6] Winkler, S. Haufe, and M. Tangermann. **Automatic classification of artifactual ICA-components for artifact removal in EEG signals**. Behavioral and Brain Functions, 7:30, 2011.
- [7] Seung-Hyeon Oh, Yu-Ri Lee, and Hyoung-Nam Kim. **A Novel EEG Feature Extraction Method Using Hjorth Parameter**. International Journal of Electronics and Electrical Engineering Vol. 2, No. 2, 2014.
- [8] HJORTH, B. **Eeg analysis based on time domain properties**. Electroencephalography and Clinical Neurophysiology, 29(3):306 – 310, 1970.
- [9] Panagiotis Petrantonakis, Leontios Hadjileontiadis. **Emotion Recognition From EEG Using Higher Order Crossings**. IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society. Outubro, 2009.
- [10] Ana Carolina Lorena, André C. P. L. F. de Carvalho. **Uma Introdução às Support Vector Machine**. Centro de Matemática, Computação e Cognição, Universidade Federal do ABC, Santo André - São Paulo, 2007.
- [11] Milos Radovic, Mohamed Ghalwash, Nenad Filipovic, Zoran Obradovic. **Minimum redundancy maximum relevance feature selection approach for temporal gene**

expression data. BMC Bioinformatics, 2017.

[12] MAZZANTINI, S. **“MRMR” Explained Exactly How You Wished Someone Explained to You.** Towards Data Science, 12 de fevereiro de 2021. Disponível em: <<https://towardsdatascience.com/mrmr-explained-exactly-how-you-wished-someone-explained-to-you-9cf4ed27458b>>. Acesso em: 12 de outubro de 2022.

[13] SATHEESH, V. **Machine Learning Algorithms: Support Vector Machine.** Analytics Vidhya, 19 de março de 2021. Disponível em: <<https://medium.com/analytics-vidhya/machine-learning-algorithms-support-vector-machines-ddfc413540d2>>. Acesso em: 13 de outubro de 2022.

[14] C. GUYTON, Artur; E. HALL, John. **Tratado de Fisiologia Médica.** 9a Edição. 9º ed. [S.l: s.n.], 1997.

[15] Python Software Foundation. **Python Language Reference.** Disponível em: <www.python.org>

[16] Mathworks Inc. **MATLAB and Statistics Toolbox.** Natick, Massachusetts, United States.

[17] DELORME, A, MAKEIG, S. **EEGLAB: an open-source toolbox for analysis of single trial EEG dynamics.** Journal of Neuroscience Methods 134: 9 – 21. Disponível em: <www.eeglab.org>

[18] SERAFIM, L. GRILO JR, M. MORAES, C. FANTINATO, D. NEVES, A. RAMOS, R. **MRMR aplicado a sinais de EEG para a classificação de emoções.** XXXIX Simpósio brasileiro de telecomunicações e processamento de sinais - SBRT 2021, 26–29 de setembro de 2021. Fortaleza, CE.

9. Anexos

Conforme mencionado durante o artigo, as tabelas presentes nesta seção representam a acurácia média, que é indicada pela letra A e seu desvio padrão, que é indicado por ΔA .

Conjunto de dados	Vídeos 1 e 4		Vídeos 1 e 5	
	Infomax	FastICA	Infomax	FastICA
	$A \pm \Delta A$			
10%	$0,78 \pm 0,09$	$0,81 \pm 0,09$	$0,65 \pm 0,15$	$0,72 \pm 0,11$
20%	$0,73 \pm 0,09$	$0,76 \pm 0,07$	$0,7 \pm 0,11$	$0,7 \pm 0,12$
30%	$0,76 \pm 0,1$	$0,78 \pm 0,12$	$0,73 \pm 0,1$	$0,72 \pm 0,13$
40%	$0,8 \pm 0,11$	$0,74 \pm 0,1$	$0,75 \pm 0,12$	$0,76 \pm 0,1$
50%	$0,78 \pm 0,1$	$0,74 \pm 0,11$	$0,76 \pm 0,12$	$0,74 \pm 0,1$
60%	$0,76 \pm 0,09$	$0,78 \pm 0,09$	$0,72 \pm 0,1$	$0,74 \pm 0,1$
70%	$0,77 \pm 0,09$	$0,79 \pm 0,11$	$0,73 \pm 0,08$	$0,78 \pm 0,1$
80%	$0,79 \pm 0,1$	$0,76 \pm 0,09$	$0,71 \pm 0,11$	$0,65 \pm 0,13$
90%	$0,82 \pm 0,08$	$0,76 \pm 0,11$	$0,69 \pm 0,1$	$0,73 \pm 0,1$
100%	$0,78 \pm 0,09$	$0,8 \pm 0,07$	$0,71 \pm 0,07$	$0,75 \pm 0,12$

Tabela 1: Acurácia do classificador para diferentes cenários (Hjorth)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,81 \pm 0,1$	$0,82 \pm 0,08$	$0,84 \pm 0,08$	$0,79 \pm 0,1$	$0,81 \pm 0,08$
20%	$0,82 \pm 0,07$	$0,79 \pm 0,1$	$0,77 \pm 0,1$	$0,8 \pm 0,1$	$0,81 \pm 0,08$
30%	$0,75 \pm 0,12$	$0,8 \pm 0,11$	$0,8 \pm 0,07$	$0,8 \pm 0,1$	$0,75 \pm 0,09$
40%	$0,77 \pm 0,11$	$0,78 \pm 0,11$	$0,78 \pm 0,09$	$0,78 \pm 0,1$	$0,76 \pm 0,11$
50%	$0,76 \pm 0,1$	$0,81 \pm 0,1$	$0,77 \pm 0,11$	$0,74 \pm 0,13$	$0,72 \pm 0,13$
60%	$0,74 \pm 0,13$	$0,72 \pm 0,12$	$0,68 \pm 0,15$	$0,73 \pm 0,13$	$0,64 \pm 0,17$
70%	$0,73 \pm 0,11$	$0,71 \pm 0,12$	$0,7 \pm 0,15$	$0,66 \pm 0,12$	$0,53 \pm 0,2$
80%	$0,71 \pm 0,18$	$0,61 \pm 0,16$	$0,61 \pm 0,18$	$0,56 \pm 0,15$	$0,49 \pm 0,14$
90%	$0,7 \pm 0,12$	$0,55 \pm 0,18$	$0,51 \pm 0,16$	$0,5 \pm 0,18$	$0,51 \pm 0,12$
100%	$0,53 \pm 0,13$	$0,57 \pm 0,17$	$0,54 \pm 0,16$	$0,4 \pm 0,11$	$0,46 \pm 0,07$

Tabela 2: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 4 e FastICA - HOC)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,71 \pm 0,13$	$0,75 \pm 0,14$	$0,85 \pm 0,11$	$0,84 \pm 0,1$	$0,84 \pm 0,1$
20%	$0,79 \pm 0,09$	$0,76 \pm 0,09$	$0,78 \pm 0,11$	$0,82 \pm 0,13$	$0,8 \pm 0,08$
30%	$0,7 \pm 0,16$	$0,74 \pm 0,14$	$0,71 \pm 0,12$	$0,72 \pm 0,13$	$0,81 \pm 0,08$
40%	$0,69 \pm 0,15$	$0,71 \pm 0,14$	$0,68 \pm 0,2$	$0,66 \pm 0,18$	$0,67 \pm 0,15$
50%	$0,66 \pm 0,15$	$0,54 \pm 0,19$	$0,55 \pm 0,15$	$0,66 \pm 0,17$	$0,55 \pm 0,18$
60%	$0,63 \pm 0,17$	$0,57 \pm 0,2$	$0,57 \pm 0,15$	$0,5 \pm 0,13$	$0,51 \pm 0,15$
70%	$0,53 \pm 0,18$	$0,55 \pm 0,1$	$0,47 \pm 0,13$	$0,46 \pm 0,16$	$0,46 \pm 0,1$
80%	$0,49 \pm 0,14$	$0,44 \pm 0,13$	$0,46 \pm 0,15$	$0,43 \pm 0,1$	$0,42 \pm 0,11$
90%	$0,49 \pm 0,16$	$0,41 \pm 0,11$	$0,44 \pm 0,09$	$0,4 \pm 0,11$	$0,45 \pm 0,14$
100%	$0,42 \pm 0,12$	$0,50 \pm 0,18$	$0,54 \pm 0,2$	$0,4 \pm 0,12$	$0,41 \pm 0,09$

Tabela 3: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 5 e FastICA - HOC)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,77 \pm 0,09$	$0,77 \pm 0,1$	$0,81 \pm 0,07$	$0,75 \pm 0,12$	$0,77 \pm 0,11$
20%	$0,77 \pm 0,09$	$0,74 \pm 0,11$	$0,75 \pm 0,11$	$0,79 \pm 0,1$	$0,74 \pm 0,12$
30%	$0,77 \pm 0,09$	$0,81 \pm 0,12$	$0,75 \pm 0,07$	$0,73 \pm 0,12$	$0,77 \pm 0,13$
40%	$0,7 \pm 0,14$	$0,78 \pm 0,09$	$0,79 \pm 0,09$	$0,76 \pm 0,11$	$0,73 \pm 0,13$
50%	$0,75 \pm 0,15$	$0,79 \pm 0,09$	$0,76 \pm 0,13$	$0,73 \pm 0,14$	$0,66 \pm 0,17$
60%	$0,74 \pm 0,14$	$0,75 \pm 0,1$	$0,74 \pm 0,11$	$0,71 \pm 0,16$	$0,72 \pm 0,12$
70%	$0,74 \pm 0,12$	$0,77 \pm 0,12$	$0,74 \pm 0,13$	$0,66 \pm 0,15$	$0,6 \pm 0,17$
80%	$0,76 \pm 0,09$	$0,63 \pm 0,15$	$0,58 \pm 0,21$	$0,57 \pm 0,12$	$0,55 \pm 0,16$
90%	$0,66 \pm 0,15$	$0,62 \pm 0,14$	$0,51 \pm 0,14$	$0,55 \pm 0,15$	$0,49 \pm 0,12$
100%	$0,56 \pm 0,14$	$0,54 \pm 0,17$	$0,51 \pm 0,11$	$0,46 \pm 0,11$	$0,47 \pm 0,11$

Tabela 4: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 4 e Infomax - HOC)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,78 \pm 0,1$	$0,81 \pm 0,08$	$0,76 \pm 0,09$	$0,8 \pm 0,1$	$0,76 \pm 0,09$
20%	$0,79 \pm 0,09$	$0,81 \pm 0,1$	$0,74 \pm 0,09$	$0,75 \pm 0,09$	$0,83 \pm 0,1$
30%	$0,77 \pm 0,07$	$0,83 \pm 0,09$	$0,74 \pm 0,13$	$0,72 \pm 0,12$	$0,79 \pm 0,15$
40%	$0,78 \pm 0,09$	$0,79 \pm 0,08$	$0,77 \pm 0,12$	$0,74 \pm 0,13$	$0,75 \pm 0,15$
50%	$0,75 \pm 0,12$	$0,74 \pm 0,12$	$0,76 \pm 0,11$	$0,71 \pm 0,15$	$0,72 \pm 0,16$
60%	$0,66 \pm 0,15$	$0,69 \pm 0,18$	$0,67 \pm 0,12$	$0,68 \pm 0,19$	$0,65 \pm 0,15$
70%	$0,61 \pm 0,18$	$0,56 \pm 0,22$	$0,57 \pm 0,19$	$0,65 \pm 0,16$	$0,55 \pm 0,16$
80%	$0,58 \pm 0,15$	$0,52 \pm 0,14$	$0,55 \pm 0,19$	$0,51 \pm 0,15$	$0,5 \pm 0,15$
90%	$0,51 \pm 0,18$	$0,5 \pm 0,2$	$0,5 \pm 0,17$	$0,44 \pm 0,08$	$0,42 \pm 0,09$
100%	$0,42 \pm 0,12$	$0,42 \pm 0,13$	$0,45 \pm 0,17$	$0,43 \pm 0,12$	$0,42 \pm 0,09$

Tabela 5: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 5 e Infomax - HOC)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,81 \pm 0,08$	$0,81 \pm 0,09$	$0,83 \pm 0,09$	$0,78 \pm 0,11$	$0,76 \pm 0,1$
20%	$0,81 \pm 0,13$	$0,79 \pm 0,09$	$0,81 \pm 0,1$	$0,82 \pm 0,1$	$0,78 \pm 0,13$
30%	$0,78 \pm 0,12$	$0,79 \pm 0,08$	$0,8 \pm 0,09$	$0,81 \pm 0,11$	$0,78 \pm 0,1$
40%	$0,79 \pm 0,12$	$0,81 \pm 0,09$	$0,79 \pm 0,08$	$0,79 \pm 0,12$	$0,81 \pm 0,12$
50%	$0,79 \pm 0,13$	$0,77 \pm 0,11$	$0,79 \pm 0,12$	$0,78 \pm 0,1$	$0,79 \pm 0,09$
60%	$0,79 \pm 0,12$	$0,76 \pm 0,09$	$0,78 \pm 0,09$	$0,74 \pm 0,09$	$0,78 \pm 0,1$
70%	$0,76 \pm 0,11$	$0,75 \pm 0,09$	$0,76 \pm 0,1$	$0,78 \pm 0,13$	$0,76 \pm 0,11$
80%	$0,78 \pm 0,1$	$0,75 \pm 0,09$	$0,75 \pm 0,08$	$0,76 \pm 0,11$	$0,7 \pm 0,12$
90%	$0,81 \pm 0,1$	$0,7 \pm 0,15$	$0,78 \pm 0,14$	$0,73 \pm 0,1$	$0,69 \pm 0,17$
100%	$0,72 \pm 0,11$	$0,76 \pm 0,12$	$0,75 \pm 0,11$	$0,7 \pm 0,15$	$0,67 \pm 0,12$

Tabela 6: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 4 e FastICA - Hjorth e HOC)

k	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,74 \pm 0,11$	$0,68 \pm 0,08$	$0,65 \pm 0,14$	$0,78 \pm 0,09$	$0,65 \pm 0,13$
20%	$0,75 \pm 0,09$	$0,75 \pm 0,1$	$0,74 \pm 0,11$	$0,73 \pm 0,12$	$0,75 \pm 0,12$
30%	$0,73 \pm 0,15$	$0,71 \pm 0,14$	$0,73 \pm 0,11$	$0,79 \pm 0,08$	$0,75 \pm 0,11$
40%	$0,68 \pm 0,17$	$0,72 \pm 0,12$	$0,76 \pm 0,09$	$0,82 \pm 0,09$	$0,74 \pm 0,1$
50%	$0,75 \pm 0,13$	$0,76 \pm 0,18$	$0,72 \pm 0,12$	$0,72 \pm 0,09$	$0,76 \pm 0,09$
60%	$0,69 \pm 0,17$	$0,71 \pm 0,11$	$0,73 \pm 0,09$	$0,7 \pm 0,15$	$0,71 \pm 0,14$
70%	$0,71 \pm 0,12$	$0,71 \pm 0,1$	$0,75 \pm 0,1$	$0,71 \pm 0,13$	$0,76 \pm 0,08$
80%	$0,67 \pm 0,15$	$0,7 \pm 0,13$	$0,69 \pm 0,13$	$0,69 \pm 0,14$	$0,68 \pm 0,16$
90%	$0,62 \pm 0,17$	$0,63 \pm 0,2$	$0,67 \pm 0,12$	$0,63 \pm 0,17$	$0,66 \pm 0,16$
100%	$0,63 \pm 0,14$	$0,65 \pm 0,14$	$0,62 \pm 0,11$	$0,63 \pm 0,13$	$0,62 \pm 0,13$

Tabela 7: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 5 e FastICA - Hjorth e HOC)

K	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,84 \pm 0,1$	$0,84 \pm 0,08$	$0,78 \pm 0,1$	$0,77 \pm 0,07$	$0,83 \pm 0,12$
20%	$0,8 \pm 0,1$	$0,83 \pm 0,1$	$0,78 \pm 0,11$	$0,74 \pm 0,13$	$0,82 \pm 0,11$
30%	$0,76 \pm 0,09$	$0,78 \pm 0,1$	$0,78 \pm 0,1$	$0,79 \pm 0,09$	$0,75 \pm 0,12$
40%	$0,8 \pm 0,08$	$0,79 \pm 0,11$	$0,77 \pm 0,12$	$0,76 \pm 0,12$	$0,77 \pm 0,1$
50%	$0,78 \pm 0,09$	$0,76 \pm 0,12$	$0,77 \pm 0,11$	$0,77 \pm 0,12$	$0,79 \pm 0,09$
60%	$0,77 \pm 0,11$	$0,78 \pm 0,11$	$0,75 \pm 0,15$	$0,77 \pm 0,12$	$0,79 \pm 0,1$
70%	$0,76 \pm 0,11$	$0,8 \pm 0,11$	$0,76 \pm 0,14$	$0,73 \pm 0,11$	$0,74 \pm 0,09$
80%	$0,75 \pm 0,12$	$0,81 \pm 0,09$	$0,75 \pm 0,12$	$0,78 \pm 0,11$	$0,78 \pm 0,13$
90%	$0,72 \pm 0,07$	$0,74 \pm 0,1$	$0,74 \pm 0,1$	$0,73 \pm 0,13$	$0,77 \pm 0,09$
100%	$0,72 \pm 0,07$	$0,77 \pm 0,11$	$0,75 \pm 0,07$	$0,72 \pm 0,14$	$0,7 \pm 0,14$

Tabela 8: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 4 e Infomax - Hjorth e HOC)

K	8	9	10	11	12
	$A \pm \Delta A$				
10%	$0,73 \pm 0,1$	$0,72 \pm 0,1$	$0,72 \pm 0,1$	$0,78 \pm 0,11$	$0,75 \pm 0,09$
20%	$0,77 \pm 0,09$	$0,73 \pm 0,08$	$0,79 \pm 0,1$	$0,75 \pm 0,11$	$0,71 \pm 0,1$
30%	$0,76 \pm 0,11$	$0,8 \pm 0,1$	$0,8 \pm 0,09$	$0,76 \pm 0,12$	$0,71 \pm 0,12$
40%	$0,77 \pm 0,1$	$0,83 \pm 0,09$	$0,82 \pm 0,08$	$0,76 \pm 0,1$	$0,79 \pm 0,11$
50%	$0,76 \pm 0,13$	$0,76 \pm 0,09$	$0,79 \pm 0,1$	$0,79 \pm 0,1$	$0,76 \pm 0,1$
60%	$0,76 \pm 0,09$	$0,76 \pm 0,11$	$0,77 \pm 0,12$	$0,77 \pm 0,1$	$0,77 \pm 0,09$
70%	$0,74 \pm 0,11$	$0,77 \pm 0,09$	$0,75 \pm 0,09$	$0,74 \pm 0,14$	$0,74 \pm 0,14$
80%	$0,72 \pm 0,14$	$0,76 \pm 0,1$	$0,74 \pm 0,16$	$0,74 \pm 0,1$	$0,7 \pm 0,15$
90%	$0,71 \pm 0,09$	$0,7 \pm 0,11$	$0,67 \pm 0,08$	$0,72 \pm 0,12$	$0,71 \pm 0,1$
100%	$0,69 \pm 0,12$	$0,66 \pm 0,11$	$0,67 \pm 0,12$	$0,65 \pm 0,15$	$0,56 \pm 0,16$

Tabela 9: Acurácia do classificador para diferentes cenários (Sinais dos vídeos 1 e 5 e Infomax - Hjorth e HOC)