

Influência da Janela de Estimação no Desempenho de Classificação em Sistemas BCI-SSVEP

Henrique Luiz Voni Giuliani, Patrick O. de Paula, Paula G. Rodrigues,
Diogo C. Soriano, Ricardo Suyama, Denis G. Fantinato

Resumo— Interfaces Cérebro-Computador (BCI, do inglês, *Brain-Computer Interface*) provêm um canal alternativo de comunicação entre o usuário e uma máquina. Idealmente, o sistema deve ser capaz de classificar os dados coletados com alta acurácia e baixa latência. Dessa forma, nesse trabalho estudamos o efeito do tamanho da janela de amostras no desempenho da interface, utilizando técnicas de regressão logística e rede perceptron multicamadas para classificação dos dados, visando diminuir o tempo necessário para identificação da intenção do usuário. Os resultados revelam que, a despeito da simplicidade dos classificadores, é possível obter taxas de acerto superiores a 78% mesmo para janelas de apenas 1 segundo de duração.

Palavras-Chave— BCI, Regressão Logística, Redes Neurais Artificiais, MLP

Abstract— Brain Computer Interfaces (BCI) provide an alternative communication channel between the user and a machine. Ideally, the system should be able to classify the collected data with high accuracy and low latency. Therefore, in this work, we studied the effect of the sample window length on the performance of the interface, using logistic regression and multilayer perceptron network for data classification, aiming to decrease the overall system latency. The results reveal that, despite the simplicity of the classifiers, it is possible to obtain a classification accuracy higher than 78% even for 1s sample windows.

Keywords— BCI, Logistic Regression, Artificial Neural Networks, MLP

I. INTRODUÇÃO

A comunicação das intenções de um usuário a uma máquina pode ser realizada sem o uso das vias convencionais (nervos e/ou músculos), como acontece em sistemas de Interface Cérebro-Computador (BCI, do inglês *Brain-Computer Interface*) [1]. Nesses sistemas a comunicação é realizada a partir da interpretação dos sinais eletrofisiológicos corticais, sendo particularmente úteis para tecnologias assistivas e aplicações ligadas à saúde, ou até mesmo para entretenimento [1], [2].

Tipicamente, a operação dos sistemas BCI não invasivos se dá através da aquisição de sinais de eletroencefalografia (EEG), o qual se associa com o registro dos campos elétricos decorrentes da atividade eletroquímica neuronal por meio de eletrodos posicionados no escalpo [2]. No entanto, devido à baixa amplitude dos sinais e à presença demasiada de

Henrique L. V. Giuliani, Patrick O. de Paula, Paula G. Rodrigues, Diogo C. Soriano, Ricardo Suyama e Denis G. Fantinato, Universidade Federal do ABC, Santo André, e-mails: patrick.oliveira@aluno.ufabc.edu.br, {henrique.voni, paula.rodrigues, diogo.soriano, ricardo.suyama, denis.fantinato}@ufabc.edu.br. Este trabalho foi parcialmente financiado por FAPESP (2019/17997-4 e 2020/10014-2) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

ruído e artefatos, os dados de EEG precisam ser tratados para sua correta interpretação. Existem diversas técnicas em processamento de sinais passíveis de serem utilizadas para esse propósito, mas a escolha adequada depende do paradigma (i.e., da forma de indução de um padrão elétrico estável e passível de detecção e mapeamento) empregado [2]. No contexto de potenciais evocados visualmente em regime estacionário (SSVEP, do inglês *Steady State Visually Evoked Potential*), fontes visuais externas oscilam em um conjunto de frequências pré-definidas e geram uma resposta em sinais cerebrais, de forma inconsciente, que apresentam padrões bem definidos [1], [3]. Nesse caso, métodos seletores de bandas e de análise espectral (Transformada de Fourier, de Welch e Wavelet) ou mesmo métodos estatísticos (análise de correlação canônica) podem ser utilizados [2], [3].

Após o processamento, a etapa de classificação recebe os dados tratados e gera uma saída que indica a ação ou comando a ser tomado na máquina. O classificador é geralmente baseado em técnicas de aprendizado de máquina, com estruturas que podem ser linear – como a regressão logística – ou não linear – como as redes neurais artificiais (ANNs, do inglês *Artificial Neural Networks*).

O uso dessas ferramentas permite o desenvolvimento de um sistema BCI robusto com taxas de acerto em torno de 90% em uma quantidade considerável de experimentos no contexto de SSVEP [1], [2], [3]. Geralmente, o bom desempenho observado está relacionado ao uso de uma *janela de estimação* dos sinais de EEG consideravelmente extensa (em torno de 2 ou 3 segundos). Entretanto, para um sistema BCI com resposta rápida, torna-se imprescindível o uso de janelas de estimação menores, o que permitiria, além de um aprimoramento para os sistemas atuais, um maior horizonte de aplicações. Nesse sentido, o presente trabalho busca avaliar o desempenho desses sistemas usando períodos de aquisição de apenas 1 segundo em dados de (SSVEP) EEG. O sistema BCI proposto envolverá o uso de técnicas como o CAR (do inglês *Common Average Reference*), FFT (do inglês *Fast Fourier Transform*) e, para a classificação, serão comparadas duas estruturas: (i) linear baseado em regressão logística e, (ii) não linear, através de um Perceptron Multicamadas (MLP, do inglês *Multilayer Perceptron*) [4]. A comparação entre essas duas estruturas irá delinear qual abordagem é mais promissora para a redução da janela de estimação.

O trabalho está organizado da seguinte forma. Na Seção II, o sistema BCI-SSVEP é detalhado, sendo explicado os blocos que o compõem e as técnicas usadas em cada etapa. A Seção III apresenta a configuração do conjunto de dados

utilizado nos experimentos. Na Seção IV, estão organizados os resultados e discussões dos experimentos. Por fim, na Seção V, resumimos as principais conclusões do trabalho e sugerimos possíveis trabalhos futuros decorrentes da presente pesquisa.

II. SISTEMAS DE INTERFACE CÉREBRO-COMPUTADOR

Uma Interface Cérebro-Computador é um sistema de comunicação que permite a transmissão direta de informação entre o cérebro e uma aplicação de computador. Embora a implementação de um sistema BCI robusto seja uma tarefa tecnológica complexa, por envolver a integração de diversas áreas do conhecimento, a mesma mostra-se extremamente relevante no âmbito das tecnologias assistivas, oferecendo melhor qualidade de vida a seus usuários, principalmente àqueles com severas limitações motoras ou de comunicação associadas a uma ampla variedade de traumas (e.g. lesão da medula espinhal) ou doenças neurodegenerativas (e.g. esclerose lateral amiotrófica) [1], [2].

Dentre os paradigmas de BCI existentes, um dos mais utilizados é o SSVEP. O paradigma caracteriza-se pelo sincronismo da atividade elétrica do córtex visual com a estimulação visual externa em uma dada frequência [1]. Mediante diferentes opções de estímulos simultâneos apresentados (variando em frequência ou em fase), tem-se a possibilidade de seleção de um em especial (a ser mapeado em um comando externo) por meio da modulação da atenção do usuário [1], [2].

Uma vez que a abordagem SSVEP induz uma resposta nas frequências de estímulo que são conhecidas, é possível obter bom desempenho mesmo sem sessões de treinamento, tal como exigido, por exemplo, no paradigma BCI de imagética motora. Entretanto, surgem dificuldades relativas ao processamento/extração da informação de interesse, visto que os sinais elétricos cerebrais possuem amplitude extremamente baixa (5–20 μV) [5] e estão sujeitos a um processo de mistura intrínseco a comunicação eletroquímica neuronal, repleta de artefatos e interferência. Um sistema BCI, sob a perspectiva de uma tarefa de processamento de sinais, pode ser compreendido em quatro etapas [3]: pré-processamento, extração de características, seleção de atributos e classificação, conforme ilustrado na Fig. 1. Estas serão detalhadas a seguir.

A. Pré-processamento

A fim de remover dos sinais a presença de artefatos decorrentes de movimentos musculares e piscar de olhos, ou análogos, os dados são processados por um filtro espacial utilizando, por exemplo, o método CAR (do inglês *Common Average Reference*) [1], [3], que subtrai a média dos eletrodos, amostra a amostra:

$$V_i^{CAR} = V_i^{ER} - \frac{1}{N_e} \sum_{j=1}^{N_e} V_j^{ER} \quad (1)$$

em que V_j^{ER} é o potencial do i -ésimo eletrodo medido com respeito a um mesmo referencial e N_e é o número de eletrodos. Pela remoção da média das medições, o método é capaz de remover artefatos comuns ao conjunto de eletrodos.

B. Extração de características

Um dos métodos mais comuns para a extração de características em sinais de EEG no contexto de SSVEP é a Transformada Discreta de Fourier [6], dada por

$$X_i(k) = \sum_{n=0}^{N-1} x_i(n) \cdot \exp(-j2\pi f_k n) \quad (2)$$

em que $x_i(n)$ é a n -ésima amostra do sinal correspondente ao i -ésimo eletrodo (aqui assumido como $x_i(n) = V_i^{CAR}(n)$), $f_k = k/N$, $k = 0, \dots, N-1$, sendo N o tamanho da janela ou período. A Eq. (2) geralmente é calculada pelo algoritmo da Transformada Rápida de Fourier (FFT, do inglês *Fast Fourier Transform*) [6]. O sinal transformado para o domínio da frequência X_i pode ser então analisado. Usualmente, obtêm-se a Densidade Espectral de Potência (PSD, do inglês *Power Spectral Density*) [6]:

$$P_i(k) = \frac{1}{N} |X_i(k)|^2 \quad (3)$$

em que $|\cdot|$ é o operador de valor absoluto, sendo possível identificar as frequências de maior energia, que, no contexto de SSVEP, estarão associadas às frequências de estímulo [3].

C. Seleção de atributos

A etapa de seleção de atributos tem por objetivo determinar as características mais informativas a serem usadas para se conceber o sistema de classificação. Esta etapa é importante para tornar mais rápido o treinamento e a operação da máquina, e tornar o processo de aprendizado mais efetivo em termos de capacidade de generalização a partir da eliminação de características muito ruidosas ou com baixo teor informativo [7]. Destacam-se nesse âmbito duas estratégias de seleção de características: *filtros* e *wrappers* [3], [7]. A primeira delas faz uso de medidas estatísticas gerais para quantificar a relevância de cada característica, dentre os quais podemos citar o coeficiente de correlação de Pearson e o discriminante de Fisher [8]. Por outro lado, na segunda estratégia, utiliza-se o treinamento da máquina na avaliação comparativa dos atributos [7], sendo comum a utilização de heurísticas para a construção de *wrappers*, e.g., nos métodos de *forward selection* e *backward selection* [3].

Embora pesquisas relacionadas [9] – que utilizaram o mesmo conjunto de dados que o presente trabalho – façam uso da etapa de seleção de características, estas podem exibir elevada complexidade computacional (principalmente na abordagem *wrapper*) e, além disso, apresentar resultados que podem variar ao longo do tempo. Por este motivo, não consideraremos seu uso neste trabalho, o que objetivou estabelecer um comparativo de desempenho no âmbito de diferentes estruturas de classificação.

D. Classificação

A etapa de classificação é a parte do sistema responsável por mapear os sinais de entrada às saídas (ou classes) correspondentes, sendo que cada sinal de entrada é representado por um conjunto de características e está associado a apenas uma possível saída [4].

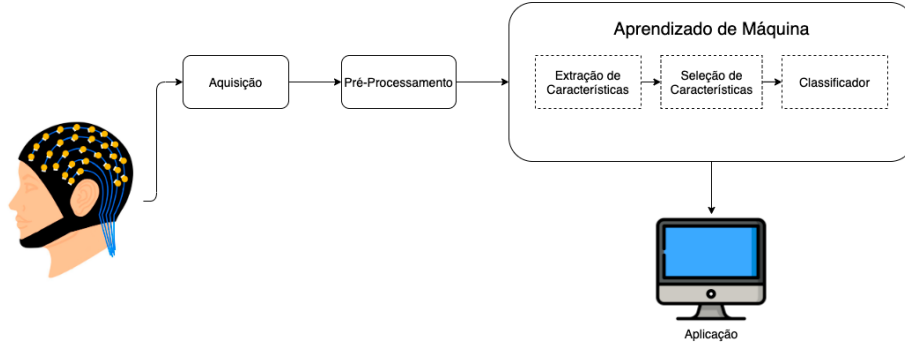


Fig. 1: Etapas de um sistema BCI

O objetivo da tarefa de classificação é encontrar, através do conjunto de características das entradas, a fronteira de decisão que delimita corretamente o espaço de cada classe. Tal fronteira pode ser obtida empregando algoritmos de aprendizado de máquina no caso supervisionado, onde os parâmetros da estrutura utilizada adapta-se iterativamente a casos conhecidos (rotulados) e é capaz de classificar entradas desconhecidas. Neste trabalho, foram utilizadas duas estruturas distintas: a Regressão Logística (LR, do inglês *Logistic Regression*), para o caso linear, e a rede MLP, para o caso não linear [4], [10].

1) *Regressão Logística*: Dado um conjunto de T amostras $\{\mathbf{x}_i\}_{i=1}^T$, sendo $\mathbf{x}_i \in \mathbb{R}^m$ o vetor coluna com os atributos, uma estrutura linear realiza uma ponderação sobre os valores de \mathbf{x}_i , i.e. $\mathbf{w}^T \mathbf{x}_i + b$, em que $\mathbf{w} \in \mathbb{R}^m$ é um vetor coluna com pesos ajustáveis e $b \in \mathbb{R}$ um termo de viés. No contexto de classificação de N_c classes, é possível decompô-lo em $N_c - 1$ problemas de classificação binária (i.e. envolvendo duas classes denotadas por 0 e 1, ou negativa e positiva) e podem ser abordados pela técnica de Regressão Logística [4]. Nesse caso, cada amostra é mapeada no intervalo $(0, 1)$ pela função logística:

$$\sigma(\mathbf{x}_i) = \frac{1}{1 + e^{-(\mathbf{w}^T \mathbf{x}_i + b)}}, \quad (4)$$

em que $\sigma(\mathbf{x}) = y_i$ é interpretada como a probabilidade de x_i pertencer à classe positiva. Para cada amostra \mathbf{x}_i está associado um elemento $\hat{y}_i \in \{0, 1\}$ indicando se \mathbf{x}_i pertence à classe positiva, possibilitando a construção da função de verossimilhança $\mathcal{L}(\mathbf{W}, b)$ com a qual o modelo é avaliado, definida como,

$$\mathcal{L}(\mathbf{w}, b) = \prod_{i=1}^T p(\hat{y}_i | \mathbf{x}_i, \mathbf{w}, b), \quad (5)$$

onde pode-se escrever

$$p(\hat{y}_i | \mathbf{x}_i, \mathbf{w}, b) = \sigma(\mathbf{x}_i)^{\hat{y}_i} (1 - \sigma(\mathbf{x}_i))^{1 - \hat{y}_i}. \quad (6)$$

O problema de ajuste do modelo passa a ser o de encontrar os pesos \mathbf{w} e b com os quais $\mathcal{L}(\mathbf{w}, b)$ é maximizada. Com alguma manipulação algébrica, chega-se ao problema de otimização equivalente [4]

$$\min_{\mathbf{W}, b} \sum_{i=1}^N \log(1 + \exp[-\hat{y}_i(\mathbf{W}^T \mathbf{x}_i + b)]). \quad (7)$$

Nos problemas de classificação envolvendo múltiplas classes (abordagem *one-versus-all*), a classe final de x_i é, assim, aquela de maior probabilidade.

2) *Multilayer Perceptron*: Uma rede *Multilayer Perceptron* é a arquitetura fundamental de ANNs. Sua estrutura é construída por uma sequência de L camadas, onde as camadas 1 e L são, respectivamente, as camadas de entrada e saída. Para cada camada $l = 2, \dots, L$ define-se uma matriz de pesos $\mathbf{W}^{[l]} \in \mathbb{R}^{n_l \times n_{l-1}}$ e vieses $\mathbf{b}^{[l]} \in \mathbb{R}^{n_l}$, todos ajustáveis. De modo geral, a rede é um mapeamento de \mathbb{R}^{n_1} em \mathbb{R}^{n_L} . Dado um input $\mathbf{x} \in \mathbb{R}^{n_1}$ e denotando por $y^{[l]}$ a saída na camada l , tem-se que $\mathbf{y}^{[1]} = \mathbf{x}$ para a primeira camada e, para as demais:

$$\mathbf{y}^{[l]} = \sigma(\mathbf{W}^{[l]} \mathbf{y}^{[l-1]} + \mathbf{b}^{[l]}) \quad (8)$$

em que a função $\sigma(\cdot)$ é denominada *função de ativação*, podendo ser, por exemplo, as funções Sigmoid, Tangente Hiperbólica e ReLU [11].

Com um conjunto de dados \mathbf{X} de T amostras em \mathbb{R}^{n_1} , $\{\mathbf{x}_i\}_{i=1}^T$, associados a valores objetivo $\{\hat{y}_i\}_{i=1}^T$ em \mathbb{R}^{n_L} , pode-se definir uma função de custo $\mathcal{L}(\mathbf{W}, \mathbf{b})$ parametrizada pelos pesos adaptativos da rede. Para problemas de classificação com $C \geq 2$ classes, usualmente a função de custo utilizada é a Entropia Cruzada, definida como

$$\mathcal{L}(\mathbf{W}, \mathbf{b}) = -\frac{1}{T} \sum_{i=1}^T \sum_{j=1}^C y_{ij}^{[L]} \log \hat{y}_{ij}. \quad (9)$$

Assim, o problema de aprendizado dos melhores parâmetros para a rede pode ser expresso de acordo com o seguinte problema de otimização:

$$\min_{\mathbf{W}, \mathbf{b}} \mathcal{L}(\mathbf{W}, \mathbf{b}). \quad (10)$$

O método iterativo mais usual para o ajuste dos pesos da rede é o algoritmo de *Backpropagation* [10], em que o ajuste se faz proporcionalmente e na direção contrária ao do gradiente do custo \mathcal{L} em relação aos pesos:

$$\mathbf{W}_{i+1} \rightarrow \mathbf{W}_i - \eta \frac{\partial \mathcal{L}(\mathbf{W}_i, \mathbf{b}_i)}{\partial \mathbf{W}_i} \quad (11)$$

$$\mathbf{b}_{i+1} \rightarrow \mathbf{b}_i - \eta \frac{\partial \mathcal{L}(\mathbf{W}_i, \mathbf{b}_i)}{\partial \mathbf{b}_i} \quad (12)$$

onde $\eta \in \mathbb{R}$ é a *taxa de aprendizado*.

III. CONJUNTO DE DADOS DE EEG

O conjunto de dados utilizado para a realização dos experimentos deste trabalho é uma contribuição de pesquisas precedentes desenvolvidas pelo Laboratório de Métodos Computacionais para a Bioengenharia da Universidade Federal do ABC¹ [9]. A coleta de sinais de EEG foi feita utilizando-se uma touca com 16 eletrodos, posicionados de acordo com o sistema 10-20. 15 voluntários (5 homens e 10 mulheres) saudáveis participaram das 12 sessões de 6 segundos para cada uma das 4 frequências de interesse (10, 11, 12 e 13Hz) – somando 48 sessões por sujeito, conforme demonstrado na Fig. 2. A taxa de amostragem para a coleta dos dados é de 256Hz.

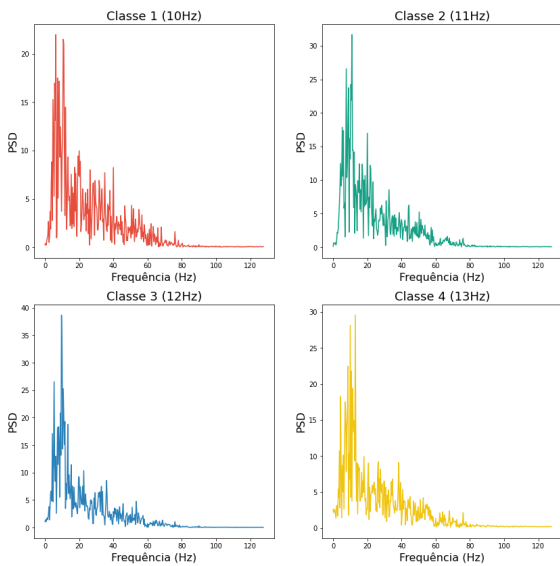


Fig. 2: PSD do sinal captado pelo eletrodo Oz para as quatro frequências de interesses em uma amostra de 3 segundos.

As sessões de coleta fizeram uso de um monitor de 22 polegadas e atualização de 60Hz que apresenta ao usuário uma interface com um número determinado de fontes luminosas em padrão xadrez, que cintilam em diferentes frequências, sendo cada uma delas associada a um determinado comando externo conforme ilustrado pela Figura 3. A classificação de um sinal pode ser feita através da identificação do aumento de energia em uma dada frequência.

Para a realização dos experimentos, os dados foram particionadas em janelas de tempo de 3, 2 e 1 segundos com uma sobreposição de 0,5 segundos (ou 128 pontos); totalizando para cada caso, respectivamente: 192, 240 e 288 amostras por indivíduo. Vale ressaltar que o uso de janelas ou períodos de aquisição iguais a 3 e 2 segundos já foram previamente utilizados na literatura [9]. No entanto, o caso com período de 1 segundo, por possuir uma resolução menor e ser consequentemente mais desafiador, ainda carece de maior atenção.

IV. RESULTADOS

O sistema BCI considerado será composto por: (i) o filtro CAR para o pré-processamento; (ii) o algoritmo FFT para

¹Aprovado pelo Comitê de Ética em Pesquisa da UFABC, CAAE:51005615.2.0000.5594.

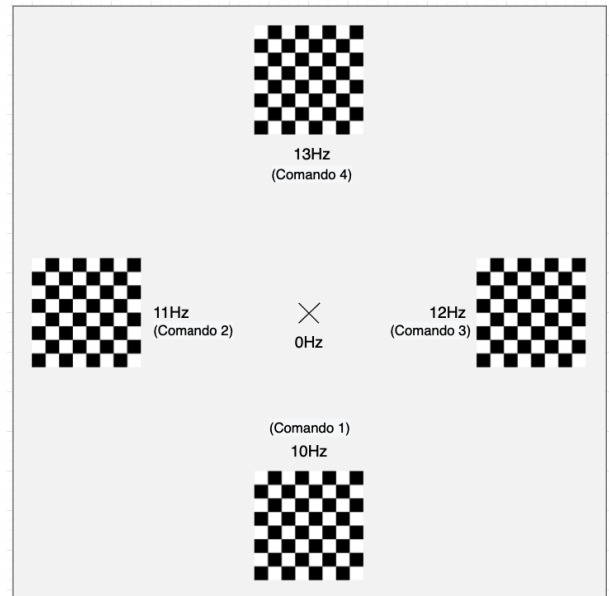


Fig. 3: Estratégia de coleta de dados pela BCI-SSVEP

extração das características, i.e., as três primeiras harmônicas da frequência de interesse (os dados foram previamente normalizados com média zero e variância unitária, com a finalidade de manter todos os atributos na mesma escala); e (iii) um classificador, que poderá ser linear (regressão logística) ou não linear (MLP). A etapa usual de seleção de atributos não será utilizada, pois exige processamento considerável e pode variar consideravelmente ao longo do tempo, podendo fazer o desempenho do classificador se degradar. Dessa forma, ao suprimirmos essa etapa, buscamos um sistema com menor tempo de calibração e um classificador mais robusto.

As redes MLP foram configuradas com 1 camada intermediária com 15 neurônios, função de ativação sigmóide e número máximo de 2000 épocas. Para o classificador linear baseado em Regressão Logística, o único hiperparâmetro configurado foi o limite máximo de 2000 iterações.

A respeito das configurações do experimento, foram utilizados 50% dos dados para treinamento e o restante para avaliação/teste adotando a acurácia como métrica principal.

1) *Resultados com Janelas de 3 e 2 Segundos:* As Tabs. I e II exibem a taxa de acurácia, para os casos com períodos de 3 e 2 segundos, respectivamente, usando-se o sistema BCI-SSVEP com os classificadores LR e a rede MLP. Torna-se evidente que a LR e a MLP possuem um desempenho médio equivalente, mantendo altas taxas de acerto. Assim, não há uma vantagem expressiva que justifique utilizar estruturas mais complexas. Dado que a quantidade de informação relevante contida nas amostras de 2 e 3 segundos seja adequada para realizar a classificação, um classificador linear é suficiente para alcançar bons resultados. Em muitos dos casos, principalmente com a janela de 3 segundos, ambos os classificadores alcançaram acurácia máxima. Também é interessante que o desempenho médio decaiu sutilmente entre os cenários de 3 e 2 segundos, mas ficando ainda acima dos 90% de acurácia.

2) *Resultados com Janela de 1 Segundo:* Com a finalidade de investigar o comportamento do sistema BCI-SSVEP em

TABELA I: Resultados obtidos com janelas de 3 segundos.

Indivíduo	LR	MLP
1	95,83%	95,83%
2	98,95%	97,91%
3	100,00%	100,00%
4	100,00%	100,00%
5	96,88%	96,87%
6	96,87%	95,83%
7	79,16%	85,41%
8	100,00%	100,00%
9	87,50%	86,45%
10	98,95%	100,00%
11	100,00%	100,00%
12	100,00%	100,00%
13	98,95%	100,00%
14	100,00%	100,00%
15	100,00%	100,00%
MÉDIA	96,87%	97,22%

TABELA III: Resultados obtidos com janelas de 1 segundo.

Indivíduo	LR	MLP
1	87,50%	87,50%
2	65,27%	68,05%
3	87,50%	88,19%
4	90,97%	92,36%
5	81,25%	86,11%
6	83,33%	79,86%
7	48,61%	49,30%
8	92,36%	93,75%
9	51,38%	56,25%
10	76,38%	79,16%
11	84,72%	84,72%
12	86,11%	86,11%
13	72,22%	75,69%
14	99,30%	99,30%
15	90,97%	92,36%
MÉDIA	79,86%	81,25%

TABELA II: Resultados obtidos com janelas de 2 segundos.

Indivíduo	LR	MLP
1	93,33%	94,16%
2	91,66%	92,50%
3	99,16%	98,33%
4	99,16%	98,33%
5	93,33%	95,00%
6	98,33%	96,66%
7	75,00%	69,16%
8	99,16%	99,16%
9	86,66%	90,00%
10	94,16%	93,33%
11	100,00%	98,33%
12	99,16%	99,16%
13	95,00%	92,50%
14	100,00%	100,00%
15	99,16%	98,33%
MÉDIA	94,88%	94,33%

cenários mais desafiadores, utilizamos janela de estimação de 1 segundo. Os resultados obtidos em termos da taxa de acerto para o sistema BCI com classificador LR e MLP são discriminados na Tab. III.

A partir do desempenho dos classificadores neste cenário, fica evidente uma maior dificuldade dos classificadores em lidar com os dados e suas acurácias apresentam maior diferença. Ao mesmo tempo, os resultados sugerem que classificadores como a MLP, cuja natureza permite lidar com casos não lineares, começam a se sobressair e obter maiores taxas de acerto. Essa observação fica aparente nos dados dos indivíduos 5 e 9, onde os ganhos chegam a 5% nas taxas de acerto.

Em outras palavras, a complexidade dos classificadores pode estar associada ao tamanho da janela de estimação utilizada, de forma que janelas de tempo menores exijam classificadores mais robustos.

De uma forma geral, com a janela de estimação de 1 segundo, conseguiu-se alcançar uma taxa de acerto médio de 81,25% com classificadores MLP. Isto indica que é possível adotar o uso de janelas de estimação de tamanho reduzido, diminuindo a latência nos sistemas BCI.

V. CONCLUSÕES

No contexto dos sistemas BCI-SSVEP, diminuir a janela de estimação utilizada para realizar a classificação do sinal

é uma tarefa desafiadora. Neste trabalho, sinais de EEG de 15 indivíduos foram utilizados para a análise de desempenho adotando diferentes janelas de tempo, que incluía janelas de 3 e 2 segundos, e como contribuição desse trabalho, janelas de 1 segundo. O sistema BCI proposto combinou técnicas de CAR e FFT, e um classificador que poderia ser linear (ajustado através da Regressão Logística) ou não linear (uma rede MLP). Com as técnicas empregadas, foi possível obter um desempenho médio relativamente alto mesmo com a janela de 1 segundo, alcançando 81,25% de acurácia com a MLP. Além disso, os resultados indicaram que, à medida que a janela de tempo usada diminui, classificadores não lineares tornam-se necessários. Como continuidade deste trabalho e seguindo na mesma linha de pesquisa, faz-se valiosa a aplicação de *ensembles* para classificar os dados em janelas de tempo iguais ou inferiores a 1 segundo.

REFERÊNCIAS

- [1] C. S. Nam, A. Nijholt, and F. Lotte, *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*. CRC Press, 2018.
- [2] X. Gu, Z. Cao, A. Jolfaei, P. Xu, D. Wu, T.-P. Jung, and C.-T. Lin, "EEG-Based Brain-Computer Interfaces (BCIs): A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and their Applications," *arXiv preprint arXiv:2001.11337*, 2020.
- [3] S. N. Carvalho, T. B. Costa, L. F. Uribe, D. C. Soriano, G. F. Yared, L. C. Coradine, and R. Attux, "Comparative Analysis of Strategies for Feature Extraction and Classification in SSVEP BCIs," *Biomedical Signal Processing and Control*, vol. 21, pp. 34–42, 2015.
- [4] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.
- [5] A. Kübler and K.-R. Müller, "An Introduction to Brain-Computer Interfacing," *Toward brain-computer interfacing*, pp. 1–25, 2007.
- [6] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.
- [7] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.
- [8] W. Duch, "Filter Methods," in *Feature Extraction*, pp. 89–117, Springer, 2006.
- [9] J. I. da Silva Jr., "Comparativo De Desempenho De Sistemas BCI-SSVEP Off-line E Em Tempo De Execução Utilizando Técnicas De Estimação De Espectro E Análise De Correlação Canônica," Master's thesis, UFABC, 2017.
- [10] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. MIT press Cambridge, 2016.
- [11] B. Ding, H. Qian, and J. Zhou, "Activation functions and their characteristics in deep neural networks," in *2018 Chinese Control And Decision Conference (CCDC)*, pp. 1836–1841, IEEE, 2018.