

Aprendizado Profundo Aplicado a Interfaces Cérebro-Computador Baseadas em Potenciais Visualmente Evocados em Regime Estacionário

Palavras-Chave: Processamento de sinais, interface cérebro-computador, aprendizado de máquina, *deep learning*

Aluno: Wesna Simone Bulla de Araujo

Orientador: Prof. Levy Boccatto

Faculdade de Engenharia Elétrica e de Computação (FEEC)

1 Introdução

Uma interface cérebro-computador (BCI, do inglês *brain-computer interface*) é um sistema de comunicação que visa permitir ao usuário realizar tarefas, como digitar uma tecla ou mover-se para uma direção, sem a utilização de nervos ou músculos periféricos, isto é, sem ter de produzir o movimento propriamente dito. Para isso, a interface deve interpretar os sinais cerebrais registrados a fim de gerar uma resposta correspondente à ação desejada [1].

De maneira geral, uma BCI inicia sua operação com a aquisição dos sinais cerebrais. Em seguida, o conjunto de observações passa por algumas etapas de processamento, as quais visam extrair as características mais representativas desses sinais, para, finalmente, classificá-los, gerando uma resposta correspondente ao comando escolhido pelo usuário, a qual é enviada ao dispositivo envolvido na aplicação [2].

Com o advento do aprendizado profundo (DL, do inglês *deep learning*), e suas várias aplicações de sucesso [3, 4], alguns estudos têm sido realizados buscando substituir algumas ou todas as etapas de tratamento dos dados por modelos profundos [5], com o propósito de automatizar a etapa de extração de características e, com isso, beneficiar o processamento dos sinais cerebrais em BCIs [6, 5].

No entanto, o projeto de uma BCI explorando o paradigma baseado em potenciais visualmente evocados em regime estacionário (SSVEP, do inglês *steady-state visually evoked potentials*) constitui um cenário de aprendizado que impõe alguns fatores possivelmente complicadores para DL. Primeiramente, a disponibilidade de dados de EEG é relativamente escassa: as bases públicas não são volumosas, e a preparação de uma base própria exige a criação de um rigoroso protocolo e envolve um processo relativamente demorado de coleta de dados com novos voluntários.

Em segundo lugar, dada a natureza do paradigma SSVEP, explorar o domínio da frequência parece ser a opção mais natural, já que é esperado que haja um aumento de energia no espectro em torno da frequência do estímulo (e também em suas harmônicas) em resposta à estimulação visual, especialmente nos sinais de EEG capturados por eletrodos

situados na região occipital do córtex, a qual é o centro do processamento visual do ser humano. Sendo assim, é natural considerar atributos espectrais durante o processamento, uma vez que a evocação tende a se manifestar mais nitidamente no domínio de frequência. Neste sentido, o treinamento de uma estrutura profunda, que parte de extratores de atributos genéricos e inicialmente aleatórios, parece ser, em certa medida, um caminho que coloca em segundo plano o conhecimento que já se tem das especificidades do problema.

O presente trabalho de iniciação científica buscou realizar um estudo sobre BCIs para avaliar os modos clássico e baseado em DL para a construção da interface considerando o paradigma SSVEP. Inicialmente, foram estudadas algumas das técnicas mais utilizadas em cada estágio: no pré-processamento, o filtro de média de referência comum (CAR, do inglês *common average reference*); na extração de características, a análise de correlações canônicas (CCA, do inglês *canonical correlation analysis*), a transformada discreta de Fourier (DFT, do inglês *discrete Fourier transform*) e o método do periodograma de Welch; na classificação, as máquinas de vetores-suporte (SVMs, do inglês *support-vector machines*) e as redes neurais convolucionais (CNNs, do inglês *convolutional neural networks*). Dentro da perspectiva de DL, foram exploradas as propostas de [7], que se baseia numa CNN-1D para processar os EEGs no domínio do tempo (ou seja, os dados brutos), e de [8], que explora uma CNN-1D cuja entrada corresponde à coleção de densidades espectrais de potência dos diferentes eletrodos.

Dessa forma, foi possível fazer comparações entre as duas metodologias e chegar a observações importantes sobre as possíveis vantagens e desvantagens de cada abordagem no contexto de BCI-SSVEP. O presente resumo está organizado da seguinte maneira: a Seção 2 apresenta as etapas constituintes de um sistema BCI-SSVEP com ênfase no processamento de sinais. As Seções 3 e 4 mostram a metodologia empregada para a realização da pesquisa, bem como os resultados obtidos. Finalmente, a Seção 5 traz as conclusões extraídas ao longo de todo o desenvolvimento do trabalho.

2 Processamento de Sinais em BCI-SSVEP

2.1 Pré-Processamento

A primeira etapa do processamento de sinais, denominada de pré-processamento, visa eliminar artefatos presentes nos sinais cerebrais através de técnicas de filtragem, e, com isso, melhorar a relação sinal-ruído [2]. Uma técnica simples, mas bastante utilizada para tal finalidade, é o filtro CAR.

2.1.1 CAR

Esta técnica realiza uma filtragem espacial que reduz as componentes que estão em um maior número de eletrodos [2], de modo a proporcionar uma referência neutra para os registros de EEG. Em cada instante de tempo, o filtro CAR faz a subtração entre o potencial aferido em cada eletrodo (medido em relação a um referencial neutro), e o valor médio do potencial observado em todos os eletrodos. A Equação (1) descreve a operação do filtro:

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

onde $V_i^{\text{ER}}(n)$ denota o potencial do i -ésimo eletrodo no n -ésimo instante de amostragem, e N_e representa o número de eletrodos disponíveis.

É pertinente mencionar que na abordagem SSVEP os ruídos tendem a aparecer em boa parte dos eletrodos com intensidade parecida, fazendo com que o filtro CAR ajuste a referência de potencial dos eletrodos e já atenua parte dos artefatos [2].

2.2 Extração De Características

Esta etapa é responsável por extrair os atributos mais discriminativos do conjunto de dados, criando uma representação mais compacta e informativa com o intuito de favorecer o desempenho do classificador. No contexto de BCI-SSVEP, o uso de técnicas que capturem informações ligadas ao domínio da frequência é muito comum, uma vez que a própria noção de SSVEP está diretamente relacionada com a aparição de componentes espectrais na frequência do estímulo e/ou em suas harmônicas. A seguir descrevemos brevemente três técnicas empregadas na segunda etapa do processamento de sinais.

2.2.1 CCA

A análise de correlações canônicas é uma ferramenta estatística multivariada utilizada para quantificar a semelhança entre dois conjuntos de dados. Para isso, ambos os conjuntos passam por uma redução de dimensionalidade, isto é, uma combinação linear, com o objetivo de se encontrar a máxima correlação entre eles [9].

Assim, supondo dois vetores de variáveis aleatórias, X e Y , cuja combinação linear é dada, respectivamente, por $a^T X$ e $b^T Y$, a CCA busca maximizar a correlação canônica entre as novas variáveis $a^T X$ e $b^T Y$:

$$\max_{a,b} \rho(a^T X, b^T Y) \quad (2)$$

Através de algumas manipulações algébricas [9], é possível reescrever o coeficiente de correlação da seguinte forma:

$$\rho(a,b) = \frac{a^T \sum_{XY} b}{(a^T \sum_{XX} a)^{1/2} (b^T \sum_{YY} b)^{1/2}}, \quad (3)$$

o qual deve ser maximizado com respeito aos vetores de coeficientes a e b , respeitando as seguintes restrições:

$$a^T \sum_{XX} a = 1$$

$$b^T \sum_{YY} b = 1,$$

sendo que $\sum_{XX} = \text{cov}(X, X)$, $\sum_{XY} = \text{cov}(X, Y)$ e $\sum_{YY} = \text{cov}(Y, Y)$, onde $\text{cov}(A, B)$ denota a matriz de covariância entre as variáveis aleatórias dos vetores A e B .

No caso de sistemas BCI-SSVEP, a matriz X corresponde ao conjunto de sinais de EEG e possui dimensão $(n \times m)$, onde n corresponde ao número de eletrodos e m denota a quantidade de amostras do sinal por janelamento [10]. Por sua vez, o conjunto Y , intitulado de matriz de referência, possui sinais formados por senos e cossenos que simulam as frequências de estímulo presentes em X . A dimensão de Y é dada por $(q \times m)$, em que q é o número de harmônicas consideradas para cada frequência de estímulo [10]. A Equação (4) traz um exemplo de matriz Y para duas harmônicas [11], onde f_0 denota a frequência de estímulo considerada, S é a taxa de amostragem do sinal e T é a quantidade de amostras.

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ y_4(t) \end{pmatrix} = \begin{pmatrix} \sin(2\pi f_0 t) \\ \cos(2\pi f_0 t) \\ \sin(4\pi f_0 t) \\ \cos(4\pi f_0 t) \end{pmatrix}, t = \frac{1}{S}, \frac{2}{S}, \dots, \frac{T}{S} \quad (4)$$

A solução do problema de otimização atrelado à CCA pode ser obtida com o auxílio da decomposição em valores singulares (SVD), conforme detalhado em [9].

No contexto de BCIs, a CCA pode ser empregada tanto na etapa de pré-processamento quanto na etapa de extração de características [12]. No segundo caso, foco do presente trabalho, a CCA fornece um vetor de atributos em que cada elemento é o valor máximo de correlação atingido entre os sinais de EEG dos eletrodos e a matriz Y de referência para cada frequência de estímulo possível.

Por fim, é pertinente mencionar que a quantidade de eletrodos deve ser maior ou igual ao dobro do número de frequências de estímulo, pois os coeficientes de correlação canônicos se tornam nulos ao ultrapassar o número de canais [12, 10].

2.2.2 FFT e Periodograma de Welch

A transformada discreta de Fourier é uma ferramenta matemática que permite transformar uma sequência discreta $x[n]$ de comprimento finito em outra sequência de mesmo comprimento, mas com informação no domínio da frequência. A Equação (5) mostra tal mapeamento [13]:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j(\frac{2\pi}{N})kn}, k = 0, \dots, N-1 \quad (5)$$

onde N é a quantidade de amostras em um período da transformada de Fourier. Para realizar o cálculo da DFT de forma eficiente, usamos os algoritmos da família FFT (do inglês, *fast Fourier transform*) [13].

Outra abordagem para a extração de características consiste em estimar a densidade espectral de potência (PSD, do inglês *power spectral density*) de segmentos do sinal

através de periodogramas [14]. A ideia por trás do método de Welch é dividir os dados disponíveis em K segmentos de mesmo comprimento N , calcular o periodograma para cada segmento e, finalmente, tomar a média dos resultados. Dessa forma, há uma redução na variância associada ao espectro de potência, tornando o estimador mais consistente [13]. Matematicamente, o método é descrito pela seguinte expressão:

$$\hat{I}(\omega) = \frac{1}{KNU} \sum_{k=1}^K \left| \sum_{n=1}^N w[n]x[n+kD]e^{-j\omega n} \right|^2, \quad (6)$$

onde $x[n]$ corresponde ao sinal de entrada amostrado, D denota a sobreposição de amostras entre janelas consecutivas, U é uma constante de normalização cuja fórmula encontra-se em [13], e $w[n]$ denota a janela temporal utilizada (e.g., retangular ou Hamming [13]).

2.3 Classificadores

A última etapa do módulo de processamento de sinais de uma BCI é a classificação. Esta tarefa está associada ao desafio de designar cada padrão de entrada a uma determinada classe, sendo que os objetos pertencentes à mesma classe tipicamente revelam um grau significativo de similaridade [15]. No contexto de BCI-SSVEP, a entrada do classificador é descrita pelo conjunto de atributos extraído nas etapas anteriores, e as classes existentes são as possíveis frequências de estímulo [2, 12].

Neste projeto, estudamos duas abordagens bastante famosas de classificação: SVMs e CNNs.

2.3.1 SVM

As máquinas de vetores-suporte são modelos de classificação que se apoiam na teoria de aprendizado estatístico [4] e buscam uma melhor generalização ao formalmente explorarem a noção de maximização da margem.

Classicamente, a marca das SVMs é a busca pelo hiperplano que separe os padrões de duas classes corretamente e que deixe a maior separação entre essas classes [15]. Intuitivamente, quanto mais próxima de uma amostra estiver a fronteira projetada, mais susceptível a erros de generalização estará o classificador. A separação entre as classes é quantificada através do conceito de margem, que corresponde à distância perpendicular entre a fronteira de classificação e os dados mais próximos a ela.

Duas extensões contribuíram significativamente para a consolidação das SVMs como uma abordagem poderosa de classificação [15]: (1) a incorporação de variáveis de relaxação à formulação do problema, o que permitiu que o modelo lidasse com dados que não fossem linearmente separáveis; e (2) o emprego de funções *kernel* para calcular os produtos internos entre as amostras em um espaço de características implícito (tipicamente de dimensão maior que o original), no qual as classes possivelmente seriam mais facilmente separáveis. Desta forma, ao implicitamente projetar o hiperplano de máxima margem no espaço de características, as SVMs são capazes de construir uma fronteira de separação não-linear no espaço original dos dados.

2.3.2 CNN

As redes neurais convolucionais são estruturas caracterizadas pelo emprego de camadas que aplicam a

operação de convolução para processar os dados, sendo que os parâmetros dos *kernels* de convolução são ajustados durante o treinamento. Estas arquiteturas vêm sendo amplamente utilizadas em tarefas de visão computacional, como classificação, detecção de objetos e segmentação de imagens [4].

Embora as CNNs tenham maior aplicação sobre dados bi ou tridimensionais (e.g., imagens), elas também podem ser utilizadas para processar dados unidimensionais, como acontece no caso de um sistema BCI-SSVEP.

A primeira arquitetura de CNN-1D utilizada neste trabalho foi baseada na proposta de [7]. A Figura 1 ilustra a estrutura desta rede convolucional.

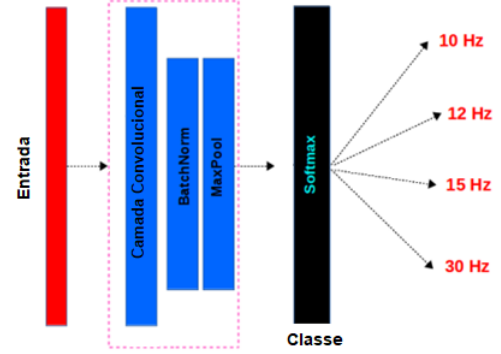


Figura 1: Estrutura da rede CNN-1D adaptada de [7].

Esta rede é formada por uma camada convolucional 1D com função de ativação *ReLU*, seguida pelas camadas de *batch normalization* e *max-pooling*, e, por fim, a camada de saída com função de ativação *Softmax*, que classificará os sinais recebidos (no cenário tratado, há quatro classes (frequências) distintas). Além disso, para testar a eficiência de tal metodologia, nenhum pré-processamento foi aplicado aos dados de entrada, isto é, a rede foi aplicada diretamente sobre os dados temporais.

A segunda abordagem explorada neste trabalho também é uma CNN-1D, mas cuja entrada corresponde ao conjunto de PSDs dos diferentes eletrodos, conforme estudado em [8], estimadas pela técnica do periodograma de Welch, e limitadas ao intervalo de 8 a 50 Hz. A Figura 2 mostra a arquitetura empregada. Os parâmetros da camada convolucional indicados na figura fazem com que, nessa camada, ocorra em paralelo a combinação dos valores das PSD dos diferentes eletrodos em cada frequência [8].

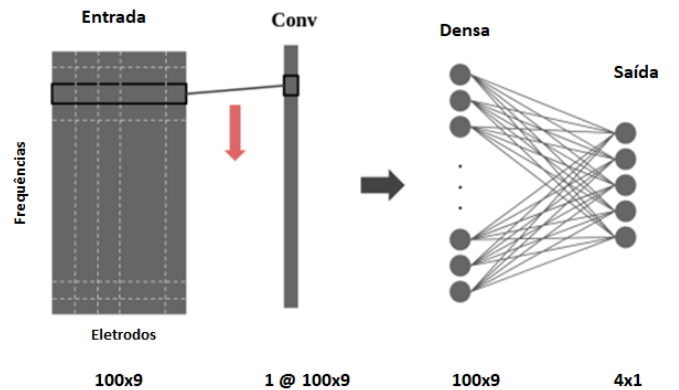


Figura 2: Estrutura da rede CNN-1D baseada em [8]. A entrada corresponde ao conjunto de PSDs dos vários eletrodos.

3 Metodologia

Para a realização das atividades propostas foi utilizado o banco de dados público do Tsinghua BCI Lab [16], o qual contém os sinais de EEG obtidos de 35 voluntários saudáveis, de ambos os gêneros, tendo ou não experiência no uso de sistemas BCI-SSVEP, usando 64 eletrodos e uma taxa de amostragem de 250 Hz. Ao todo há 40 frequências de estímulo no intervalo de 8 a 15,8 Hz e 6 *trials* (ou sessões) para cada frequência. Cada sessão tem duração de 6 s.

No presente trabalho, foram selecionados apenas nove eletrodos localizados nas regiões parietal e occipital do crânio, a saber: P1, Pz, P2, PO3, POz, PO4, O1, Oz e O2 [12]. Ademais, para respeitar o critério do número de eletrodos imposto pela CCA, foram escolhidas quatro frequências com um bom espaçamento, sendo elas, 8, 10, 12 e 15 Hz [12]. Em relação ao número de voluntários, foram utilizados os dados de oito deles de modo a atender igualmente os critérios de idade, gênero e experiência. A Tabela 1 mostra as informações sobre cada indivíduo selecionado. Além disso, para aplicar as técnicas anteriormente exploradas utilizamos um janelamento com duração de 2 s e sobreposição de 1 s, removendo antes os trechos inicial e final de 0,5 s de cada *trial*, uma vez que neles não há estimulação visual.

Tabela 1: Seleção de 8/35 indivíduos do banco de dados do Tsinghua BCI Lab [16].

Índice	Suj.	Genêro	Idade	Experiência
1	S02	Masculino	29	Sim
2	S07	Feminino	28	Sim
3	S23	Masculino	34	Não
4	S09	Feminino	31	Não
5	S03	Masculino	26	Sim
6	S08	Feminino	23	Sim
7	S24	Masculino	17	Não
8	S13	Feminino	17	Não

É importante mencionar, em especial para a etapa de classificação, que os dados, devidamente janelados, foram divididos em dois conjuntos: 83% (5 *trials*) para treinamento e validação, e 17% (1 *trial*) para teste, tomando o cuidado de colocar as janelas de um mesmo *trial* sempre no mesmo conjunto. Esta divisão foi feita uma única vez e todos os modelos consideraram os mesmos conjuntos de treinamento/validação e teste.

Para a obtenção dos hiperparâmetros dos diferentes modelos de classificação, realizamos uma busca em grade baseada em uma espécie de validação cruzada com 5-*folds*, tal que cada *trial* dentre os 5 reservados para essa etapa foi empregado uma vez como conjunto de validação. Ao final, a combinação de valores dos hiperparâmetros que levou o modelo ao melhor desempenho em termos da acurácia média de validação foi a escolhida. Além disso, no caso das CNNs, a etapa de treinamento já com os melhores parâmetros foi realizada 5 vezes para cada indivíduo, a fim de amenizar o impacto da inicialização dos pesos sobre o desempenho de teste aferido. Sendo assim, foi registrado o desempenho médio para esses modelos.

Dessa forma, foram construídos quatro modelos clássicos

de BCI e dois modelos baseados em DL. Em relação aos modelos clássicos, o primeiro corresponde a uma BCI-CCA, em que a etapa de extração de características é feita pela CCA e a decisão é pela frequência cujo coeficiente de correlação (em módulo) é o máximo, conforme proposto em [11]. Os modelos restantes utilizam o filtro CAR para o pré-processamento e uma SVM para a classificação, diferindo apenas na técnica adotada na etapa de extração de características. Assim, o segundo modelo (BCI-FFT-SVM) utiliza como vetor de características o conjunto de amostras espectrais obtidas pela DFT completa para cada eletrodo. Por sua vez, o terceiro modelo (BCI-FFT*-SVM) computa a DFT somente nas frequências de estímulo. Por fim, o último modelo (BCI-CCA-SVM) se baseia na CCA, extraindo os coeficientes de correlação canônica dos sinais de EEG.

No caso das BCIs baseadas em CNNs, foram implementados os dois modelos descritos na Seção 2.3.2, aqui denotados por BCI-CNN e BCI-Welch-CNN, os quais recebem como entrada os dados no domínio do tempo e no domínio da frequência, respectivamente. Em ambos os casos, o treinamento foi feito com o algoritmo Adam [4] considerando como função de perda a entropia cruzada categórica. É pertinente mencionar que para a BCI-Welch-CNN, o tamanho do *kernel* e o *stride* de convolução permanecem fixos e iguais a um, conforme proposto por [8], e o filtro CAR foi utilizado como etapa de pré-processamento.

4 Resultados e Discussão

As Tabelas 2 e 3 apresentam os desempenhos obtidos por cada modelo de BCI em termos da acurácia e F_1 -score, respectivamente, para cada um dos indivíduos selecionados da base de dados.

De maneira geral, nos experimentos realizados com modelos clássicos de BCI, as diferentes versões da BCI-SSVEP se mostraram satisfatórias, com destaque para os modelos BCI-CCA e BCI-FFT* (DFT aplicada apenas nas frequências de estímulo), que atingiram acurácias médias de 90,63% e 86,72%, respectivamente, conforme mostra a Tabela 2. Em relação aos modelos profundos, a BCI-CNN teve um desempenho superior ao da BCI-Welch-CNN, com acurácias médias de 88,44% e 79,22%, respectivamente. Conclusões parecidas podem também ser extraídas da Tabela 3 com relação ao F_1 -score.

Considerando os resultados das BCIs clássicas separados por voluntário, é possível constatar que usar os atributos em frequências extraídos pela DFT apenas nas frequências de estímulo levou a um desempenho melhor que explorar a DFT completa na maior parte dos casos. Além disso, a CCA combinada com SVM obteve sempre um desempenho pior que a BCI-CCA, o que ressalta ainda mais a relevância deste último modelo dada a sua maior simplicidade, especialmente no estágio de classificação.

Ademais, nota-se que os modelos BCI-FFT completa e BCI-Welch-CNN tiveram resultados bastante semelhantes, assim como os modelos BCI-FFT* (aplicada somente nas frequências de estímulo) e BCI-CNN. O principal destaque, porém, fica para o modelo BCI-CCA, que atingiu o melhor desempenho médio mesmo sendo uma

abordagem relativamente leve. Dessa forma, considerando a complexidade do modelo, seu desempenho e o total controle sobre todas as etapas do processamento de uma BCI-SSVEP, verificamos que os modelos clássicos são competitivos com opções inspiradas em DL, podendo até superar abordagens que aprendem a extrair os atributos dos dados, como a BCI-CNN.

Por fim, não é possível perceber uma tendência consistente de variação do desempenho das BCIs em função dos critérios de gênero, idade e experiência prévia dos voluntários.

Tabela 2: Desempenho dos modelos de BCI com métrica de acurácia.

Suj.	CCA	SVM FFT	SVM FFT*	SVM CCA	CNN	CNN Welch
1	87,50	75,00	81,25	43,75	90,00	51,25
2	87,50	75,00	87,50	75,00	92,50	77,50
3	68,75	68,75	75,00	81,25	72,50	57,50
4	93,75	75,00	68,75	75,00	90,00	92,50
5	93,75	100,0	100,0	75,00	100,0	98,75
6	100,0	93,75	87,50	75,00	85,00	83,75
7	93,75	87,50	93,75	93,75	98,75	98,75
8	100,0	62,50	100,0	75,00	78,75	73,75
Média (%)	90,63	79,69	86,72	74,22	88,44	79,22

* FFT analisada apenas nas frequências de estímulo.

Tabela 3: Desempenho dos modelos de BCI com métrica de F_1 -score.

Suj.	CCA	SVM FFT	SVM FFT*	SVM CCA	CNN	CNN Welch
1	87,86	75,32	81,15	37,25	89,42	48,62
2	86,11	65,40	86,11	71,53	92,25	77,25
3	66,28	68,75	75,60	81,15	70,58	55,49
4	93,65	66,67	64,84	71,93	89,62	91,90
5	93,65	100,0	100,0	74,76	100,0	98,73
6	100,0	93,65	86,67	73,74	84,56	82,86
7	93,65	86,67	93,65	93,65	98,73	98,73
8	100,0	60,00	100,0	73,10	76,73	72,57
Média (%)	90,15	77,06	86,00	72,14	87,74	78,27

* FFT analisada apenas nas frequências de estímulo.

5 Conclusão

Este trabalho de iniciação científica mostrou através de diferentes modelos de BCIs baseados no paradigma SSVEP que o emprego de técnicas consagradas de processamento de sinais podem levar a desempenhos tão bons, se não melhores, que as interfaces baseadas em modelos profundos. A principal vantagem de tais modelos encontra-se na relativa simplicidade, rapidez e controlabilidade, e no fato de explorarem o conhecimento que se tem das características do problema. Dentre as opções testadas, a BCI-CCA foi a que atingiu o melhor desempenho médio; a BCI-CNN, que parte das janelas de EEG no domínio do tempo, alcançou resultados bastante satisfatórios, mas um pouco inferiores aos da BCI-CCA. Por fim, a maior simplicidade da BCI-CCA ressalta ainda mais sua adequação ao problema.

Referências

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer

interfaces for communication and control,” *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

- [2] S. N. d. C. Leite, “Contribuições ao desenvolvimento de interfaces cérebro-computador baseadas em potenciais evocados visualmente em regime estacionário,” Ph.D. dissertation, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2016.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [4] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*, 2nd ed. O’Reilly Media, 2019.
- [5] X. Zhang, L. Yao, X. Wang, J. J. Monaghan, D. Mcalpine, and Y. Zhang, “A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers,” *Journal of Neural Engineering*, 2020.
- [6] W. Rampazzo, “Aprendizado profundo e interfaces cérebro-computador,” Master’s thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2019.
- [7] N. K. N. Aznan, S. Bonner, J. Connolly, N. Al Moubayed, and T. Breckon, “On the classification of ssvep-based dry-ecg signals via convolutional neural networks,” in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 3726–3731.
- [8] R. Pereira, “An analysis of the use of convolutional neural network for classification of steady-state visually evoked potencial-based bcis,” Master’s thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2021.
- [9] W. K. Härdle and L. Simar, *Applied multivariate statistical analysis*. Springer Nature, 2019.
- [10] S. N. d. Carvalho, G. V. Vargas, T. B. da Silva Costa, H. M. de Arruda Leite, L. Coradine, L. Boccato, D. C. Soriano, and R. Attux, “Space-time filter for ssvep brain-computer interface based on the minimum variance distortionless response,” *Medical & Biological Engineering & Computing*, vol. 59, no. 5, pp. 1133–1150, 2021.
- [11] Z. Lin, C. Zhang, W. Wu, and X. Gao, “Frequency recognition based on canonical correlation analysis for ssvep-based bcis,” *IEEE transactions on biomedical engineering*, vol. 54, no. 6, pp. 1172–1176, 2007.
- [12] G. Vettorazzi, “Filtragem espaço-temporal baseada no princípio mvdr aplicada a interfaces cérebro-computador sob o paradigma ssvep,” Master’s thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2021.
- [13] A. V. Oppenheim, J. R. Buck, and R. W. Schaffer, *Discrete-time signal processing*. Upper Saddle River, NJ: Prentice Hall, 1999.
- [14] P. Welch, “The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms,” *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
- [15] C. M. Bishop, *Pattern Recognition and Machine Learning*. springer, 2006.
- [16] Y. Wang, X. Chen, X. Gao, and S. Gao, “A benchmark dataset for SSVEP-based brain-computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1746–1752, 2017.