



Estudo do Método de Máxima Discrepância de Classificação em Interfaces Cérebro-Computador Baseadas no Paradigma de Imagética Motora

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

Autores:

Gabriel da Costa, FEEC – UNICAMP

Prof. Denis Fantinato (orientador), FEEC - UNICAMP

INTRODUÇÃO:

As interfaces cérebro-computador (BCIs) são tecnologias emergentes que estabelecem uma comunicação direta entre a atividade cerebral humana e sistemas computacionais, abrindo caminho para aplicações transformadoras em reabilitação motora, controle de próteses robóticas, sistemas de comunicação alternativa para indivíduos com limitações motoras e até mesmo em plataformas de entretenimento [1,6]. A eficácia desses sistemas depende da precisão com que os algoritmos de classificação interpretam os padrões neurais e os transformam em comandos executáveis. Atualmente, este trabalho concentra-se em sua maioria em BCIs baseadas em eletroencefalografia (EEG) de superfície, modalidade que se destaca por seu caráter não invasivo, relativa acessibilidade econômica e menor complexidade operacional quando comparada a técnicas invasivas ou semi-invasivas [6].

Um sistema típico de BCI envolve uma cadeia sequencial de processamento que se inicia com a aquisição de sinais elétricos através de eletrodos posicionados estrategicamente no couro cabeludo, seguida por etapas de pré-processamento onde técnicas de filtragem (como filtros passa-banda) e algoritmos de remoção de artefatos (que podem ser eletromagnéticos, musculares ou oculares) são aplicados para purificar o sinal bruto [2,7]. Posteriormente, na fase de extração de características, transformações matemáticas e métodos de análise espectral/temporal (como a transformada de Fourier) são empregados para identificar padrões discriminativos nos dados, convertendo-os em representações mais adequadas para alimentar os classificadores. A etapa subsequente de seleção de características procura reduzir a dimensionalidade do problema, preservando apenas os atributos mais informativos, o que culmina no estágio final de classificação, onde algoritmos de aprendizado de máquina (tipicamente redes neurais) mapeiam as características extraídas para comandos ou intenções específicas do usuário.

Para esta pesquisa, adotou-se o paradigma de imagética motora, no qual o usuário imagina movimentos corporais específicos, como mover a mão esquerda ou os pés, gerando assim padrões característicos de ativação nas diferentes regiões cerebrais, que são refletidos nos sinais de tensão capturados pelos eletrodos, sobretudo os mais próximos ao córtex motor [7]. Apesar do potencial demonstrado em ambientes controlados, desafios significativos ainda existem quando se almeja a implantação de BCIs em cenários reais. Os sinais de EEG apresentam amplitude extremamente baixa, na faixa de microvolts, o que os torna altamente suscetíveis a interferências tanto de origem intrínseca, como atividade cognitiva concorrente, quanto extrínsecas, como ruídos eletromagnéticos da rede ou artefatos de movimento [6]. Além disso, observa-se uma elevada variabilidade intra e inter-indivíduo, fenômeno que se manifesta como mudanças significativas nas distribuições estatísticas dos dados entre diferentes sessões de aquisição ou entre indivíduos distintos, comprometendo a capacidade de generalização dos modelos de classificação convencionais, que são tipicamente treinados sob a suposição ingênua de homogeneidade dos dados [5].

Diante deste cenário desafiador, o presente trabalho investiga o Método de Máxima Discrepância de Classificação (MDC), uma abordagem proposta originalmente para *support vector machines* (SVM), mas que foi

implementada por Saito et al. [9] no contexto de adaptação de domínio, como estratégia promissora para mitigar os problemas de transferência de domínio em BCIs com base em sinais de EEG.

A iniciação como um todo propõe uma investigação que abrange desde o aprendizado da teoria básica subjacente ao aprendizado de máquina tradicional e redes neurais profundas até ao método MDC e sua implementação prática e avaliação em um cenário mais realista de classificação de sinais de EEG associados à imagética motora, utilizando para tal conjuntos de dados publicamente disponíveis e validados pela comunidade científica. O objetivo da aplicação do MDC é melhorar a robustez e generalização dos classificadores para múltiplos indivíduos, superando as limitações das abordagens convencionais intra-indivíduo, contribuindo assim para o desenvolvimento de BCIs mais confiáveis e práticas.

METODOLOGIA:

O desenvolvimento desta pesquisa seguiu uma abordagem prática, integrando o estudo teórico dos fundamentos de aprendizado de máquina profundo, a implementação computacional de redes neurais especializadas para processamento de sinais de EEG, e a aplicação experimental de técnicas mais avançadas de aprendizado por transferência em cenários de classificação de sinais neurais. A primeira fase consistiu em um estudo dos princípios matemáticos e algorítmicos que sustentam as arquiteturas de redes neurais, começando desde redes monocamadas até as redes convolucionais (CNNs).

Houve uma ênfase particular na capacidade das CNNs de capturar padrões espaciais e temporais complexos em dados de alta dimensionalidade, característica intrínseca aos sinais de EEG devido à multiplicidade de canais de aquisição [1,3,4]. A operação de convolução bidimensional, constitui o cerne dessas arquiteturas e proporciona duas vantagens fundamentais para o processamento de sinais de EEG: conexões esparsas, nas quais cada neurônio na camada seguinte conecta-se apenas a uma região pequena da camada anterior, e compartilhamento de parâmetros, onde os mesmos pesos são reutilizados em diferentes posições espaciais, reduzindo consideravelmente a complexidade computacional e possibilitando a existência de arquiteturas profundas [3].

Para a pesquisa, adotou-se como base a arquitetura EEGNet v4, uma rede neural convolucional projetada especificamente para classificação eficiente de EEG, que combina operações convolucionais temporais e espaciais em uma estrutura hierárquica [10]. As convoluções temporais, implementadas através de kernels alongados no eixo do tempo, são capazes de extrair características espectrais relevantes associadas aos ritmos/espectro em frequência dos sinais, enquanto as convoluções espaciais, aplicadas via kernels profundos que operam através dos múltiplos canais de EEG, integram informações de diferentes regiões do escalpo, capturando assim os padrões de ativação distribuídos característicos da imagética motora [7]. Esta abordagem modular permite que a rede aprenda representações hierárquicas dos sinais EEG, evoluindo desde características locais de baixo nível até padrões globais de alto nível associados às intenções motoras do usuário.

Na fase prática de implementação computacional, realizada integralmente na linguagem Python, planejou-se um conjunto de três experimentos sequenciais e comparativos utilizando a mesma base de dados (BNCI2014_001), que possui dados rotulados de 4 classes de imagética motora para 9 indivíduos [8]. O primeiro experimento consistiu na aplicação detalhada da abordagem tradicional de tarefa única (Single-Task Learning - STL), onde a rede EEGNet foi treinada e testada individualmente para cada sujeito, utilizando metade dos dados para treinamento (com subdivisão interna de 2/3 para ajuste fino dos pesos e 1/3 para validação) e a outra metade reservada exclusivamente para teste.

O segundo experimento implementou a abordagem de aprendizado alternativa, ainda em STL, mas utilizando a estratégia de Leave-One-Subject-Out (LOSO), na qual, em cada iteração do processo, dados de oito sujeitos foram agregados para formar o conjunto de treinamento, enquanto o nono sujeito foi completamente reservado para avaliação de generalização.

Complementando, nos testes também foram implementadas técnicas de pré-processamento, incluindo filtragem passa-banda na faixa de 4 a 38 Hz para isolar as componentes espectrais relevantes para a imagética motora, e aplicação do referenciamento CAR (Common Average Reference) para mitigar artefatos de referência e melhorar a relação sinal-ruído [7]. Todo o trabalho computacional, desde o carregamento dos dados até a classificação final, foi implementado utilizando bibliotecas especializadas, como PyTorch para construção e treinamento das redes neurais e MOABB (Mother of All BCI Benchmarks) para padronização e reprodutibilidade dos experimentos com conjuntos de dados públicos.

Na etapa final o projeto avança para a implementação do Método de Máxima Discrepância de Classificação (MDC), seguindo a abordagem proposta por Saito et al. [9]. Esta técnica será adaptada e aplicada para enfrentar o cerne do problema de generalização em BCIs: a variabilidade das distribuições entre diferentes indivíduos.

Mais formalmente, a técnica MDC opera através de uma arquitetura adversarial que incorpora dois classificadores diferentes que trabalham de forma concorrente: enquanto um extrator de características compartilhado gera representações dos dados, os dois classificadores são otimizados para maximizar sua discrepância nas previsões sobre exemplos do domínio de destino, criando assim um sinal de treinamento que força o extrator a aprender representações invariantes às mudanças de domínio [9]. O método está exemplificado na figura 1. Esta abordagem contrasta com métodos tradicionais de alinhamento de domínio, oferecendo um mecanismo mais refinado para reduzir a divergência entre as distribuições de dados de origem (treinamento) e destino (teste/aplicação) no espaço latente de características. A hipótese norteadora desta pesquisa é que a aplicação do MDC na etapa de classificação pode melhorar significativamente a adaptabilidade do modelo a novos usuários ou sessões experimentais, reduzindo assim a necessidade de processos de recalibração, que são caros e demorados, e limitam a adoção prática das BCIs.

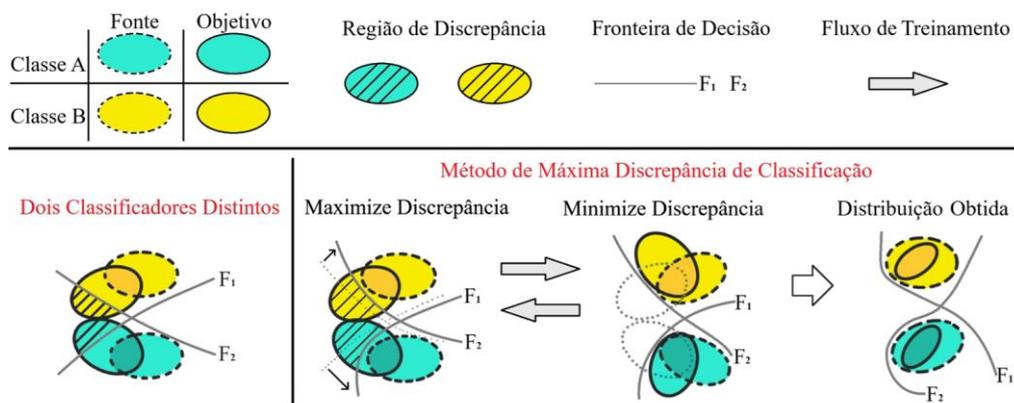
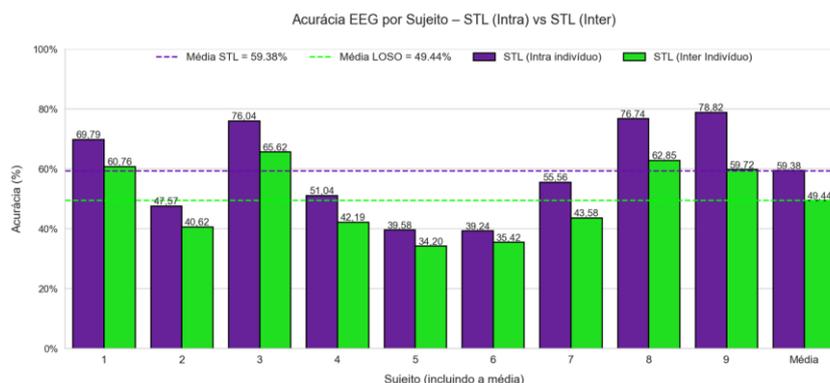


Figura 1: Ilustração do método de máxima discrepância de

RESULTADOS E DISCUSSÃO:

Os experimentos realizados até o momento já proporcionam insights valiosos sobre o comportamento comparativo das diferentes abordagens de classificação. Na configuração STL, onde cada modelo é treinado e testado exclusivamente com dados do mesmo indivíduo, observou-se desempenho razoável, com acurácias médias superiores a 75% na classificação das intenções motoras de três indivíduos, embora com variações significativas entre sujeitos que atingiram até 45 pontos percentuais, reflexo direto da conhecida variabilidade inter-indivíduo em respostas neurais [5]. Este bom desempenho intra-sujeito, entretanto, mascara uma limitação fundamental: os modelos desenvolveram forte dependência de características específicas de cada indivíduo, resultando péssimos resultados quando expostos a dados de novos usuários. Isso foi observado quando se tentou aplicar os modelos treinados em dados de um único sujeito, a outro, foram obtidas quedas de acurácia da ordem de 40 pontos percentuais. Os resultados obtidos como um todo estão apresentados na figura 2.



A abordagem com protocolo LOSO apresentou um perfil de desempenho distinto. Ao agregar dados de múltiplos usuários durante o treinamento, os modelos demonstraram alguma capacidade de generalização para novos indivíduos, alcançando acurácias médias na faixa de 50% no conjunto de teste composto por sujeitos completamente ausentes do treinamento. Esta melhoria na generalização, contudo, veio acompanhada de uma redução média de 10 pontos percentuais em relação aos resultados STL, padrão já reportado na literatura para paradigmas similares. A variabilidade de desempenho entre sujeitos também se mostrou pronunciada nesta configuração, com oscilações de até 30 pontos percentuais, indicando que certos padrões neurais individuais permanecem difíceis de capturar por representações apreendidas de forma agregada.

CONCLUSÕES:

A iniciação trata sobre uma investigação abrangente sobre a aplicação de métodos de aprendizado de máquina para classificação de sinais de EEG e almeja aplicar o Método de Máxima Discrepância de Classificação (MDC) em interfaces cérebro-computador baseadas no paradigma de imagética motora, abordando o problema da variabilidade intrínseca observada entre diferentes indivíduos. Os resultados obtidos, ainda que preliminares, apontam para conclusões significativas e de amplo alcance. Em primeiro lugar, confirmou-se que abordagens convencionais, tanto na configuração STL intra e inter indivíduo, enfrentam limitações intrínsecas impostas pela incompatibilidade fundamental nas distribuições estatísticas de sinais EEG entre usuários, um fenômeno já documentado na literatura [5] mas ainda carente de soluções ideais.

A proposta do MDC vem como uma estratégia promissora neste cenário, demonstrando potencial para superar as limitações das abordagens estabelecidas. Objetiva-se manter um desempenho competitivo, acima de 65% em indivíduos completamente ausentes do conjunto de treinamento. Caso isso seja atingido, estaríamos mais próximos de uma arquitetura BCI que requer mínima recalibração ao ser transferida para novos usuários. Para os próximos passos da pesquisa, planeja-se finalizar a implementação do método MDC no conjunto de dados que já vem sendo trabalhado, para avaliar sua escalabilidade e robustez e finalizar a fase de implementação prática da iniciação.

BIBLIOGRAFIA

- [1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [2] T. Costa, S. Carvalho, D. Soriano, e R. Attux. Discriminantes lineares para uso em sistemas SSVEP-BCI. *Artificial intelligence*, 2014.
- [3] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT press, 2016.
- [4] S. Haykin. *Neural Networks and Learning Machines*. Pearson Education, 2010.
- [5] M. H. N. Z. X. L. Zitong Wan, Rui Yang. A review on transfer learning in EEG signal analysis. *Neurocomputing*, 421:1–14, 2021.
- [6] A. Kübler and K.-R. Müller. An Introduction to Brain-Computer Interfacing. *Toward brain-computer interfacing*, pages 1-25, 2007.
- [7] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. 1989.
- [8] C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlogl, and G. Pfurtscheller. BCI Competition 2008 – Graz data set A, 2008
- [9] K. Saito, K. Watanabe, Y. Ushiku, and T. Harada. Maximum Classifier Discrepancy for Unsupervised Domain Adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [10] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance. EEGNet: A Compact Convolutional Neural Network for EEG-based Brain-Computer Interfaces. *Journal of Neural Engineering*, 15(5):056013, 2018.