

Processamento de Sinais Cerebrais para Determinação do Foco de Atenção Visual

Leonardo Verçosa de Oliveira Valle, Sarah Negreiros de Carvalho Leite

Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, São José dos Campos — Brasil

Resumo— A implementação de Interfaces Cérebro-Computador (BCIs) baseadas em Potenciais Evocados Visualmente de Estado Estacionário (SSVEP) em cenários operacionais depende de metodologias de processamento de sinais que aliem acurácia e baixo custo computacional. Este artigo investiga a viabilidade de um pipeline composto pela referência pela média comum, extração de características no domínio da frequência através da Transformada Rápida de Fourier e classificação por meio de uma rede neural *Multilayer Perceptron*. O desempenho do modelo foi avaliado em um paradigma de classificação binária (estímulos visuais em 8 Hz e 10 Hz), comparando estratégias de treinamento individualizadas com uma abordagem intersujeitos (generalista). Os resultados indicam que o modelo generalista, ao ser treinado com um volume de dados maior proveniente de múltiplos usuários, consegue um desempenho equivalente ao do modelo individualizado, atingindo cerca de 88% de acurácia média entre os 35 indivíduos avaliados. Conclui-se que arquiteturas de redes neurais simples possuem um notável potencial de generalização para sistemas BCI-SSVEP, o que reforça sua aplicabilidade em contextos de alta demanda, como o controle de veículos não tripulados e o monitoramento cognitivo.

Palavras-Chave— Interface Cérebro-Computador, Eletroencefalografia, Redes Neurais Artificiais

I. INTRODUÇÃO

A evolução da interface humano-máquina é importante para a superioridade operacional em defesa e segurança. Nesse cenário, as Interfaces Cérebro-Computador (BCI) representam uma alternativa, estabelecendo canais de comando e controle rápidos e sigilosos. Dentre as tecnologias BCI, os sistemas baseados em Potenciais Evocados Visualmente em Regime Estacionário (SSVEP) são particularmente promissores devido à sua alta taxa de transferência de informação (ITR) e robustez. A abordagem SSVEP emprega a resposta síncrona do córtex visual a estímulos cintilantes, permitindo que a simples fixação do olhar em um alvo piscando em uma frequência específica seja traduzida em um comando digital. Essa capacidade abre portas para o controle de enxames de drones, a navegação em menus táticos e a operação de sistemas remotos de forma intuitiva e sem o uso das mãos (hands-free) [1], [2], [3], [4].

A Figura 1 apresenta o fluxo de processamento de um sistema BCI-SSVEP. Tipicamente, ele compreende a aquisição, amplificação e digitalização de sinais de eletroencefalografia (EEG) (1-3), seguido por etapas de pré-processamento (4) para otimizar a razão sinal-ruído (SNR), extração de características (5) para isolar as assinaturas neurais de interesse, e classificação (6) para decodificar a intenção do usuário em um comando de controle (7). O loop é fechado com o feedback ao operador (8) [3].

Leonardo Verçosa de Oliveira Valle, leonardo.valle.102087@ga.ita.br; Sarah Negreiros de Carvalho Leite, sarahnc@ita.br.

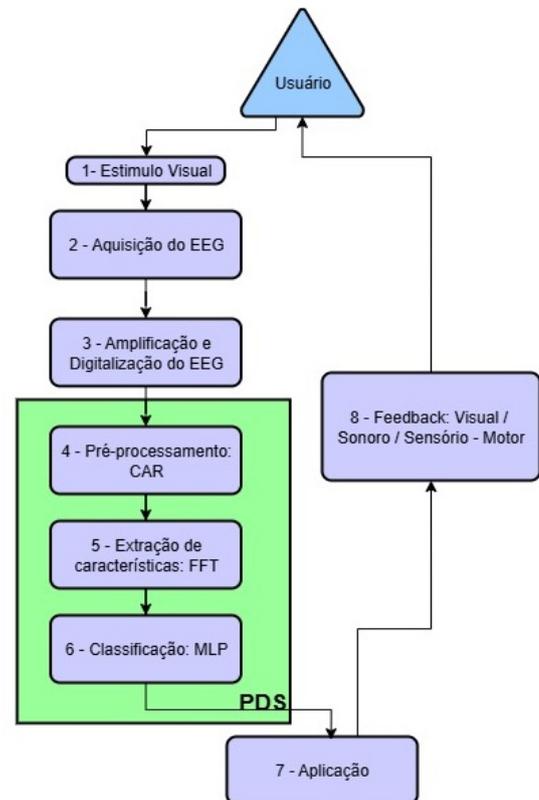


Fig. 1: Diagrama das etapas de um sistema BCI-baseado em SSVEP.

Apesar do potencial, a transição de protótipos de BCI para sistemas militarizados (*military-grade*) enfrenta diversas barreiras. Dentre as quais, se pode indicar a variabilidade neurofisiológica, ou seja, os padrões de EEG variam entre operadores (variabilidade intersujeitos) e até mesmo para o mesmo operador em diferentes momentos, dependendo do estresse ou fadiga (variabilidade intrasujeito). Modelos de classificação tradicionais frequentemente exigem sessões de calibração demoradas e específicas para cada indivíduo, um requisito impraticável em ambientes operacionais, onde o tempo é crítico e os sistemas devem ser interoperáveis entre diferentes membros da equipe [5].

A superação deste obstáculo é fundamental para viabilizar aplicações, como a emissão de comandos silenciosos em missões de reconhecimento, a interação com sistemas de informação em *cockpits* de aeronaves de alta performance ou o monitoramento passivo da carga cognitiva e da atenção de controladores de tráfego aéreo e operadores de radar [6]. Um sistema BCI eficaz para defesa deve ser generalista, ou seja, capaz de operar com alta precisão para qualquer operador com

um mínimo de ajustes.

Diante desse desafio, este trabalho investiga a hipótese de que uma arquitetura de aprendizado de máquina computacionalmente leve, como a rede *MultiLayer Perceptron* (MLP) rasa, pode alcançar um grau de generalização satisfatório quando alimentada por características espectrais bem definidas, extraídas pela Transformada Rápida de Fourier (FFT). Propõe-se avaliar se tal abordagem é capaz de oferecer um desempenho adequado, com baixo custo computacional e com capacidade de operar de forma consistente entre diferentes indivíduos, alinhando-se aos requisitos de sistemas de missão crítica.

II. BCI-SSVEP

O SSVEP é uma resposta cerebral ressonante, gerada no córtex visual quando o indivíduo foca sua atenção em um estímulo visual que se repete a uma frequência constante, geralmente entre 3.5 Hz e 75 Hz [7], [1]. Essa sincronização neural faz com que a frequência do estímulo e seus harmônicos se manifestem como picos de potência no espectro do sinal de EEG, especialmente nos canais da região occipital. A Figura 2 apresenta as regiões do cérebro, com destaque para o córtex visual.

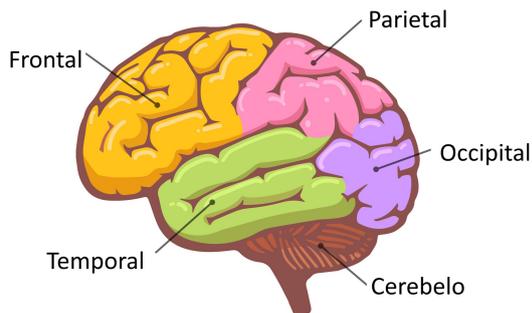


Fig. 2: Representação esquemática dos principais lobos do cérebro humano. É na região do lobo occipital onde os sinais SSVEP são primariamente gerados em resposta a estímulos visuais intermitentes.

A detecção dessa assinatura neural permite inferir com alta precisão qual estímulo o usuário está observando. Um sistema BCI-SSVEP típico utiliza uma interface com alvos visuais, cada um associado a um comando e modulado em uma frequência única.

A resposta SSVEP tem natureza involuntária, o que praticamente elimina a necessidade de treinamento do operador.

III. METODOLOGIA

Esta seção apresenta a metodologia empregada para a detecção do foco de atenção visual. Foi utilizado um conjunto de dados público com sinais de EEG de 35 indivíduos submetidos a estímulos visuais cintilantes [7]. O pipeline de processamento investigado consiste em três etapas sequenciais: (i) pré-processamento via Referência de Média Comum (CAR), (ii) extração de características espectrais (frequências fundamentais e harmônicas) por meio da FFT, e (iii) classificação utilizando uma rede MLP.

Para avaliar o desempenho do sistema BCI e sua capacidade de generalização para indivíduos não observados durante o

treinamento, foram projetados três cenários de teste. Cada cenário busca simular diferentes níveis de variabilidade e desafiar a robustez do modelo em condições realistas, especialmente quando os indivíduos não participaram da fase de treinamento.

- Cenário 1: Apenas os dados do próprio indivíduo foram utilizados para o treinamento e teste, sendo empregados 48 ensaios para o treinamento e 12 para o teste, por indivíduo. Neste cenário, foi utilizada a primeira arquitetura de MLP, composta por duas camadas ocultas com 64 e 32 neurônios e função de ativação \tanh .
- Cenário 2: A BCI foi treinada com os dados de 34 participantes e testada com os dados do 35º, adotando uma estratégia de validação cruzada do tipo *leave-one-subject-out* (LOSO), em que cada indivíduo foi utilizado como teste uma vez. Para este cenário, foi utilizada a segunda MLP, composta por três camadas ocultas com 150, 70 e 35 neurônios, respectivamente.
- Cenário 3: A BCI foi treinada com os dados de 29 participantes, excluindo seis indivíduos (7, 11, 17, 18, 29, 33) que apresentaram acurácia inferior a 70% no Cenário 1, conforme metodologia proposta por [8]. Assim como no Cenário 2, a avaliação entre os indivíduos foi realizada utilizando a estratégia LOSO, com a mesma MLP.

A. Base de Dados

O conjunto de dados de EEG utilizado neste estudo foi fornecido pelo Laboratório de BCI da Universidade Tsinghua [7]. Ele é composto por sinais de EEG multicanais (64 canais) de 35 participantes saudáveis — incluindo 8 indivíduos experientes e 27 indivíduos inexperientes. Durante a coleta dos sinais cerebrais, os participantes deveriam focar em um caractere de um teclado virtual, sendo 40 estímulos visuais codificados por modulação conjunta de frequência e fase, variando de 8 Hz a 15,8 Hz com incrementos de 0,2 Hz. Cada coleta possui 1500 amostras, correspondendo a 6 segundos de EEG amostrados a 250 Hz, dos quais, 5 s são sinais durante a efetiva exposição aos estímulos.

B. Pré-processamento

Para o pré-processamento do sinal cerebral, aplicou-se um filtro passa-bandas na faixa de 0,15 Hz a 200 Hz, a fim de limitar a banda de frequência e atender ao critério de Nyquist [9]. Em seguida, aplicou-se o filtro CAR, que atenuou artefatos comuns entre os canais e melhorou a razão sinal-ruído [10].

O filtro CAR, aplicado ao canal i , denotado por x_i^{CAR} , pode ser expresso como:

$$x_i^{CAR} = x_i - \frac{1}{N} \sum_{j=1}^N x_j \quad (1)$$

onde N representa o número total de canais de EEG ($N = 64$) e x são as amostras do sinal, num dado instante de tempo, para cada canal.

O sinal foi janelado em segmentos de 1 s, com 250 amostras não sobrepostas extraídas da porção central do sinal, sendo descartados os 500 ms iniciais e finais para evitar efeitos transitórios.

com 64 e 32 neurônios, respectivamente, ambas utilizando a função de ativação \tanh , que introduz não-linearidade e contribui para a modelagem de padrões complexos nos dados. A camada de saída empregou a função de ativação softmax . O modelo foi treinado por 100 épocas com tamanho de lote igual a 8, utilizando o otimizador Adam, que combina *momentum* e adaptação de taxa de aprendizado para acelerar a convergência. A taxa de aprendizado foi fixada em 10^{-4} , buscando maior estabilidade no processo de otimização.

A segunda MLP apresenta três camadas ocultas totalmente conectadas, com 150, 70 e 35 neurônios, respectivamente, também empregando a função de ativação \tanh em todas as camadas ocultas. A saída utiliza novamente a função softmax . Este modelo foi treinado por 100 épocas, com tamanho de lote igual a 64, utilizando o mesmo otimizador Adam. A taxa de aprendizado, neste caso, foi ajustada para 0,0005, buscando um equilíbrio entre estabilidade e velocidade na convergência.

Para quantificar a eficiência da abordagem proposta, foram mensurados o tempo de treinamento e o tempo de inferência dos modelos MLP. O tempo de treinamento foi registrado para cada um dos três cenários, enquanto o tempo de inferência foi calculado como a média da predição para um único ensaio. Todos os testes foram executados em um computador com as seguintes especificações: Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz (1.99 GHz), 20,0 GB de RAM, com os modelos operando exclusivamente em CPU.

IV. RESULTADOS

A Tabela I apresenta a acurácia do BCI-SSVEP de cada indivíduo nos três cenários de avaliação. As acurácias médias dos modelos nos três cenários foram muito semelhantes, o teste de normalidade de Shapiro–Wilk indicou que os dados de todos os cenários não seguem distribuição normal ($p < 0,001$). Portanto, foi utilizado o teste de Friedman, não paramétrico para medidas repetidas, que não revelou diferenças estatisticamente significativas entre os cenários ($p = 0,804$). As comparações post-hoc pareadas com o teste de Wilcoxon e correção de Bonferroni também não mostraram diferenças significativas entre quaisquer pares de cenários. Estes resultados indicam que, apesar da variabilidade individual, os três cenários apresentaram desempenhos equivalentes em termos de acurácia.

No Cenário 1, em que houve o treinamento da MLP de forma personalizada, a acurácia média obtida pelos 35 indivíduos foi de 0.88 ± 0.15 , havendo uma grande variabilidade de desempenho intersujeitos. Dos 35 indivíduos, 13 alcançaram uma acurácia de 100% na identificação dos dados de validação, ao passo que o modelo de dois deles (Indivíduos 18 e 33) não foi capaz de aprender a distinguir os padrões cerebrais em cada estimulação visual, apresentando um desempenho aleatório. Isso evidencia que as poucas amostras disponíveis para o treinamento customizado tornaram o modelo pouco generalizável para alguns indivíduos, com tendência ao *overfitting*. Isso se deve, principalmente, à limitação da quantidade de dados disponíveis por usuário, o que restringe a capacidade do modelo em aprender padrões robustos.

No Cenário 2, que empregou a estratégia LOSO usando os dados de todos os indivíduos da base para treinar o modelo,

TABELA I: Acurácia dos modelos avaliados nos 3 cenários para cada indivíduo. Os indivíduos 7, 11, 17, 18, 29 e 33 foram os que apresentaram menor desempenho no Cenário 1 e, portanto, não entraram no treinamento do modelo do Cenário 3.

Indivíduo	Cenário 1 (%)	Cenário 2 (%)	Cenário 3 (%)
1	100.00	93.33	93.33
2	91.67	85.00	86.67
3	75.00	90.00	91.67
4	100.00	83.33	88.33
5	91.67	91.67	93.33
6	100.00	93.33	93.33
7	66.67	86.67	86.67
8	100.00	96.67	96.67
9	83.33	93.33	95.00
10	83.33	90.00	86.67
11	66.67	75.00	75.00
12	83.33	95.00	91.67
13	100.00	71.67	66.67
14	100.00	83.33	88.33
15	100.00	88.33	93.33
16	91.67	86.67	90.00
17	66.67	88.33	90.00
18	50.00	80.00	81.67
19	91.67	91.67	86.67
20	100.00	93.33	91.67
21	100.00	90.00	91.67
22	100.00	93.33	98.33
23	91.67	88.33	88.33
24	91.67	96.67	93.33
25	100.00	88.33	88.33
26	100.00	95.00	91.67
27	100.00	93.33	95.00
28	91.67	96.67	98.33
29	58.33	68.33	68.33
30	83.33	90.00	86.67
31	91.67	98.33	96.67
32	100.00	95.00	96.67
33	50.00	66.67	61.67
34	91.67	96.67	96.67
35	91.67	95.00	95.00
Média	88.10	88.31	88.95
Desvio Padrão	14.89	8.03	8.75

a acurácia média observada foi de 0.89 ± 0.08 , com um comportamento mais homogêneo entre os indivíduos. Nenhum dos indivíduos alcançou acurácia de 100%, entretanto somente dois voluntários ficaram com desempenho inferior a 70%. Nota-se que o modelo conseguiu generalizar intersujeitos, sendo capaz de identificar padrões no sinal EEG, mesmo sem exposição prévia a sinais do indivíduo que está sendo avaliado.

O Cenário 3 apresenta um desempenho muito próximo ao observado no Cenário 2, com acurácia média de 89% e desvio padrão de 8.8%. A princípio, a curadoria dos dados, com a exclusão de sinais de EEG de indivíduos com menor desempenho no Cenário 1, não parece impactar na estabilidade e capacidade de generalização de modelos de BCI-SSVEP.

É interessante notar que, indivíduos com baixo desempenho no Cenário 1, tendem a atingir melhores acurácias nos Cenários 2 e 3.

A Tabela II apresenta uma comparação entre as métricas de classificação entre os três cenários.

TABELA II: Comparação das métricas de classificação para os três cenários.

Métrica	Cenário 1	Cenário 2	Cenário 3
Acurácia	88.1%	88.8%	88.9%
Precisão	89.6%	88.6%	88.7%
Recall	86.2%	89.0%	89.1%
F1 Score	87.9%	88.8%	88.9%

Em todos os três cenários avaliados, observa-se um desempenho equilibrado do modelo na classificação das duas classes de estímulo (8 Hz e 10 Hz), indicando ausência de viés e uma capacidade consistente de discriminação entre as frequências.

A análise de desempenho apresentada na Tabela III indica a viabilidade da rede MLP para aplicações críticas. O tempo médio de inferência foi baixo (cerca de 70 ms por ensaio), adequado para interação fluida em sistemas BCI. O modelo também apresentou treinamento rápido (< 15 s por *fold*) em CPU convencional, sem necessidade de GPUs, evidenciando sua praticidade e eficiência. Esses resultados indicam que o MLP é uma solução de alto desempenho e baixo custo, ideal para sistemas embarcados.

TABELA III: Análise de desempenho computacional dos modelos MLP nos três cenários.

Métrica Computacional	Cenário 1	Cenário 2	Cenário 3
Tempo Médio de Treinamento (s)	10.541	14.757	14.301
Tempo Médio de Inferência (ms)	72.768	67.989	70.699

V. CONCLUSÃO

Este trabalho apresentou uma abordagem baseada em redes neurais para a classificação de sinais EEG do tipo SSVEP, com o objetivo de desenvolver um sistema BCI eficaz mesmo na ausência de treinamento com dados do próprio indivíduo. O pipeline proposto combinou técnicas de pré-processamento (filtro passa-bandas e CAR), extração espectral nas frequências fundamentais e suas harmônicas via FFT e um classificador do tipo MLP, sendo avaliado em três cenários distintos de validação.

No Cenário 1 foram gerados modelos personalizados para cada indivíduo. Neste cenário, a acurácia média foi de 88,10%, com variação significativa entre participantes, e apresentou um equilíbrio na identificação dos estímulos de 8 e 10 Hz. Nos Cenários 2 e 3, com validação cruzada do tipo LOSO, o modelo demonstrou capacidade de generalização intersujeitos, atingindo uma acurácia média de cerca de 89% e apresentando menor variabilidade.

Os resultados indicam que a abordagem proposta é eficaz tanto em cenários personalizados quanto generalistas, sendo capaz de manter desempenho elevado mesmo sem calibragem individual. Isso torna o sistema particularmente promissor para aplicações em contextos operacionais críticos, como os do domínio da defesa e segurança, onde a interação entre humanos e sistemas complexos exige respostas rápidas, precisas e confiáveis. Em especial, a possibilidade de comandos silenciosos, monitoramento do estado cognitivo e controle *hands-free* em ambientes restritivos ou de alto risco.

VI. DISPONIBILIDADE DE DADOS E CÓDIGOS

O conjunto de dados original utilizado neste estudo está publicamente disponível [7]. Os códigos-fonte desenvolvidos neste trabalho para o pré-processamento dos si-

nais, extração de características e treinamento dos modelos de classificação foram disponibilizados em um repositório público no GitHub. O acesso pode ser feito através do seguinte link: <https://github.com/leo-valle/EEG-SSVEP-with-Neural-Networks-MLP>

REFERÊNCIAS

- [1] S. Liu, D. Zhang, Z. Liu, M. Liu, Z. Ming, T. Liu, D. Suo, S. Funahashi, and T. Yan, "Review of brain-computer interface based on steady-state visual evoked potential," *Brain Science Advances*, vol. 8, no. 4, pp. 258–275, 2022. [Online]. Available: <https://www.sciopen.com/article/10.26599/BSA.2022.9050022>
- [2] Y. Sun, L. Liang, Y. Li, X. Chen, and X. Gao, "Dual-alpha: a large eeg study for dual-frequency ssvep brain-computer interface," *GigaScience*, vol. 13, 01 2024.
- [3] D. Xu, F. Tang, Y. Li, Q. Zhang, and X. Feng, "An analysis of deep learning models in ssvep-based bci: A survey," *Brain Sciences*, vol. 13, no. 3, 2023. [Online]. Available: <https://www.mdpi.com/2076-3425/13/3/483>
- [4] K. M. Hossain, M. A. Islam, S. Hossain, A. Nijholt, and M. A. R. Ahad, "Status of deep learning for EEG-based brain-computer interface applications," *Frontiers in Computational Neuroscience*, vol. 16, 2023. [Online]. Available: <https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2022.1006763>
- [5] A. Binnendijk, T. Marler, and E. Bartels, *Brain-Computer Interfaces: U.S. Military Applications and Implications, An Initial Assessment*. RAND Corporation, jan 2020.
- [6] R. Ortner, C. Guger, R. Prueckl, E. Grünbacher, and G. Edlinger, *SSVEP Based Brain-Computer Interface for Robot Control*, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
- [7] 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.
- [8] S. Saha and M. Baumert, "Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: A review," *Frontiers in Computational Neuroscience*, vol. 13, p. 87, dec 2019.
- [9] R. Arie, A. Brand, and S. Engelberg, "Compressive sensing and subnyquist sampling," *IEEE Instrumentation Measurement Magazine*, vol. 23, no. 2, pp. 94–101, 2020.
- [10] B. Phanikrishna, P. Pławiak, and A. Jayaprakash, "A brief review on EEG signal pre-processing techniques for real-time brain-computer interface applications," sep 2021.
- [11] J. Xie, "A novel filter bank and fourier transform convolutional neural network for ssvep classification," pp. 9–16, 2024.
- [12] L. S. Snyder, M. Karimzadeh, C. Stober, and D. S. Ebert, "Situational awareness enhanced through social media analytics: A survey of first responders," in *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, 2019, pp. 1–8.
- [13] J. Thomas, T. Maszczyk, N. Sinha, T. Kluge, and J. Dauwels, "Deep learning-based classification for brain-computer interfaces," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 234–239.
- [14] S. Aggarwal and N. Chugh, "Review of machine learning techniques for EEG based brain computer interface," *Archives of Computational Methods in Engineering*, vol. 29, p. 1, jan 2022.