REPRODUÇÃO E VALIDAÇÃO DA ARQUITETURA EMPT EM *BENCHMARKS* DA MOABB

SANTIAGO D. V. A. MARTINEZ[1]; MARILTON S. DE AGUIAR[2]

[1]*Universidade Federal de Pelotas (UFPel) – santiago.dvam@inf.ufpel.edu.br*
[2]*Universidade Federal de Pelotas (UFPel) – marilton@inf.ufpel.edu.br*

1. INTRODUÇÃO

O cérebro constitui uma estrutura vasta e complexa, mas um de seus princípios de funcionamento já está bem estabelecido: a comunicação neuronal ocorre por meio de sinais elétricos. Esse conhecimento impulsionou a criação de máquinas capazes de decodificar tais sinais e convertê-los em comandos, as chamadas interfaces cérebro-computador (BCI) (Wolpaw et al., 2002).

O desenvolvimento de BCIs especializadas depende de diversos *datasets* criados para capturar sinais neurais, entre eles os associados à imaginação motora (MI), ligada ao movimento físico. Esses dados auxiliam na recuperação de movimentos por meio do controle de próteses e exoesqueletos usados em treinamento terapêutico (Wolpaw et al., 2002), relevante no Brasil, onde mais de 14 milhões de pessoas têm alguma deficiência, a física sendo a mais prevalente (IBGE, 2025). Repositórios abertos, como a *Mother of All BCI Benchmarks* (MOABB), cumprem papel central ao permitir testar BCIs em um ambiente padronizado (Jayaram et al., 2018).

O Eletroencefalograma (EEG) figura como o principal método de captura desses sinais, por ser não invasivo e de fácil acesso (Qazzaz et al., 2016). Seus registros, porém, são ruidosos e frequentemente contaminados por artefatos de ações involuntárias, o que exige processamento avançado (Zhang et al., 2020). Nesse cenário, métodos de aprendizado de máquina despontam como ferramentas poderosas para o tratamento de sinais EEG. A arquitetura convolucional EEGNet manteve-se, por um período, competitiva em diversos *datasets* (Lawhern et al., 2018); mais recentemente, no entanto, *Transformers* esparsos passaram a superar seus *benchmarks* (Liu et al., 2024).

A arquitetura de *Transformers* esparsos EMPT constitui o objeto deste estudo. Desenvolvida por Liu et al. (2024), ela combina uma *Mixture of Experts* (MoE) com auto-atenção ProbSparse para conferir ao modelo especialização e esparsidade. O trabalho original, contudo, não disponibiliza código aberto, o que compromete sua reprodutibilidade e limita seu impacto científico (Zhou et al., 2023), além de recorrer à Transformada-S Modificada e ao *Common Spatial Pattern* (CSP) no pré-processamento, técnicas distintas das adotadas neste trabalho. Propõe-se, assim, reproduzir a arquitetura EMPT e avaliá-la frente a *datasets* da MOABB, adotando-se o janelamento via *Braindecode* para melhor integração com a MOABB. As contribuições deste artigo resumem-se, portanto, a: (i) a primeira reprodução em código aberto da arquitetura EMPT; e (ii) sua avaliação padronizada frente a *datasets* da MOABB.

2. METODOLOGIA

Alinhada aos padrões contemporâneos de BCIs, esta arquitetura organiza-se em três fases: aquisição de sinais, pré-processamento e extração de padrões. Os sinais provêm de *datasets* da MOABB voltados à MI de movimentos da mão

esquerda e direita, num total de quatro conjuntos selecionados: BNCI2014_1, o mais empregado na literatura; BNCI2014_4, outra referência consolidada na comunidade; PhysionetMI, um *dataset* extenso e ruidoso; e Cho2017, cujos sinais densos apresentam maior dificuldade de decodificação.

O janelamento da biblioteca *Braindecode* capta os dados diretamente da MOABB, incluindo o sinal, os canais gravados, a frequência de amostragem e os metadados associados. Para isolar as informações relevantes, filtram-se primeiro os canais associados a artefatos involuntários e, em seguida, aplica-se um filtro passa-bandas entre 8 e 30 Hz, restrito aos ritmos Mu e Beta, responsáveis pelos movimentos físicos.

Os dados são organizados em janelas temporais fixas, definidas pela duração de cada triagem de EEG, no formato FXCXT (F: épocas, C: canais, T: largura temporal). Assim como no EMPT original (Liu et al., 2024), selecionam-se apenas os sinais relativos a movimentos das mãos, reorganizados no formato FXTXC exigido pelo *Transformer*, no qual os segmentos temporais funcionam como tokens sequenciais e os canais determinam a dimensionalidade.

O sinal é então processado pelo *Transformer*, estruturado em três blocos. No primeiro, uma camada de autoatenção multicabeças alimenta uma camada MoE, capturando dependências temporais globais; o roteamento esparso do MoE, por sua vez, permite que os *Experts* se especializem em variações espectrais, estruturas temporais ou segmentos ruidosos.

No segundo bloco, a autoatenção multicabeças dá lugar à ProbSparse: as informações mais relevantes são identificadas a partir da diferença entre a correlação da amostra máxima e a correlação média, e o contexto resultante é propagado (Zhou et al., 2021) para captar relações temporais de alto nível entre segmentos e canais.

O *pipeline* conclui-se com um *global average pooling* ao longo do eixo temporal, condensando a sequência em um vetor fixo, seguido por duas camadas feedforward totalmente conectadas, que conferem maior estabilidade às barreiras de decisão sobre as representações já abstraídas.

O treinamento foi conduzido em regime supervisionado e intrassujeito, com cada sujeito treinado e avaliado de forma independente por meio de validação cruzada de 5 dobras (5-*fold*) e embaralhamento das amostras, reduzindo a variância estatística. Utilizou-se o otimizador Adam (taxa de aprendizado de 5×10^{-5}) com perda de entropia cruzada categórica esparsa, por até 300 épocas, com parada antecipada após 15 épocas sem melhoria na acurácia de validação, restaurando-se os pesos de melhor desempenho.

3. RESULTADOS E DISCUSSÃO

Para avaliar a competitividade do *pipeline* frente ao EMPT original e a outras arquiteturas, analisaram-se os sujeitos do dataset BNCI2014_1, tomando como valor de referência a Melhor Acurácia de Validação, protocolo comum na literatura, com K-*Fold* elevado para 10 e paciência do *early stopping* para 30 épocas, de modo a reduzir a variância da estimativa. Os resultados constam na Tabela 1.

A Tabela 1 mostra que o EMPT original supera este *pipeline* na maioria dos sujeitos; no sujeito S8, porém, os resultados ficam próximos, quando não superiores. Essa variabilidade interindividual se confirma na análise por sujeito: os melhores desempenhos ocorrem em S3, S8 e S9 (acima de 90%), enquanto S2 e S5 caem mais acentuadamente; sujeitos "fáceis" e "difíceis" tendem a coincidir

entre *pipelines*, com exceção de S7 (aqui inferior ao usual) e S6 (com leve melhora). A diferença geral frente ao EMPT decorre, em grande parte, do pré-processamento adotado, e os resultados superam ainda arquiteturas já consolidadas, como EEGNet e MI-DABAN.

Tabela 1 - Acurácia entre *pipelines* no *dataset* BNCI2014_1. Valores em porcentagem (%).

Sujeitos	S1	S2	S3	S4	S5	S6	S7	S8	S9	Média
Este *pipeline*	84.73	69.47	91.32	72.89	63.05	74.29	68.08	96.51	90.26	**79.01**
EMPT	93.72	90.03	96.72	93.54	92.61	90.84	95.51	94.11	93.42	**93.39**
EEGNet	83.68	63.89	90.97	64.24	59.72	52.08	87.85	82.29	86.81	**74.61**
MI- DABAN	88.54	55.56	91.32	77.43	60.42	58.68	87.15	83.68	82.64	**76.16**
CNN- LSTM	89.23	72.53	97.23	76.28	82.48	69.15	94.76	86.14	86.10	**82.84**

Para testar a capacidade de generalização, comparou-se a acurácia obtida em diferentes datasets com a de outras arquiteturas presentes na MOABB. Os resultados encontram-se na Tabela 2.

Tabela 2 - Acurácia de *pipelines* em diferentes *datasets* de MI presentes na MOABB. Dados em porcentagem (%), com desvio padrão representado (±).

Pipeline	BNCI2014_1	BNCI2014_4	PhysionetMI	Cho2017
Este *pipeline*	79.01±11.88	69.49±15.23	66.94±16.56	59.00±14.67
EEGNet_8_2	77.15±19.33	69.50±19.50	59.55±15.95	66.79±16.34
ShallowConvNet	86.17±13.74	72.36±18.05	65.19±15.80	73.84±14.95
DeepConvNet	82.07±15.52	72.36±18.53	59.57±16.77	71.67±12.91
MDM	81.69±14.94	77.66±15.78	54.76±16.79	63.39±13.69

Os resultados obtidos no PhysionetMI mostram-se promissores, evidenciando capacidade de extrair padrões relevantes mesmo em meio a sinais ruidosos. Já no Cho2017, o desempenho é inferior, o que aponta limitações do *pipeline* diante de sinais com baixa discriminabilidade. Os desvios padrão elevados na Tabela 2 refletem a variabilidade interindividual do EEG, de modo que a estabilidade entre sujeitos importa tanto quanto a acurácia média.

4. CONCLUSÕES

Este trabalho reproduziu a arquitetura EMPT em código aberto e a testou frente a *datasets* da MOABB. Embora não tenha alcançado os valores obtidos por Liu et al. (2024), superou *benchmarks* como EEGNet e MI-DABAN e demonstrou boa capacidade de generalização no *dataset* ruidoso PhysionetMI.

Este estudo apresenta, contudo, algumas limitações. O elevado custo computacional associado à atenção esparsa e ao roteamento entre *experts* pode dificultar a implementação em tempo real ou o treinamento em larga escala na

ausência de hardware especializado. Observou-se também que o *pipeline* enfrenta dificuldades para extrair características de sinais sem padrões de MI claros, como no dataset Cho2017, sugerindo que a atenção esparsa depende de um nível mínimo de qualidade do sinal para realizar um roteamento eficaz.

Trabalhos futuros podem se dedicar à busca por hiperparâmetros, por meio de *grid search* e otimização Bayesiana, e à avaliação de outros métodos de pré-processamento. Sugere-se ainda testar o *pipeline* em cenários de tempo real, como o controle de próteses e exoesqueletos, ou com *datasets* de pacientes em tratamento, como CHB-MIT e BONN, que reúnem indivíduos com crises epilépticas. Para facilitar a continuidade da pesquisa, o código utilizado neste trabalho está disponível em repositório aberto[1].

Em conclusão, este trabalho reforça a viabilidade de arquiteturas baseadas em *Mixture of Experts* para lidar com a variabilidade e o ruído inerentes aos sinais de EEG. Ao promover a transparência metodológica, esta pesquisa busca contribuir para tecnologias assistivas mais robustas e acessíveis, aproximando as interfaces cérebro-computador da aplicação prática para os mais de 14 milhões de brasileiros com deficiência.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AL-QAZZAZ, N.K. et al. EEG wavelet spectral analysis during a working memory task in stroke-related mild cognitive impairment patients. In: **INTERNATIONAL CONFERENCE FOR INNOVATION IN BIOMEDICAL ENGINEERING AND LIFE SCIENCES**, 2016. p.82-85.

ARISTIMUNHA, B. et al. **Mother of all BCI benchmarks**. 2025. Acessado em 04 mar. 2026. Online. Disponível em: https://github.com/NeuroTechX/moabb

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo 2022**: Brasil tem 14,4 milhões de pessoas com deficiência. 2025. Online. Disponível em: <https://www.ibge.gov.br/>

JAYARAM, V.; BARACHANT, A. Moabb: Trustworthy algorithm benchmarking for bcis. **Journal of Neural Engineering**, Bristol, v.15, n.6, p.066011, 2018.

LAWHERN, V.J. et al. EEGNet: A compact convolutional neural network for eeg-based brain–computer interfaces. **Journal of Neural Engineering**, v.15, n.5, p.056013, 2018.

LIU, M. et al. EMPT: A sparsity transformer for EEG-based motor imagery recognition. **Frontiers in Neuroscience**, v.18, p.1366294, 2024.

WOLPAW, J.R. et al. Brain–computer interfaces for communication and control. **Clinical Neurophysiology**, v.113, n.6, p.767-791, 2002.

ZHANG, X. et al. A survey on deep learning-based brain signal decoding and its applications to BCIs. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, v.28, n.8, p.1734-1748, 2020.

ZHOU, H. et al. Informer: Beyond efficient transformer for long sequence time-series forecasting. In: **AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE**, 2021. v.35, p.11106-11115.

ZHOU, S. et al. What is the impact of releasing code with publications? Statistics from the machine learning, robotics, and control communities. **arXiv preprint**, 2023. Acessado em 04 mar. 2026. Online. Disponível em: https://arxiv.org/abs/2308.10008

[1] https://github.com/capysanti/Reproducing-the-EMPT-Pipeline