

Estudo do Método de Aprendizado Multitarefa em Interfaces Cérebro-Computador Baseadas no Paradigma de Imagética Motora

Palavras-Chave: Interface Cérebro-Computador; Aprendizado de Máquina.

Autores:

Daniel Neri de Jesus, FEEC – UNICAMP

Prof. Denis Fantinato (orientador), FEEC - UNICAMP

INTRODUÇÃO:

As interfaces cérebro-máquina (BCI, do inglês *Brain-Computer Interface*) configuram-se como dispositivos capazes de estabelecer um canal de comunicação bidirecional entre o sistema nervoso humano e plataformas computacionais [2]. Tais sistemas apresentam enorme potencial em diversas áreas, incluindo tecnologias assistivas para pessoas com deficiências motoras, equipamentos de reabilitação neurológica e aplicações de entretenimento interativo. Em virtude dessa amplitude de aplicações, o desenvolvimento de BCIs tem atraído considerável atenção tanto da comunidade acadêmica quanto da indústria.

Um sistema típico de BCI pode ser decomposto em cinco etapas fundamentais. A primeira consiste na (i) *aquisição* dos sinais bioelétricos gerados pela atividade cerebral; em BCIs baseadas em eletroencefalografia (EEG), essa captura é feita por meio de eletrodos posicionados estrategicamente sobre o escalpo, em regiões de interesse neurológico. A segunda fase é o (ii) *pré-processamento*, durante a qual são aplicados filtros e algoritmos de remoção de artefatos, por exemplo, ruídos musculares e interferências eletromagnéticas, para melhorar a relação sinal-ruído. Na sequência, ocorre a (iii) *extração de características*, cujo objetivo é projetar os sinais brutos em um espaço de representação onde padrões relevantes fiquem mais facilmente distinguíveis. Em seguida, faz-se a (iv) *seleção de características*, processo que visa identificar e reter os elementos mais importantes, reduzindo dimensionalidade e melhorando a eficiência dos classificadores. Por fim, a quinta etapa corresponde à (v) *classificação* propriamente dita, na qual algoritmos de aprendizado de máquina atribuem rótulos aos sinais, determinando, por exemplo, a que tipo de comando certo sinal está associado.

Neste projeto, o foco foi sobre a etapa final, ou seja, o desenvolvimento e avaliação de métodos de classificação de sinais EEG. Em particular, foram investigadas técnicas de aprendizado de máquina de redes neurais profundas, com vistas a aperfeiçoar a acurácia e o potencial de generalização na classificação. Além disso, como diversos dispositivos podem ser considerados sistemas BCI, o enfoque foi em interfaces Cérebro-Máquina baseadas em eletroencefalogramas de superfície (EEG). Esta abordagem, além de bastante difundida, permite a aquisição de dados de maneira não invasiva e com equipamento relativamente simples.

Dentro do universo de BCIs baseadas em EEG, destacam-se sobretudo dois paradigmas de utilização [2]. O primeiro, denominado SSVEP (*Steady-State Visually Evoked Potential*), explora a sincronização da atividade elétrica do córtex visual em resposta a estímulos luminosos que piscam em frequências específicas. O segundo paradigma, conhecido como imagética motora, solicita ao usuário que imagine movimentos pré-definidos (por exemplo, movimentar mentalmente a mão direita, os pés ou a língua). Essa tarefa gera padrões característicos de ativação no córtex motor, os quais podem ser capturados, processados e classificados para inferir a intenção de movimento.

Apesar do caráter promissor dessas abordagens, ainda há importantes desafios a serem superados para que BCIs atinjam desempenho confiável em ambientes reais. A natureza não invasiva da EEG implica em sinais de amplitude muito baixa (tipicamente entre 5 e 20 μV)[5], o que os torna suscetíveis a ruídos tanto internos

(atividade cerebral não relacionada ao paradigma de interesse) quanto externos (interferências eletromagnéticas, artefatos de movimento) [5]. Assim, a qualidade do processamento de sinal e a eficiência dos algoritmos de classificação são cruciais para reduzir erros de interpretação e garantir respostas em tempo hábil, especialmente em aplicações que demandam interação em tempo real.

Diante desse contexto, o presente trabalho propõe um estudo aprofundado dos métodos de classificação de sinais EEG, focando no uso de arquiteturas de redes neurais artificiais e técnicas de *deep learning*. O objetivo é aprimorar a extração de características relevantes, otimizar a separação de classes e elevar a precisão na detecção das intenções do usuário, contribuindo para a disseminação de BCIs mais acessíveis e eficientes.

METODOLOGIA:

O projeto, que consistiu em uma revisão bibliográfica e em aplicações de códigos desenvolvidos em Python, teve como foco principal o estudo de redes convolucionais (CNNs, do inglês *Convolutional Neural Networks*) e de redes de aprendizado profundo (*Deep Learning*), combinadas com a abordagem multitarefa (MTL, do inglês *Multi-Task Learning*). A partir do aprofundamento destas estruturas, soluções com combinações das mesmas foram utilizadas para classificação de dados de EEG presentes em repositórios públicos, bem como comparativos com métodos tradicionais. Para o projeto, foi utilizado a estratégia chamada *cross-subject*, que se constitui de utilizar, para o treinamento da rede, dados de diferentes indivíduos, a fim de se otimizar a tarefa de classificação de um único indivíduo. Essa abordagem foi comparada com a abordagem tradicional, em que apenas um indivíduo fornece dados tanto para o treinamento, quanto para a classificação.

A seguir, apresenta-se uma breve descrição dos principais conceitos estudados:

-*Deep Learning*: No escopo deste projeto, redes profundas podem ser entendidas como aquelas que reúnem várias camadas além do que se considera convencional para uma rede neural tradicional. O interesse em utilizá-las vem dos experimentos que mostram que o aumento da profundidade da rede traz ganhos no desempenho em tarefas de classificação, não apenas pelo acréscimo de complexidade, mas pela capacidade de extrair melhores representações [11]. Desse modo, o estudo de arquiteturas profundas se torna um passo natural na evolução das redes neurais para estruturas de maior complexidade, como as exploradas ao longo deste trabalho.

-Redes convolucionais (CNNs): Redes convolucionais são aquelas redes que possuem ao menos em uma camada a operação de convolução [11]. A operação de convolução em redes neurais é feita de maneira discretizada, da forma que sua equação, em duas dimensões, pode ser escrita como:

$$s(t) = (X * W)(i, j) = \sum_{k=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} X(k, y)W(i - k, j - y). \quad (1)$$

onde X é a matriz de entrada (*input*) e W é a matriz de pesos (*kernel*). A operação de convolução pode ser interpretada visualmente como a sobreposição de um kernel sobre a matriz de entrada, em que cada posição possível de sobreposição resulta em uma saída, que é a soma ponderada dos elementos cobertos nesta região. A figura 1 ilustra o raciocínio, onde temos um *input* de (5x5) e um *kernel* de (3x3).

Usualmente, a aplicação da convolução em redes neurais utiliza *kernels* muito menores que as matrizes de entrada [11], abordagem que é fundamental para as principais vantagens de sua aplicação. Entre elas, destaca-se o fato de que a convolução proporciona conexões esparsas e compartilhamento de parâmetros. Para explicar a relevância destes fatos, é importante ressaltar que nas redes neurais tradicionais cada saída depende de um peso distinto para cada elemento da matriz de entrada. Já no processo de convolução, utilizando um kernel menor que a matriz de entrada, o número de pesos a ser calculado

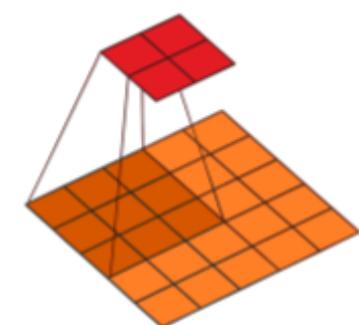


Figura 1: Ilustração de convolução: Região do input (5x5): Laranja; Região de sobreposição input (5x5) com kernel (3x3): Laranja escuro; Matriz de saída (2x2): Vermelho. Adaptado de [12]

para cada saída da rede é limitado ao tamanho do kernel, diminuindo assim a quantidade de parâmetros substancialmente [11] e viabilizando o processamento de arquiteturas muito mais profundas. Além disso, ao utilizar um *kernel* fixo, é de se imaginar que diferentes posições da matriz de entrada compartilham o mesmo parâmetro, permitindo que a rede correlacione pontos vizinhos entre si, em vez de tratá-los isoladamente. Uma importante aplicação desta propriedade é o uso de redes convolucionais para reconhecimento de imagens. Enquanto uma rede neural tradicional iria aprender os parâmetros pixel por pixel, a abordagem convolucional, “agrupa” os pixels vizinhos, tratando-os em conjunto, de forma que a rede analisa a vizinhança de cada pixel para

criar um mapa de características que ajuda a classificar a objeto em diferentes porções da imagem, ao invés de tratar a imagem como um todo. Aplicação das redes convolucionais no projeto se beneficia destas duas vantagens, já que há o interesse na aplicação de redes profundas (consequentemente, é necessário diminuir a quantidade de parâmetros para viabilizar computacionalmente certas arquiteturas profundas) e é compreensível a intuição de que dados de EEG possuam algum grau de correlação com as vizinhanças, tantos espaciais (como eletrodos próximos), quanto temporais (como sinais emitidos em instantes próximos).

-*Multi-Task Learning* (MTL): No contexto do aprendizado de máquina, uma tarefa refere-se à realização de uma previsão, classificação ou inferência a partir de um conjunto de dados disponíveis, com base em um objetivo específico de modelagem [13]. Nesse cenário, o *Multi-Task Learning* é um paradigma em que múltiplas tarefas relacionadas são aprendidas de forma conjunta, por meio da otimização simultânea do modelo [14]. Essa abordagem contrasta com a abordagem tradicional de aprendizado de tarefa única (STL, do inglês *Single-Task Learning*), onde cada tarefa é tratada de forma independente, sem qualquer interação ou compartilhamento de conhecimento entre tarefas distintas [14]. A diferença entre as abordagens é ilustrada abaixo:

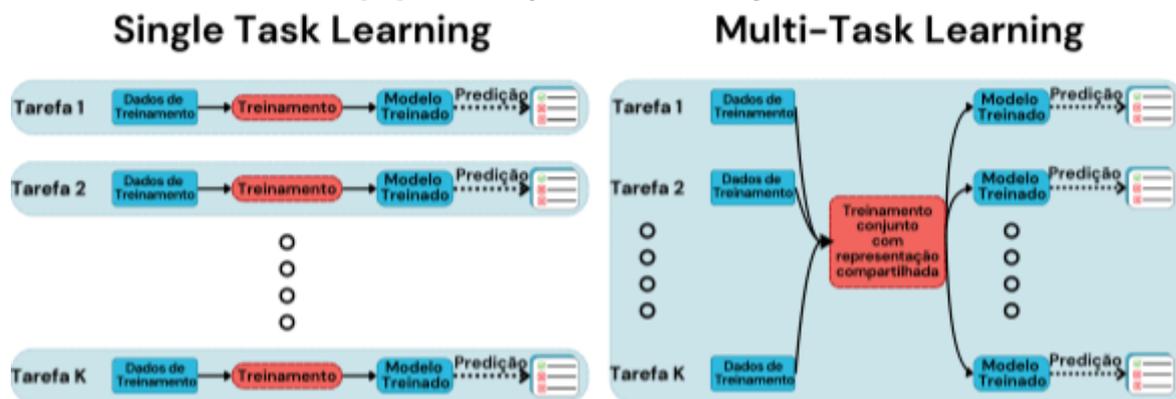


Figura 2: Metodologia de aprendizado de tarefa única e aprendizado de múltiplas tarefas. Adaptado de [15]

A principal motivação para o uso de MTL reside na hipótese de que tarefas relacionadas podem fornecer sinais de aprendizado úteis umas às outras, permitindo que o modelo compartilhe informações e parâmetros de diferentes tarefas e aproveite correlações entre si. Ao aprender representações mais generalizáveis, o modelo pode obter melhor desempenho em tarefas individuais do que aquele alcançado por modelos treinados isoladamente, especialmente em cenários com dados escassos [14].

RESULTADOS E DISCUSSÃO:

Realizaram-se dois experimentos de treinamento e avaliação utilizando a arquitetura EEGNet: um baseado no treinamento tradicional e outro empregando a estratégia Leave-One-Subject-Out (LOSO). Na aplicação tradicional, selecionou-se um único sujeito do banco de dados, cujos sinais foram divididos ao meio. Metade dos registros destinou-se ao treinamento, com subdivisão interna de 80% para ajuste dos pesos da rede e 20% para validação, enquanto a outra metade compôs o conjunto de teste. Essa abordagem, embora seja a mais usual em estudos de EEG, tende a apresentar acurácias elevadas nos dados de teste, pois a rede aprende padrões específicos do próprio indivíduo, capturando particularidades que dificilmente surgiriam em outros participantes. Em contraste, na abordagem LOSO, a rede foi treinada usando dados de oito sujeitos, deixando-se um nono indivíduo completamente reservado para avaliação do teste. Do conjunto de treinamento LOSO, utilizou-se 50 % dos registros de cada participante, novamente repartidos em 80% para treinamento e 20% para validação. Para o teste, adotaram-se os mesmos 50% de dados do sujeito deixado de fora, exatamente aqueles usados para teste na configuração tradicional. Os gráficos dos resultados podem ser conferidos abaixo:

Abordagem Intra-indivíduo

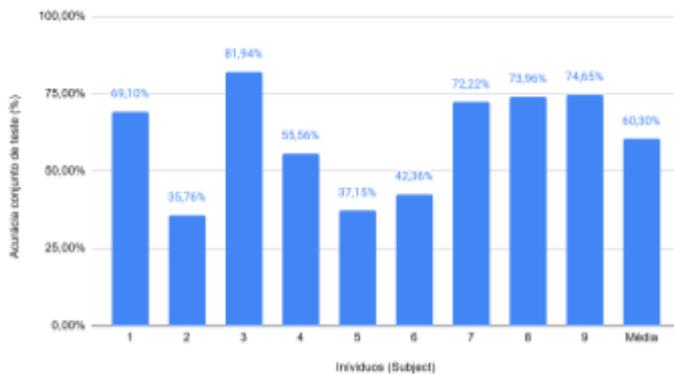


Figura 3: Gráfico de acurácias do conjunto de teste, em diferentes indivíduos, utilizando o treinamento tradicional

Abordagem LOSO

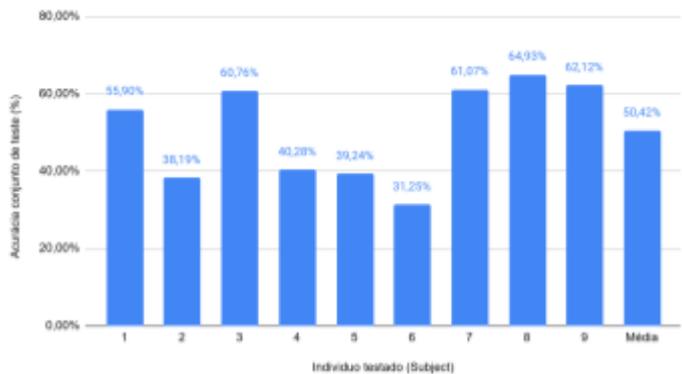


Figura 4: Gráfico de acurácias do conjunto de teste, em diferentes indivíduos, utilizando o treinamento LOSO

Os resultados obtidos indicam que o modelo tradicional apresenta acurácia ligeiramente superior ao LOSO, ao se analisar o desempenho no conjunto de testes. Esse resultado é esperado, pois, no protocolo tradicional, o mesmo indivíduo fornece tanto os dados de treino quanto os de teste. Dessa forma, a rede neural tende a aprender padrões neurais específicos daquele sujeito, como assinaturas elétricas particulares, e os utiliza como base para a classificação. Essa proximidade entre os dados de treino e teste favorece o desempenho da rede, mas compromete sua capacidade de generalização.

Ao adotar o protocolo LOSO (Leave-One-Subject-Out), observa-se uma queda significativa na acurácia, o que é natural, já que a rede precisa generalizar a partir de padrões aprendidos em múltiplos indivíduos e aplicar esse conhecimento a um sujeito nunca visto. Ainda assim, essa abordagem é mais representativa de aplicações reais, onde frequentemente não se dispõe de dados prévios do usuário final. Modelos treinados com LOSO evitam que a rede se prenda a características específicas de um único sujeito, promovendo uma representação mais robusta e generalizável dos sinais cerebrais.

Este achado exemplifica uma troca fundamental em BCIs: maximizar a acurácia em um único usuário contra promover uma aplicabilidade mais ampla. Na prática, redes treinadas pelo protocolo LOSO revelam-se mais adequadas para cenários de uso geral, onde o sistema deverá lidar com sinais de usuários não contemplados durante o treinamento. Em suma, embora o método tradicional ofereça resultados melhores em condições controladas para um indivíduo específico, o LOSO representa uma abordagem mais promissora para aplicações reais, garantindo maior confiabilidade ao processar dados cerebrais de novos participantes. Futuramente, pretendemos aplicar a abordagem MTL utilizando uma configuração dos dados de entrada que se assemelham ao LOSO.

CONCLUSÕES:

Os experimentos comparativos entre a rede EEGNet com o treinamento tradicional e com protocolo Leave-One-Subject-Out (LOSO) evidenciaram que, embora o modelo tradicional apresente acurácia ligeiramente superior no conjunto de testes, o LOSO oferece uma capacidade de generalização muito mais robusta a usuários não vistos durante o treinamento. A queda de acurácia observada no protocolo LOSO confirma o desafio inerente à variabilidade interindividual em sinais EEG de imagética motora. Ainda assim, a estratégia LOSO demonstrou-se promissora para cenários práticos, pois reduz o viés de indivíduo e promove representações mais genéricas, fundamentais para aplicações de BCI em ambientes reais, onde dados de novos usuários não estão disponíveis a priori.

Entre as limitações deste estudo, destaca-se o uso de uma única arquitetura (EEGNet) e de um número restrito de dados. Futuras pesquisas devem explorar arquiteturas alternativas de deep learning, incorporar técnicas de seleção ou fusão de características além de investigar abordagens híbridas que combinam MTL com *transfer learning* para potencialmente elevar ainda mais a generalização entre diferentes populações de usuários.

BIBLIOGRAFIA

- [1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [2] P. O. de Paula, T. B. da Silva Costa, R. Attux, and D. G. Fantinato. *Classification of image encoded SSVEP-based EEG signals using Convolutional Neural Networks. Expert Systems with Applications*, 214:119096, 2023.
- [3] I. Guyon and A. Elisseeff. *An Introduction to Variable and Feature Selection. Journal of machine learning research*, 3(Mar):1157-1182, 2003.
- [4] S. Haykin. *Neural Networks and Learning Machines*. Pearson Education, 2010.
- [5] A. Kübler and K.-R. Müller. *An Introduction to Brain-Computer Interfacing. Toward brain- computer interfacing*, pages 1-25, 2007.
- [6] C. S. Nam, A. Nijholt, and F. Lotte. *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*. CRC Press, 2018.
- [7] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. 1989.
- [10] C. Brunner , R. Leeb , G. R. Muller-Putz , A. Schlogl , and G. Pfurtscheller *BCI Competition 2008 – Graz data set A, 2008*
- [11] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [12] T. Esteves. *Uma introdução aos diferentes tipos de convoluções*. Medium, Aug. 22, 2022. Disponível em: <https://estevestoni.medium.com/uma-introdução-aos-diferentes-tipos-de-convoluções-d3ce7cd81759>.
- [13] Microsoft. *Tarefas de machine learning*. Microsoft Ignite, Jan. 4 2025. Disponível em: <https://learn.microsoft.com/pt-br/dotnet/machine-learning/resources/tasks>.
- [14] J. Yu, Y. Dai, X. Liu, J. Huang, Y. Shen, K. Zhang, R. Zhou, E. Adhikarla, W. Ye, Y. Liu, Z. Kong, K. Zhang, Y. Yin, V. Namboodiri, B. D. Davison, J. H. Moore, and Y. Chen. *Unleashing the Power of Multi-Task Learning: A Comprehensive Survey Spanning Traditional, Deep, and Pretrained Foundation Model Eras*. arXiv preprint arXiv:2404.18961v1 [cs.LG], 29 Apr. 2024.
- [15] A. R. Gonçalves. *Sparse and structural multitask learning = Aprendizado multitarefa estrutural e esparso*. PhD thesis, [sn], 2016.