



UNIVERSIDADE FEDERAL DO PARÁ  
CAMPUS UNIVERSITÁRIO DE TUCURUÍ  
FACULDADE DE ENGENHARIA ELÉTRICA

GUILHERME DA COSTA DOURADO

**Otimização de Interfaces Cérebro-Máquina Baseadas em Imagética Motora  
por Seleção de Features e Combinação de Modelos: Compilado de artigos  
publicados em 2024 e 2025**

TUCURUÍ

2026

GUILHERME DA COSTA DOURADO

**Otimização de Interfaces Cérebro-Máquina Baseadas em Imagética Motora  
por Seleção de Features e Combinação de Modelos: Compilado de artigos  
publicados em 2024 e 2025**

Trabalho de Conclusão de Curso, apresentado  
como requisito parcial para a obtenção de  
grau de Bacharel em Engenharia Elétrica,  
pela Universidade Federal do Pará.

Orientador:  
Prof. Dr. Cleison Daniel Silva.

TUCURUÍ

2026

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD  
Sistema de Bibliotecas da Universidade Federal do Pará  
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

---

D739o Dourado, Guilherme da Costa.  
Otimização de Interfaces Cérebro-Máquina Baseadas em  
Imagética Motora por Seleção de Features e Combinação de  
Modelos: Compilado de artigos publicados em 2024 e 2025 /  
Guilherme da Costa Dourado. — 2026.  
33 f. : il. color.

Orientador(a): Prof. Dr. Cleison Daniel Silva  
Trabalho de Curso (Graduação) - Universidade Federal do Pará,  
Campus Universitário de Tucuruí, Faculdade de Engenharia  
Elétrica, Tucuruí, 2026.

1. Interfaces Cérebro-Máquina. 2. Imagética motora. 3.  
Seleção de features. 4. Otimização Bayesiana. 5. Aprendizado  
em conjunto. I. Título.

CDD 620

---

Guilherme da Costa Dourado

**Otimização de Interfaces Cérebro-Máquina Baseadas em Imagética Motora  
por Seleção de Features e Combinação de Modelos: Compilado de artigos  
publicados em 2024 e 2025**

Trabalho de Conclusão de Curso, apresentado como requisito parcial para a obtenção de grau de Bacharel em Engenharia Elétrica, pela Universidade Federal do Pará.

Data de aprovação: 26/05/2026.

**Banca Examinadora:**

---

**Prof. Dr. Cleison Daniel Silva**  
Orientador - FEE/CAMTUC/UFPA

---

**Prof. Dr. Davi Carvalho Moreira**  
Avaliador Interno - FEE/CAMTUC/UFPA

---

**Prof. Dr. Jefferson Souza Costa**  
Avaliador Interno - FEE/CAMTUC/UFPA

TUCURUÍ

2026



UNIVERSIDADE FEDERAL DO PARÁ  
CAMPUS UNIVERSITÁRIO DE TUCURUÍ  
FACULDADE DE ENGENHARIA ELÉTRICA

<b>Título</b>	Otimização de Interfaces Cérebro-Máquina baseadas em imagética Motora por Seleção de Features e Combinação de Modelos: Compilado de artigos publicados em 2024 e 2025.
<b>Discente</b>	Guilherme da Costa Dourado
<b>Matrícula</b>	

#	BANCA EXAMINADORA	CONDIÇÃO
1	Cleison Daniel Silva (UFPA/NDAE)	Orientador
2	Davi Carvalho Moreira (UFPA/CAMTUC)	Membro
3	Jefferson Sousa Costa (UFPA/CAMTUC)	Membro

Data da Defesa: 26/05/2026	Hora Início: 9:00	Hora Término: 10:40
----------------------------	-------------------	---------------------

Trabalho Escrito (0 a 10 pontos por critério)			
Critério	Examinador 1	Examinador 2	Examinador 3
Formatação	9,0	9,5	8,5
Linguagem (gramática e semântica)	9,0	10,0	10,0
Conteúdo técnico	9,5	10,0	10,0
<b>Média</b>	<b>9,2</b>	<b>9,8</b>	<b>9,5</b>

Defesa Oral (0 a 10 pontos por critério)			
Critério	Examinador 1	Examinador 2	Examinador 3
Sequência lógica de apresentação	10,0	10,0	10,0
Administração do tempo	10,0	10,0	10,0
Expressão oral	9,5	10,0	10,0
Domínio do tema	9,5	10,0	9,5
<b>Média</b>	<b>9,8</b>	<b>10,0</b>	<b>9,9</b>

<b>Média por examinador</b>	9,5	9,9	9,7
<b>Média Final</b>	<b>9,7</b>		
<b>Conceito Final</b>	<b>EXCELENTE</b>		

Documento assinado digitalmente  
**gov.br** JEFFERSON SOUZA COSTA  
Data: 26/05/2026 11:03:52-0300  
Verifique em <https://validar.iti.gov.br>

Jefferson Sousa Costa  
*Membro*

Documento assinado digitalmente  
**gov.br** DAVI CARVALHO MOREIRA  
Data: 26/05/2026 11:07:19-0300  
Verifique em <https://validar.iti.gov.br>

Davi Carvalho Moreira  
*Membro*

Documento assinado digitalmente  
**gov.br** CLEISON DANIEL SILVA  
Data: 28/05/2026 18:33:57-0300  
Verifique em <https://validar.iti.gov.br>

Cleison Daniel Silva  
*Orientador*

Dedico este trabalho aos meus pais e à minha vó, que sempre acreditaram em mim e me ensinaram o valor do esforço e da honestidade. A todas as pessoas que estiveram ao meu lado ao longo dessa trajetória, pelo apoio, pela paciência e pelo incentivo constante, especialmente nos momentos mais desafiadores. E a Deus, por me conceder força e sabedoria para chegar até aqui.

## AGRADECIMENTOS

Aos meus pais, pelo apoio incondicional, pelos ensinamentos e pelo exemplo de dedicação que sempre me motivaram a seguir em frente. À minha avó, pelo carinho, pelas orações e por sua constante presença, que representou uma importante fonte de força e inspiração ao longo desta jornada.

A todas as pessoas que estiveram ao meu lado durante essa trajetória, pela compreensão, incentivo e apoio nos momentos mais desafiadores, contribuindo para que eu mantivesse o foco e a determinação necessários à conclusão deste trabalho.

Ao meu orientador, pela orientação, disponibilidade e pelas contribuições fundamentais para o desenvolvimento desta pesquisa.

Aos professores do curso de Engenharia Elétrica da Universidade Federal do Pará – Campus Tucuruí, pelo comprometimento com o ensino e pela contribuição significativa na minha formação acadêmica e profissional.

Aos colegas e amigos que fizeram parte desta trajetória, pelo companheirismo e pelas trocas de conhecimento que enriqueceram minha experiência universitária.

E, finalmente, à Universidade Federal do Pará, pela oportunidade de aprendizado e pela estrutura disponibilizada para a realização deste trabalho.

"Ele fortalece ao cansado e dá grande vigor ao que está sem forças.  
Até os jovens se cansam e ficam exaustos, e os moços tropeçam e caem;  
mas aqueles que esperam no Senhor renovam as suas forças. Voam bem alto como águias;  
correm e não ficam exaustos, andam e não se cansam."  
(Isaías 40:29-31).

## RESUMO

Este Trabalho de Conclusão de Curso aborda o estudo e a otimização de sistemas de Interfaces Cérebro-Máquina (ICM) baseadas em Imagética Motora (IM), com foco na melhoria da precisão e da robustez de classificadores aplicados a sinais de Eletroencefalografia (EEG). O trabalho está dividido em dois estudos complementares, ambos utilizando o conjunto de dados público 2a da IV Competição de ICM. No primeiro estudo, foi investigado o impacto da seleção de *features* sobre o desempenho de sistemas ICM-IM. A metodologia incluiu filtragem por sub-bandas, aplicação do método *Common Spatial Pattern* (CSP) para geração de vetores de características e seleção por meio dos algoritmos *Least Absolute Shrinkage and Selection Operator* (LASSO) e *SelectKBest*, seguida de classificação com *Support Vector Machine* (SVM). Os resultados indicaram um ganho médio de 2% na acurácia da classificação, demonstrando que a seleção criteriosa de características contribui para reduzir redundâncias e aprimorar o desempenho dos classificadores. No segundo estudo, foram exploradas técnicas de aprendizado em conjunto, como *Bagging* e *Soft Voting*, aplicadas à combinação de modelos personalizados via otimização Bayesiana. As análises mostraram um aumento médio de 14,4% nos valores do índice *Kappa* médio ( $\kappa_m$ ), métrica que avalia a concordância entre as predições do classificador e os rótulos reais, em relação aos modelos individuais. Entretanto, não foi observada significância estatística nos valores de  $z_{\kappa}$ , indicando que os ganhos permaneceram sensíveis à variabilidade individual dos sinais de EEG. Em conjunto, os dois estudos configuram uma proposta de otimização em múltiplos estágios para sistemas de ICM, contribuindo para o desenvolvimento de soluções mais precisas, interpretáveis e adaptáveis, com potencial aplicação em contextos clínicos e assistivos.

**Palavras-chave:** Interfaces Cérebro-Máquina; imagética motora; seleção de *features*; otimização Bayesiana; aprendizado em conjunto.

## ABSTRACT

This undergraduate thesis addresses the study and optimization of Motor Imagery (MI)-based Brain-Computer Interface (BCI) systems, focusing on improving the accuracy and robustness of classifiers applied to Electroencephalography (EEG) signals. The work is divided into two complementary studies, both using the public dataset 2a from the IV BCI Competition. In the first study, the impact of feature selection on the performance of MI-BCI systems was investigated. The methodology included sub-band filtering, application of the Common Spatial Pattern (CSP) method for feature vector generation, and feature selection using the Least Absolute Shrinkage and Selection Operator (LASSO) and SelectKBest algorithms, followed by classification with Support Vector Machine (SVM). The results indicated an average gain of 2% in classification accuracy, demonstrating that careful feature selection contributes to reducing redundancies and improving classifier performance. In the second study, ensemble learning techniques such as Bagging and Soft Voting were explored, applied to the combination of personalized models through Bayesian optimization. The analyses showed an average increase of 14.4% in the mean *Kappa* index ( $\kappa_m$ ), a metric that evaluates the agreement between classifier predictions and true labels, compared to individual models. However, no statistical significance was observed in the values of  $z_{\kappa}$ , indicating that the gains remained sensitive to the individual variability of EEG signals. Together, the two studies constitute a multi-stage optimization approach for BCI systems, contributing to the development of more accurate, interpretable, and adaptable solutions, with potential applications in clinical and assistive contexts.

**Keywords:** Brain-Computer Interface; motor imagery; feature selection; Bayesian optimization; ensemble learning.

## SUMÁRIO

1	INTRODUÇÃO . . . . .	11
2	SELEÇÃO DE FEATURES POR SUB-BANDA APLICADA À INTERFACE CÉREBRO-MÁQUINA BASEADA EM IMA- GÉTICA MOTORA . . . . .	14
3	COMBINAÇÃO DE MODELOS PERSONALIZADOS DE IN- TERFACES CÉREBRO-MÁQUINA BASEADAS EM IMAGÉ- TICA MOTORA COM OTIMIZAÇÃO BAYESIANA . . . . .	20
4	CONCLUSÃO FINAL . . . . .	28
	REFERÊNCIAS . . . . .	29

# 1 INTRODUÇÃO

As Interfaces Cérebro-Máquina (ICMs), também conhecidas como *Brain-Computer Interfaces* (BCI), representam uma das áreas mais inovadoras e desafiadoras da engenharia contemporânea, integrando conhecimentos de processamento de sinais, aprendizado de máquina e neuroengenharia. Segundo Lotte (2014), esses sistemas permitem a comunicação direta entre o cérebro humano e dispositivos externos, sem a necessidade de ação muscular, utilizando sinais cerebrais obtidos, em sua maioria, por meio da Eletroencefalografia (EEG). Essa tecnologia tem sido amplamente investigada nas últimas décadas por seu potencial em aplicações assistivas, médicas e de reabilitação, especialmente em casos de limitações motoras severas (KIM; CARLSON; LEE, 2013; PAPANASTASIOU et al., 2020).

De acordo com Lindig e Bougrain (2015), entre as diversas abordagens de ICM, destacam-se aquelas baseadas em Imagética Motora (IM), nas quais o usuário realiza mentalmente a simulação de um movimento, como mover a mão ou o pé, gerando padrões cerebrais específicos que podem ser decodificados por algoritmos de aprendizado de máquina. Entretanto, a implementação de sistemas ICM-IM ainda enfrenta desafios significativos, como a baixa relação sinal-ruído dos dados de EEG, a variabilidade interindividual e a dificuldade de generalização dos classificadores (LOTTE, 2014). Esses fatores limitam a precisão e a confiabilidade dos sistemas, representando lacunas que motivam o desenvolvimento de novas estratégias de otimização.

Nesse contexto, o presente trabalho justifica-se pela necessidade de aprimorar o desempenho e a robustez de sistemas de ICM-IM, explorando técnicas de processamento e modelagem que possam mitigar os efeitos da variabilidade dos sinais cerebrais e aumentar a acurácia e a confiabilidade das classificações. A pesquisa busca contribuir para o avanço de metodologias mais eficazes, capazes de aliar desempenho computacional, interpretabilidade e potencial de aplicação em sistemas assistivos reais.

Durante a graduação em Engenharia Elétrica na Universidade Federal do Pará (UFPA), o autor deste trabalho desenvolveu, no âmbito do Laboratório de Inteligência, Controle e Eletrônica (LINCE), uma linha de pesquisa voltada à otimização de sistemas de Interfaces Cérebro-Máquina baseadas em Imagética Motora. O objetivo principal foi investigar métodos capazes de melhorar o desempenho da classificação de sinais EEG, tanto na etapa de seleção de características quanto na de combinação de modelos de aprendizado de máquina. Assim, o trabalho tem como objetivo geral propor e avaliar estratégias de otimização aplicadas a sistemas de ICM-IM, visando aprimorar a precisão e a confiabilidade dos classificadores. De modo complementar, busca analisar o impacto da seleção de *features* na acurácia de sistemas ICM-IM, investigar o efeito da combinação de modelos personalizados otimizados por otimização Bayesiana sobre a robustez dos

classificadores e discutir as contribuições e limitações de cada abordagem, relacionando-as às perspectivas de aplicação futura.

Para alcançar esses objetivos, foram desenvolvidos dois estudos complementares, ambos utilizando o conjunto de dados público 2a da IV Competição de ICM (BRUNNER et al., 2008). O primeiro artigo, intitulado “Seleção de Features por Sub-Banda Aplicada a Interface Cérebro-Máquina Baseada em Imagética Motora”, foi apresentado no XLII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT 2024. Nesse estudo, investigou-se o impacto da seleção de *features* em sistemas ICM-IM, utilizando os algoritmos *Least Absolute Shrinkage and Selection Operator* (LASSO) (MIAO et al., 2021) e *SelectKBest* (PEDREGOSA et al., 2011) aplicados a vetores de características extraídos com o método *Common Spatial Pattern* (CSP) (NOVI et al., 2007). Os resultados indicaram um ganho médio de 2% na acurácia de classificação, evidenciando a importância de selecionar características discriminativas para reduzir redundâncias e aprimorar o desempenho dos classificadores.

O segundo artigo, intitulado “Combinação de Modelos Personalizados de Interfaces Cérebro-Máquina Baseadas em Imagética Motora com Otimização Bayesiana”, foi publicado no XVII Congresso Brasileiro de Inteligência Computacional - CBIC 2025. Este estudo ampliou a abordagem anterior ao propor a combinação de modelos personalizados via otimização Bayesiana (VILAS-BOAS; SILVA, 2020), integrados por técnicas de aprendizado em conjunto, como o *Bagging* e o *Soft Voting* (RAHIMI et al., 2016; SEN; VERMA, 2023). A metodologia foi implementada em linguagem *Python*, com o uso de bibliotecas como *scikit-learn* (PEDREGOSA et al., 2011) e *MNE-Python* (GRAMFORT et al., 2013). Comparando os modelos combinados com os modelos individuais, os resultados indicaram um aumento médio de 14,4% no índice *Kappa* médio ( $\kappa_m$ ), que permite verificar se o desempenho obtido é realmente superior ao que seria esperado de forma aleatória (CHICCO; WARRENS; JURMAN, 2021). No entanto, ao avaliar o índice  $z_\kappa$  entre os modelos combinados e os melhores modelos individuais, verificou-se que, em nenhum dos casos, as diferenças alcançaram significância estatística. Esse resultado indica que, embora o aprendizado em conjunto tenha proporcionado um aumento médio no desempenho, tal melhoria não se traduziu em evidências estatísticas consistentes, permanecendo sensível à variabilidade individual dos sinais de EEG.

Os dois estudos estabelecem uma relação de continuidade e complementaridade. O primeiro concentra-se na melhoria da representação dos sinais cerebrais por meio da seleção de *features*, buscando identificar as componentes mais informativas para a discriminação de classes em tarefas de Imagética Motora. O segundo, por sua vez, avança para a etapa de fusão de modelos, investigando estratégias de combinação que visam aumentar a robustez e a capacidade de generalização dos classificadores. Em conjunto, essas abordagens compõem uma proposta de otimização em múltiplos estágios para sistemas de ICM, na

qual diferentes fases do fluxo de processamento, como a seleção de características e a integração de modelos, são aprimoradas de maneira sequencial e interdependente. Essa estratégia contribui para o desenvolvimento de soluções mais precisas e adaptáveis, com potencial aplicação em contextos clínicos e assistivos.

As principais dificuldades enfrentadas durante o desenvolvimento desses trabalhos estiveram relacionadas à complexidade técnica do processamento de sinais cerebrais, à necessidade de calibração individualizada dos modelos e às restrições de recursos computacionais. Essas limitações foram superadas por meio do uso de ferramentas de aprendizado de máquina, do ambiente colaborativo *Google Colab* e do apoio de docentes e pesquisadores vinculados ao LINCE.

O presente TCC, intitulado “Otimização de Interfaces Cérebro-Máquina Baseadas em Imagética Motora por Seleção de Features e Combinação de Modelos”, reúne esses dois artigos em formato de capítulos independentes e complementares. O Capítulo 2 apresenta a investigação sobre seleção de *features* por sub-banda, o Capítulo 3 descreve a abordagem baseada na combinação de modelos personalizados, e o Capítulo 4 traz as conclusões finais deste trabalho. Em conjunto, os artigos sintetizam os principais resultados alcançados, destacando as contribuições da pesquisa e apontando caminhos promissores para a continuidade dos estudos na área de Interfaces Cérebro-Máquina.

## **2 SELEÇÃO DE FEATURES POR SUB-BANDA APLICADA À INTERFACE CÉREBRO-MÁQUINA BASEADA EM IMAGÉTICA MOTORA**

Na próxima página encontra-se o arquivo original do artigo intitulado “Seleção de Features por Sub-Banda Aplicada a Interface Cérebro-Máquina Baseada em Imagética Motora”, apresentado na XLII edição do Simpósio Brasileiro de Telecomunicações e Processamento de Sinais – SBrT 2024, realizado em formato presencial no ano de 2024.

# Seleção de Features por Sub-Banda Aplicada a Interface Cérebro-Máquina Baseada em Imagética Motora

Cleison Silva, Guilherme Dourado, Raphael Teixeira, Vitor Vilas-Boas e Vitor Jorge

**Resumo**—Este artigo investiga como os métodos de seleção de *features* influenciam o desempenho de sistemas de Interfaces Cérebro-Máquina baseadas em Imagética Motora (ICM-IM). A proposta consiste em aumentar a redundância dos filtros espaciais via *Common Spatial Pattern* (CSP) e selecionar um subconjunto de *features* com os algoritmos LASSO e *SelectKBest*. Usando dados da IV Competição de ICM, a abordagem inclui filtragem por sub-banda, extração e seleção de *features*, seguida de classificação com SVM. Os resultados mostram um ganho médio de 2% na acurácia da classificação dos sinais, destacando a importância da seleção de *features* para aprimorar a performance do modelo.

**Palavras-Chave**—ICM-IM, Filtros Espaciais, Seleção de *Features*, Acurácia, Lasso, *SelectKBest*.

**Abstract**—This article investigates how methods of feature selection influences system performance of Brain-Machine Interfaces based on motor imagery (BMI-MI). The proposal consists of increasing the redundancy of spatial filters via *Common Spatial Pattern* (CSP) and select a subset of features with the LASSO and *SelectKBest* algorithms. Using data of the IV BCI Competition, the approach includes filtering by sub-band, feature extraction and selection, followed by classification with SVM. The results show a average gain of 2% in accuracy of signals classification, highlighting the importance of feature selection to improve model performance.

**Keywords**—BMI-MI, Spatial Filters, Features Selection, Accuracy, Lasso, *SelectKBest*.

## I. INTRODUÇÃO

Sistemas de Interfaces Cérebro-Máquina (ICMs) baseadas em Imagética Motora (IM) são formados por cadeias de processamento de sinais coletados durante a realização de tarefas mentais específicas, tais como a imaginação de movimentos. Padrões associados às tarefas mentais podem ser identificados em sinais coletados e utilizados para gerar comandos a serem aplicados em dispositivos externos ao corpo humano. Exemplos de aplicações de sistema

ICM são cadeiras de rodas, braços robóticos e Veículos Aéreos Não Tripulados (VANTs) comandados por ICMs [1].

A atividade elétrica cerebral observada durante a Imagética Motora é registrada por Eletroencefalografia (EEG) que produz sinais com boa resolução temporal através de equipamentos com baixo custo relativo [2]. A cadeia de processamento de sinais de EEG de uma ICM é composta por etapas de pré-processamento, extração de *features* e classificação dos sinais, conforme mostra a Figura 1. No pré-processamento, os sinais de EEG são filtrados com o objetivo de reduzir a interferência de ruídos e artefatos nas bandas de frequência de interesse, em geral, associadas à imagética motora [3]. Padrões associados à IM são extraídos dos sinais de EEG filtrados e em seguida, selecionados para formar uma representação reduzida dos sinais e assim, ser entregue a algoritmos de classificação. Em geral, durante a realização da imagética motora, um sistema ICM deve ser capaz de atribuir um rótulo ao trecho de sinal de EEG associado.

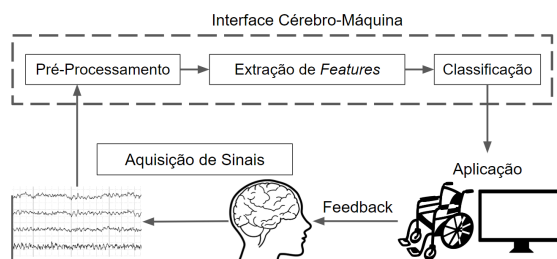


Fig. 1. Fluxograma de uma ICM-IM. Nas etapas de aquisição e pré-processamento são empregadas ferramentas que possibilitam o controle de aplicações baseadas no reconhecimento de padrões em sinais de EEG.

O algoritmo de filtragem espacial *Common Spatial Patterns* (CSP) é amplamente utilizado em sistemas ICM baseados em IM. O CSP produz filtros que combinam a informação de IM contida no domínio

espacial para ampliar a discriminação entre os sinais de EEG associados a IM diferentes.

Os filtros espaciais CSP, embora amplamente presentes em abordagens associadas a taxas de classificação elevadas dos sinais [4], são algoritmos baseados em média aritmética. Portanto, são susceptíveis a produzir filtros espaciais errôneos ao sofrerem influência de valores espúrios, presentes nos sinais de EEG coletados.

Neste trabalho é investigado uma abordagem para mitigar a influência de valores espúrios no desempenho de sistemas ICM. A abordagem consiste em reamostrar aleatoriamente com reposição os sinais de EEG disponíveis e assim, calcular, para cada reamostragem, os filtros espaciais CSP, formando um banco contendo os filtros espaciais. Esta abordagem produz uma versão aumentada da representação dos padrões, que se não for tratada adequadamente pode gerar ajuste excessivo do algoritmo de classificação aos sinais disponíveis.

Um tratamento adequado consiste em selecionar um subconjunto mais discriminativo da representação dos padrões para ser entregue ao algoritmo de classificação. Os algoritmos *Least Absolute Shrinkage and Selection Operator* (LASSO) [5] e *SelectKBest* [6] são utilizados para realizar a seleção da representação dos padrões mais discriminativos dos sinais de EEG para classificação.

O artigo está organizado da seguinte forma: a Seção I apresenta a introdução e contextualização do trabalho. Na Seção II são apresentados os conceitos básicos fundamentais para a compreensão das etapas de processamento, assim como o ferramental teórico que compõem sistemas de interfaces cérebro-máquina baseadas em imagética motora. Na Seção III são descritos os modelos desenvolvidos nesta pesquisa. A Seção IV apresenta os resultados do trabalho desenvolvido. E, por fim, a Seção V apresenta as conclusões do estudo proposto.

## II. FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os conceitos básicos das etapas e algoritmos que compõem uma ICM baseada em imagética motora por sub-banda. Primeiramente, são descritas as etapas de pré-processamento dos sinais EEG. Por fim, são apontados os algoritmos aplicados na etapa de classificação.

### A. Filtragem Temporal

A faixa de frequência em que os fenômenos que caracterizam a IM se apresentam de forma mais distinta é limitada entre  $f_l$  e  $f_h$ , frequências de corte inferior e superior, respectivamente [7]. Uma estratégia amplamente utilizada para filtragem de sinais de EEG consiste em subdividir a faixa de frequência definida por  $[f_l, f_h]$  em faixas de frequências com bandas menores através de um banco de filtros [8].

Neste estudo, os sinais são filtrados através da Transformada Rápida de Fourier (FFT) [9], formando um banco com  $N_s$  filtros temporais. Seja  $\mathbf{X}_i^\ell \in \mathcal{R}^{N_c \times N_q}$  a  $i$ -ésima época filtrada na  $\ell$ -ésima sub-banda contendo  $N_c$  canais de EEG,  $N_q$  amostras e  $\ell \in [1, \dots, N_s]$ .

### B. Banco de Filtros Espaciais - CSP

O *Common Spatial Pattern* - CSP consiste na combinação linear de sinais de EEG no domínio espacial [4], assim, é possível reduzir a quantidade dos sinais resultantes ao selecionar apenas os filtros espaciais mais discriminativos. O algoritmo CSP é projetado para maximizar a diferença entre os padrões de sinais de EEG de duas classes distintas, definidas por  $C_\rho : \rho \in [+1, -1]$ . Os filtros CSP podem ser obtidos resolvendo:

$$(\mathbf{w}^\ell)^* = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{C}_{+1}^\ell \mathbf{w}}{\mathbf{w}^T (\mathbf{C}_{+1}^\ell + \mathbf{C}_{-1}^\ell) \mathbf{w}}, \quad (1)$$

em que  $\mathbf{C}_{+1}$  e  $\mathbf{C}_{-1}$  são as matrizes de covariância médias para as classes  $+1$  e  $-1$  e  $\mathbf{w}$  é o vetor que define os filtros espaciais CSP.

As matrizes de covariâncias médias das classes são calculadas como segue,

$$\mathbf{C}_\rho^\ell = \frac{1}{N_\rho N_q} \sum_{i \in C_\rho} \mathbf{X}_i^\ell (\mathbf{X}_i^\ell)^T \quad (2)$$

em que  $\rho \in [+1, -1]$ ,  $N_\rho$  é a quantidade de exemplos da classe  $\rho$ , tal que  $N = N_{+1} + N_{-1}$  e  $(\cdot)^T$  é a operação de transposição matricial.

A redução da dimensionalidade do conjunto de sinais ocorre ao selecionar  $N_p$  filtros espaciais mais discriminativos, em que  $N_p \leq N_c$ , conforme destacado na Eq. (3), em que  $\mathbf{W}_p^\ell$  é a matriz contendo os  $N_p$  filtros espaciais da sub-banda  $\ell$ ,

$$\mathbf{W}_{N_p}^\ell = [\mathbf{w}_1 \dots \mathbf{w}_{N_p}], \quad (3)$$

onde  $\mathbf{w}_{N_p}$  são vetores colunas que compõem a matriz  $\mathbf{W}_{N_p}$ .

O banco de filtros espaciais é formado repetindo-se  $N_r$  vezes a amostragem com reposição do conjunto de sinais de EEG disponível para ajuste dos algoritmos. Para cada amostragem, calculam-se  $N_p$  filtros espaciais para cada sub-banda  $\ell$ .

A matriz  $\bar{\mathbf{W}}_{N_p}^\ell \in \mathcal{R}^{N_c \times N_p N_r}$  corresponde a matriz aumentada contendo os filtros resultantes de  $N_r$  repetições,

$$\bar{\mathbf{W}}_{N_p}^\ell = [\mathbf{W}_1^1 \dots \mathbf{W}_{N_p}^1 \quad \dots \quad \mathbf{W}_1^{N_r} \dots \mathbf{W}_{N_p}^{N_r}] \quad (4)$$

O vetor de *features*,  $\mathbf{f}_i^\ell \in \mathcal{R}^{N_p N_r}$ , da sub-banda  $\ell$  é dado por,

$$\mathbf{f}_i^\ell = \text{diag}((\bar{\mathbf{W}}_{N_p}^\ell)^T \mathbf{X}_i^\ell (\mathbf{X}_i^\ell)^T \bar{\mathbf{W}}_{N_p}^\ell) / N_q \quad (5)$$

em que  $\text{diag}(\cdot)$  é o operador que obtém os elementos da diagonal principal do argumento.

O vetor de *features* aumentado,  $\bar{\mathbf{f}}_i \in \mathcal{R}^{N_s N_p N_r}$ , é dado por,

$$\bar{\mathbf{f}}_i = [(\mathbf{f}_i^1)^T \dots (\mathbf{f}_i^{N_s})^T]^T \quad (6)$$

### C. Seleção de Features

Após serem aplicados a cada sub-banda, os filtros espaciais produzem um conjunto de *features* que devem ser capazes de representar cada uma das classes de IM utilizadas no experimento. A etapa de seleção de *features* visa reduzir o tamanho do vetor de *features* definido na Eq. (6) e utilizado nas etapas seguintes da ICM.

1) **LASSO**: O algoritmo *Least Absolute Shrinkage and Selection Operator* - LASSO funciona minimizando a função custo somada com uma penalidade L1 (norma L1) das estimativas dos coeficientes de cada variável [5]. O objetivo é encontrar um conjunto de coeficientes que minimizem a função custo, mas que também tenha um número limitado de variáveis com coeficientes diferentes de zero. A função custo é definida na Eq. (7),

$$\arg \min_{\beta, \beta_0} \left( \frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - \bar{\mathbf{f}}_i^T \boldsymbol{\beta})^2 + \alpha \sum_{j=1}^{N_s N_p N_r} |\beta_j| \right) \quad (7)$$

onde  $\bar{\mathbf{f}}_i$  é o vetor de *features*,  $y_i \in [-1, +1]$  e  $\alpha$  é o hiperparâmetro de suavização e define a posição e o número de *features* selecionadas de  $\bar{\mathbf{f}}_i$ . Assim,

as *features* mais significativas são selecionadas das sub-bandas de frequência consideradas e o vetor de *features*  $\hat{\mathbf{f}}_i \in \mathcal{R}^{N_e \times 1}$  segue para as próximas etapas.

2) **SelectKBest**: O algoritmo *SelectKBest* seleciona as melhores  $K$  *features* com base em uma pontuação predefinida. Primeiramente é selecionada uma medida estatística adequada para avaliar a relevância das *features* como, por exemplo análise de variância (ANOVA). Em seguida, com base na medida estatística escolhida, o método atribui uma pontuação a cada *feature*. A seleção pode ser feita escolhendo-se as  $K$  *features* com as pontuações mais altas ou aplicando um limiar para selecionar as *features* acima de um determinado valor de pontuação [10].

### D. Classificação

O classificador *Support Vector Machine* (SVM) é amplamente aplicado no desenvolvimento de ICMS [11]. O SVM determina o hiperplano que melhor separa as diferentes classes representadas pelos vetores de *features*. O hiperplano é determinado de forma que a margem de separação entre as classes seja maximizada. O SVM utiliza vetores de suporte, que são os exemplos mais próximos do hiperplano de decisão, para separar as diferentes classes. Na Eq. (8) tem-se a função custo do SVM,

$$\min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i, \quad (8)$$

sujeito à  $y_i (\omega^T \phi(\hat{\mathbf{f}}_i) + b) \geq 1 - \xi_i$ ,  $\xi_i \geq 0$ ,  $i = (1, \dots, N)$ . Em que  $\hat{\mathbf{f}}_i$  é o vetor de *features*,  $\phi(\hat{\mathbf{f}}_i)$  mapeia  $\hat{\mathbf{f}}_i$  para um espaço de maior dimensionalidade,  $C$  é o hiperparâmetro de regularização do SVM,  $y_i \in [-1, +1]$  e  $\xi$  é a variável de folga.

## III. METODOLOGIA

### A. Dataset IV2a - Competição de ICM

Os dados utilizados no estudo estão disponíveis no site do IV competição ICM [12]<sup>1</sup>. Os dados contêm sinais de EEG de 9 sujeitos, contendo para cada um  $N_c = 22$  canais, amostrados a uma taxa de  $F_s = 250$  Hz. Para cada sujeito estão disponíveis 288 tentativas de imaginação motora para cada uma das 4 classes registradas: mão esquerda, mão direita, pés e língua. Cada tentativa tem duração de 4

<sup>1</sup><https://www.bbci.de/competition/iv/dataset2a>

segundos. Neste trabalho são utilizadas 2 classes: a mão esquerda e a mão direita. O posicionamento dos eletrodos no escalpo segue o Sistema Internacional de Posicionamento 10 – 20 [12].

Os dados são divididos em dois conjuntos, um utilizado para realizar o ajuste dos modelos, definido como conjunto de treinamento, e outro conjunto utilizado para avaliar os modelos ajustados, definido como conjunto de teste, ambos contendo  $N = 144$  tentativas, sendo  $N_p = 72$  para cada classe de imagética motora. Os sinais de EEG são analisados em trechos com duração de 2 segundos, perfazendo  $N_q = 500$  amostras.

### B. Abordagem de referência - Baseline

Na estrutura da abordagem de referência, a filtragem temporal é realizada por  $N_s = 8$  sub-bandas. Para cada sub-banda são selecionados  $N_p = 8$  filtros espaciais, formando o vetor de *features* com dimensão de  $N_s \times N_p = 64$ , definido na Eq. (6), com  $N_r = 1$ . O algoritmo SVM, com *Kernel* linear e  $C = 100$ , é utilizado. As sub-bandas utilizadas são:  $\{[4, 12], [8, 16], [12, 20], [16, 24], [20, 28], [24, 32], [28, 36], [32, 40]\}$ .

### C. Abordagem Proposta

A abordagem proposta, ilustrada na Figura 2, consiste em aumentar a dimensão do vetor de *features*, definido na Eq. (6). O aumento da dimensão do vetor de *features* é realizado ao se amostrar aleatoriamente exemplos do conjunto de treinamento. A cada amostragem é coletado 50% dos exemplos de forma balanceada entre as classes. Repete-se o procedimento de amostragem  $N_r = 6$  vezes, para cada uma das  $N_s = 8$  sub-bandas. Os filtros resultantes formam o banco de filtros definido na Eq. (4). O vetor de *features* é formado, conforme Eq. (5) e, com a configuração utilizada, contém ao todo 384 *features*.

Na sequência, com o vetor de *features*, os algoritmos de seleção de *features*, LASSO e *SelectKBest*, são utilizados para reduzir a dimensão do vetor de *features* mais discriminativas. Os parâmetros  $\alpha$  e  $K$  dos algoritmos citados são obtidos através de um *grid search* [13] utilizando o conjunto de treinamento. A métrica de avaliação escolhida para o *SelectKBest* é a ANOVA. Após selecionadas, as *features* mais discriminativas ajustam o algoritmo SVM que é utilizado para a classificação do conjunto de avaliação.

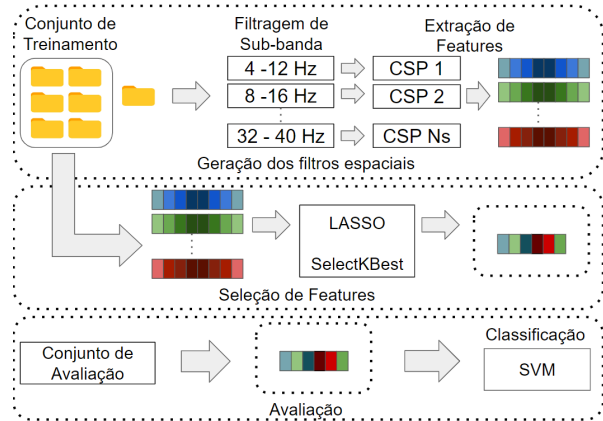


Fig. 2. Configuração proposta da ICM-IM utilizando a geração de filtros espaciais variados.

## IV. RESULTADOS

Nesta Seção são apresentados os resultados dos experimentos para a abordagem proposta e de referência descritas nas Seções III-C, Seção III-B, respectivamente.

Na Tabela I são apresentados os resultados, em termos de acurácia, para cada um dos 9 sujeitos do *dataset* para as abordagens *baseline* e a proposta. Na abordagem proposta os algoritmos de seleção de *features* LASSO e *SelectKBest* (SKB) estão em destaque. Outra informação disponível na Tabela I, são os valores de  $\alpha$  utilizados pelo algoritmo LASSO para cada sujeito, assim como o número de *features* utilizadas pelos algoritmos LASSO e *SelectKBest*, respectivamente. Ao analisar a Tabela I, no quesito número de *features*, as quantidades selecionadas para compor o modelo variam de sujeito para sujeito em ambos os algoritmos. Em especial, para os sujeitos 1, 3, 7 e 8, embora o valor de  $\alpha$  do LASSO seja o mesmo, as quantidades de *features* selecionadas são diferentes. Essa observação corrobora com a literatura, em que destaca-se a especificidade da imagética motora para cada indivíduo.

Ainda na Tabela I, no quesito Acurácia, os resultados apontam que para ambos os algoritmos de seleção de *features* avaliados, a abordagem proposta produz, em média, desempenho de classificação  $\approx 2\%$  superior em comparação à abordagem *baseline*. Por outro lado, quando a comparação é entre os algoritmos de seleção de *features*, observa-se um desempenho equivalente, 80,71% e 80,56%, respectivamente, para o LASSO e o *SelectKBest*.

Quanto a quantidade de *features* selecionadas, o algoritmo LASSO seleciona, para a maioria dos sujeitos, uma quantidade menor de *features* do que o *SelectKBest*. O que não implica, necessariamente, em desempenho de classificação superior, como pode ser observado para os sujeitos 2, 5, 6 e 8.

No geral, observa-se que os resultados para a abordagem proposta produzem resultados globais equivalentes ainda que os resultados individuais por sujeito sejam diferentes, o que sugere uma dificuldade em se estabelecer uma única abordagem capaz de produzir os melhores resultados para todos os sujeitos.

TABELA I

DESEMPENHO DE CLASSIFICAÇÃO DOS MODELOS AVALIADOS.

S	Alpha $\alpha \cdot 10^3$	Nº de Features (Total = 384)		Acurácia		
		LASSO	SKB	LASSO	SKB	Baseline
1	25	38	114	87,5	86,11	<b>88,89</b>
2	5	102	215	57,64	<b>58,33</b>	52,08
3	25	26	50	<b>92,36</b>	90,97	82,64
4	1	138	204	<b>68,75</b>	63,19	66,67
5	10	65	84	<b>85,42</b>	<b>85,42</b>	77,08
6	10	84	277	59,72	<b>69,44</b>	65,97
7	25	45	122	<b>91,67</b>	86,11	89,58
8	25	27	231	93,75	<b>97,22</b>	93,75
9	10	37	225	<b>89,58</b>	88,22	86,81
Acurácia Média				<b>80,71</b>	80,56	78,16
Desvio Padrão				$\pm$ <b>13,70</b>	$\pm$ 12,69	$\pm$ 13,10

## V. CONCLUSÃO

O trabalho apresenta uma abordagem alternativa para realizar a sintonia de sistemas ICM-IM. A abordagem propõe a formação de um banco de filtros espaciais via CSP e produz uma versão aumentada da representação de padrões dos sinais de EEG. A versão aumentada dos padrões é então reduzida por seleção via algoritmos LASSO e *SelectKBest*. A abordagem é aplicada em uma base de dados pública de competições de ICM-IM. Os resultados indicam que o banco de filtros espaciais via CSP aumentam a possibilidade de selecionar as *features* significativas da imagética motora para cada sujeito, ao passo que atenua os ruídos ocasionados por dados espúrios presentes nos sinais de EEG. Os algoritmos LASSO e *SelectKBest* apresentaram desempenho equivalente para os testes realizados. Ainda que o LASSO se destaque por selecionar uma quantidade menor de *features* em relação ao *SelectKBest*. Ademais, em termos

de acurácia, nenhuma das abordagens apresentadas resultaram em desempenho superior para todos os sujeitos avaliados. O que evidencia a ausência de uma abordagem definitiva capaz de garantir os melhores resultados para todos os sujeitos.

## REFERÊNCIAS

- [1] K.-T. Kim, T. Carlson, and S.-W. Lee, "Design of a robotic wheelchair with a motor imagery based brain-computer interface," in *2013 International Winter Workshop on Brain-Computer Interface (BCI)*, pp. 46–48, 2013.
- [2] J. S. Kumar and P. Bhuvaneshwari, "Analysis of electroencephalography (eeg) signals and its categorization—a study," *Procedia Engineering*, vol. 38, pp. 2525–2536, 2012. INTERNATIONAL CONFERENCE ON MODELLING OPTIMIZATION AND COMPUTING.
- [3] L. Zheng, W. Feng, Y. Ma, P. Lian, Y. Xiao, Z. Yi, and X. Wu, "Ensemble learning method based on temporal, spatial features with multi-scale filter banks for motor imagery eeg classification," *Biomedical Signal Processing and Control*, vol. 76, p. 103634, 2022.
- [4] S. Zhang, Z. Zhu, B. Zhang, B. Feng, T. Yu, Z. Li, Z. Zhang, G. Huang, and Z. Liang, "Overall optimization of csp based on ensemble learning for motor imagery eeg decoding," *Biomedical Signal Processing and Control*, vol. 77, p. 103825, 2022.
- [5] Y. Miao, J. Jin, I. Daly, C. Zuo, X. Wang, A. Cichocki, and T.-P. Jung, "Learning common time-frequency-spatial patterns for motor imagery classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 699–707, 2021.
- [6] G. Dueñas, S. Jimenez, and J. Baquero, "Automatic prediction of item difficulty for short-answer questions," in *2015 10th Computing Colombian Conference (10CCC)*, pp. 478–485, 2015.
- [7] Z. Tang, H. Yu, C. Lu, P. Liu, and X. Jin, "Single-trial classification of different movements on one arm based on erd/ers and corticomuscular coherence," *IEEE Access*, vol. 7, pp. 128185–128197, 2019.
- [8] Q. Novi, C. Guan, T. Dat, and P. Xue, "Sub-band common spatial pattern (sbcs) for brain-computer interface," pp. 204 – 207, 06 2007.
- [9] C. Silva, R. Duarte, and A. Trofino, "Feature extraction improvements using an lmi approach and riemannian geometry tools: An application to bci," in *2016 IEEE Conference on Control Applications (CCA)*, pp. 966–971, 2016.
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [11] F. Lotte, "A tutorial on EEG signal-processing techniques for mental state recognition in brain-computer interfaces," in *Guide to Brain-Computer Music Interfacing*, no. October 2014, pp. 133–161, Springer, London, 2014.
- [12] C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI competition 2008 – graz data set a," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces)*, vol. 16, 2008.
- [13] X. Deng, B. Zhang, N. Yu, K. Liu, and K. Sun, "Advanced ts-glue-net for motor imagery eeg-based brain-computer interfaces," *IEEE Access*, vol. 9, pp. 25118–25130, 2021.

### **3 COMBINAÇÃO DE MODELOS PERSONALIZADOS DE INTERFACES CÉREBRO-MÁQUINA BASEADAS EM IMAGÉTICA MOTORA COM OTIMIZAÇÃO BAYESIANA**

Na próxima página encontra-se o arquivo original do artigo intitulado “Combinação de Modelos Personalizados de Interfaces Cérebro-Máquina Baseadas em Imagética Motora com Otimização Bayesiana”, apresentado na XVII edição do Congresso Brasileiro de Inteligência Computacional – CBIC 2025, realizado em formato presencial no ano de 2025.

# Combinação de Modelos Personalizados de Interfaces Cérebro-Máquina Baseadas em Imagética Motora com Otimização Bayesiana

Guilherme da Costa Dourado  
Faculdade de Engenharia Elétrica  
Universidade Federal do Pará (UFPA)  
Tucuruí, Brasil  
dourado.gui19@gmail.com

Cleison Daniel Silva  
PPG em Computação Aplicada  
Universidade Federal do Pará (UFPA)  
Tucuruí, Brasil  
cleison@ufpa.br

Vitor da Silva Jorge  
PPG em Computação Aplicada  
Universidade Federal do Pará (UFPA)  
Tucuruí, Pará  
vitor.dsilvajorge@gmail.com

**Resumo**—Este trabalho investiga a combinação de modelos personalizados de Interfaces Cérebro-Máquina baseadas em Imagética Motora (ICM-IM), otimizados via otimização Bayesiana e combinados por *Bagging*. Foram gerados modelos para classificação binária de seis pares de classes a partir do conjunto de dados público 2a da Competição IV de ICM. As predições dos três melhores modelos de cada par foram combinadas por *Soft Voting*. A performance foi avaliada com o índice  $\kappa$  e sua estatística  $z_\kappa$ . Embora a combinação tenha elevado os valores médios de  $\kappa$ , os ganhos não foram estatisticamente significativos, exigindo ajustes metodológicos para melhor desempenho.

**Palavras-Chave**—Interfaces Cérebro-Máquina, Imagética Motora, Otimização Bayesiana, Aprendizado em Conjunto, *Bagging*, *Soft Voting*.

**Abstract**—This study investigates the combination of personalized Motor Imagery-based Brain-Computer Interfaces (MI-BCI) models, optimized through Bayesian optimization and combined using *Bagging*. Binary classifiers were built for six class pairs from the public dataset 2a of the BCI Competition IV. The predictions of the top three models for each pair were combined using *Soft Voting*. Performance was evaluated using the  $\kappa$  index and its  $z_\kappa$  statistic. Although the ensemble increased average  $\kappa$  values across subjects and class pairs, the improvements were not statistically significant, indicating that methodological adjustments are needed to achieve better performance in binary MI-BCI classification tasks.

**Keywords**—Brain-Computer Interfaces, Motor Imagery, Bayesian Optimization, Ensemble Learning, *Bagging*, *Soft Voting*.

## I. INTRODUÇÃO

Interfaces Cérebro-Máquina (ICMs) baseadas em Imagética Motora (IM) traduzem sinais cerebrais, como os captados por eletroencefalografia (EEG), em comandos para dispositivos externos, sendo úteis em aplicações como mobilidade assistida [1], [2], reabilitação e treinamento em contextos clínicos e educacionais [3]. A imaginação de movimentos, como os de mãos ou pés, gera padrões específicos nos sinais de EEG, que podem ser utilizados para treinar modelos de classificação. Contudo, a baixa amplitude dos sinais, sua suscetibilidade a ruídos e a alta variabilidade interindividual tornam a tarefa de classificação desafiadora [1].

O processamento de uma ICM-IM envolve várias etapas fundamentais, como a extração de épocas, pré-processamento, extração de características e classificação. ICMs personalizadas podem empregar técnicas como os Padrões Espaciais Comuns (*Common Spatial Pattern*, CSP) e Análise Discriminante Linear (*Linear Discriminant Analysis*, LDA) para extração e seleção de características [4]. Quando ajustadas individualmente por métodos como a otimização Bayesiana (*Bayesian optimization*, BO), esses modelos demonstram desempenho satisfatório ao se adaptar às particularidades de cada usuário [4], [5]. No entanto, sua generalização ainda é limitada, o que motiva a investigação de abordagens que combinem os pontos fortes desses modelos personalizados.

O aprendizado em conjunto (*ensemble learning*) surge como alternativa promissora ao explorar múltiplos classificadores para aumentar a robustez e reduzir a variância dos modelos. Métodos de combinação de modelos ICMs como *bagging*, *boosting* e *stacking* e estratégias de decisão final de classificação como votação majoritária e *soft voting* têm sido amplamente estudadas nesse contexto [6]–[9]. Em especial, o *bagging* permite combinar modelos ajustados a partir de subconjuntos amostrados aleatoriamente, promovendo diversidade entre classificadores e mitigando o *overfitting* [6], [7].

Este trabalho tem como objetivo investigar se a combinação de ICMs personalizadas, por meio da técnica de *bagging*, pode melhorar significativamente o desempenho da classificação de IM em uma abordagem binária. São avaliados seis pares de classes motoras extraídas de um conjunto de dados público, com os modelos individuais ajustados por BO. A decisão final é obtida por meio de *soft voting*, que considera as probabilidades preditas por cada classificador. O desempenho é analisado por meio do índice *kappa* ( $\kappa$ ) e do índice estatístico  $z_\kappa$ , oferecendo uma avaliação mais robusta do que métricas tradicionais como acurácia [10].

Este artigo está organizado da seguinte forma: a Seção I apresenta a introdução e o contexto do trabalho; a Seção II discute os fundamentos e as etapas que compõem a cadeia de processamento de uma ICM-IM, além das estratégias de combinação de modelos; a Seção III detalha a metodologia

utilizada em cada etapa; a Seção IV apresenta os resultados obtidos; e a Seção V traz as conclusões do estudo.

## II. FUNDAMENTAÇÃO TEÓRICA

Na Seção II, são apresentados os principais conceitos e etapas que compõem uma ICM-IM. Serão abordados os procedimentos de extração de épocas, pré-processamento dos sinais de EEG, extração e seleção de características, classificação dos padrões neurais, bem como estratégias complementares como otimização Bayesiana, aprendizado em conjunto e métricas utilizadas para avaliação do desempenho do sistema.

### A. Extração de épocas

Após a aquisição dos sinais, a primeira etapa de processamento de uma ICM-IM consiste na extração de épocas, conforme ilustrado na Fig. 1. A extração de épocas consiste em segmentar o sinal de EEG em janelas temporais fixas que capturam a atividade cerebral relacionada à imaginação motora. Esse processo ocorre durante a fase de treinamento da ICM, sendo essencial para calibrar os algoritmos de classificação. Protocolos experimentais, como o *Graz* [11], utilizam estímulos visuais para indicar ao usuário qual movimento imaginar, permitindo rotular os sinais com precisão. As épocas são geralmente extraídas entre 0,5 e 2,5 segundos após o estímulo, período em que os sinais relacionados à intenção motora estão mais evidentes.

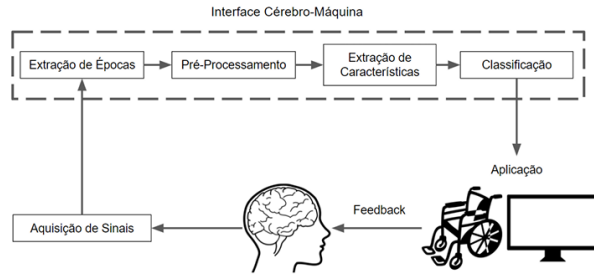


Fig. 1. Cadeia de processamento de uma ICM tradicional.

### B. Pré-Processamento

Devido à sua baixa amplitude, os sinais de EEG são altamente suscetíveis a ruídos, como mencionado anteriormente, o que exige a aplicação de técnicas de pré-processamento para melhorar a qualidade dos dados e a eficiência das ICMs. Essa etapa envolve a aplicação de filtros temporais e espaciais com o objetivo de realçar informações relevantes e suprimir interferências.

Uma abordagem comum na literatura atual, é a utilização de um banco de filtros em sub-bandas (*Sub-band Common Spatial Pattern*, SBCSP) para realizar a filtragem temporal e espacial dos sinais. Essa configuração é adotada com base em estudos que indicam as sub-bandas como a estratégia mais eficaz para observar fenômenos característicos da imagética motora, que podem se manifestar de maneira distinta no espectro de frequência de cada indivíduo [12].

O banco de filtros tem como finalidade segmentar o sinal em diversas sub-bandas, que em conjunto cobrem as faixas de frequência dos ritmos  $\mu$  (8 – 12 Hz) e  $\beta$  (12 – 30 Hz). Em seguida, é aplicado um filtro espacial em cada sub-banda e procede-se à seleção das componentes que melhor representam as intenções motoras. A decisão final do sistema é baseada na fusão dos escores derivados de cada sub-banda analisada.

A filtragem temporal é aplicada para restringir a análise do sinal a frequências relevantes à imagética motora. Normalmente, são utilizados filtros passa-faixa que isolam os ritmos  $\mu$  e  $\beta$  [13]. Para representar os dados no domínio da frequência, emprega-se a Transformada Rápida de Fourier (*Fast Fourier Transform*, FFT), convertendo o sinal bruto em componentes espectrais que refletem sua composição harmônica.

Já a filtragem espacial tem como função acentuar as informações discriminativas entre os canais, reduzindo a redundância espacial e maximizando a separação entre classes motoras. O algoritmo Padrões Espaciais Comuns é uma das técnicas mais utilizadas nessa etapa [1]. O CSP busca encontrar um conjunto de filtros espaciais que maximize a variância de uma classe motora (ex.: mão esquerda) enquanto minimiza a da outra (ex.: mão direita), como descrito em [14]. Formalmente, para duas classes com matrizes de covariância média  $C_1$  e  $C_2$ , um filtro espacial é obtido extremado a seguinte razão em (1):

$$\operatorname{argmax}_{\mathbf{w}} = \frac{\mathbf{w}^T C_1 \mathbf{w}}{\mathbf{w}^T (C_1 + C_2) \mathbf{w}}. \quad (1)$$

A solução dessa equação leva à formulação de um problema de autovalores generalizado, visto em (2):

$$C_1 \mathbf{w} = \lambda (C_1 + C_2) \mathbf{w}, \quad (2)$$

onde  $\lambda$  e  $\mathbf{w}$  são os autovalores e autovetores, respectivamente. Como descrito em (3), os vetores  $\mathbf{w}$  correspondem aos filtros espaciais, e a aplicação desses filtros aos sinais originais  $\mathbf{X}$  gera um novo conjunto de dados filtrados espacialmente  $\mathbf{R}$ :

$$\mathbf{R} = \mathbf{W}^T \mathbf{X}, \quad (3)$$

em que  $\mathbf{W}$  é a matriz de transformação cujas colunas são filtros espaciais. Somente os filtros associados aos maiores e menores autovalores são selecionados, pois são os que maximizam a separabilidade entre as intenções motoras. Esses componentes formam um vetor de características de dimensão reduzida, pronto para a etapa subsequente de extração e seleção de características.

### C. Extração e Seleção de Características

Após o pré-processamento, a extração de características visa transformar os sinais de EEG em representações mais eficientes das informações relevantes para o experimento [15]. Essa etapa é fundamental para traduzir a atividade cerebral do usuário em comandos úteis para a ICM.

Na abordagem por sub-bandas, as características são extraídas separadamente de cada faixa de frequência, e cada sub-banda é pontuada conforme sua relevância. Para isso,

aplica-se a LDA como meta-classificador, que projeta os dados maximizando a razão entre variância entre classes e dentro das classes.

Devido à presença de ruídos e à variabilidade interindividual dos sinais EEG, apenas algumas sub-bandas contribuem significativamente para a classificação. Assim, uma etapa de seleção de bandas é aplicada para manter apenas as mais informativas. Por fim, os vetores extraídos são fundidos em uma única representação, que pode ser avaliada por métodos como classificadores Bayesianos, utilizando a razão de verossimilhança para gerar uma pontuação final a ser entregue ao classificador decisivo.

#### D. Classificação

A etapa de classificação visa atribuir as amostras extraídas e processadas às suas respectivas classes motoras. Diversos algoritmos são utilizados, sendo os mais comuns a Análise Discriminante Linear (LDA), Regressão Logística (*Logistic Regression*, LR), Máquina de Vetores de Suporte (*Support Vector Machine*, SVM), Perceptron Multicamadas (*Multi-layer Perceptron*, MLP) e o algoritmo dos K-Vizinhos Mais Próximos (*K-Nearest Neighbors*, KNN).

- A LDA é um algoritmo de classificação que busca separar classes com distribuições Gaussianas e variâncias iguais, além de ser uma técnica de redução de dimensionalidade. Ela utiliza um hiperplano de separação e maximiza a distância entre as médias das classes, minimizando a variância interclasses [16].
- A LR é um algoritmo de aprendizado supervisionado voltado à classificação, especialmente binária, ao prever a probabilidade de um evento com base em variáveis de entrada por meio da função sigmoide [17].
- O SVM, assim como a LDA, utiliza um hiperplano discriminante para separar classes [16], mas seu objetivo é maximizar a margem entre os dados de diferentes classes, aumentando a confiança na classificação.
- O MLP é formado por camadas de unidades inspiradas em neurônios, capazes de aprender operações simples [18]. Suas camadas ocultas e de saída aplicam funções de ativação como sigmoide, tangente hiperbólica (*tanh*), *Rectified Linear Unit* (ReLU) e *softmax*, que introduzem não linearidade e permitem a aprendizagem de padrões complexos [19].
- O classificador KNN determina a classe de uma nova amostra com base na classe da maioria de seus K vizinhos mais próximos, definidos pela menor distância entre uma nova amostra e um conjunto de amostras de treino [20].

#### E. Otimização Bayesiana

A otimização Bayesiana é uma estratégia para ajuste automático de hiperparâmetros baseada em modelos probabilísticos. A busca por esses valores ótimos não é trivial e pode consumir vastos recursos computacionais e tempo, especialmente quando uma estrutura de ICM é ajustada de forma personalizada para cada sujeito. É nesse contexto que a utilização de algoritmos de otimização se torna essencial [5].

Essa técnica é especialmente útil no ajuste de classificadores em ICMs personalizadas, pois reduz significativamente o tempo de busca por configurações eficazes e melhora a adaptabilidade entre diferentes sujeitos. A estratégia é fundamentada no Teorema de Bayes, que calcula a probabilidade posterior de um modelo  $P(M/O)$  a partir de sua verossimilhança  $P(O/M)$ , da probabilidade a priori  $P(M)$  e da probabilidade marginal das observações  $P(D)$ , conforme mostra (4):

$$P(M/O) = \frac{P(O/M) \cdot P(M)}{P(O)}. \quad (4)$$

#### F. Aprendizado em Conjunto

O aprendizado em conjunto visa melhorar o desempenho de classificação ao combinar múltiplos modelos, explorando as vantagens individuais de cada classificador. Em ICMs, essa abordagem permite integrar diferentes estruturas ajustadas com hiperparâmetros distintos, o que contribui para maior robustez e menor taxa de erro [18]. As combinações podem ocorrer em várias etapas do processamento e são especialmente eficazes diante da variabilidade dos sinais de EEG. A seguir, são apresentados os algoritmos de aprendizado em conjunto utilizados neste estudo.

1) *Bagging*: O *Bagging* é uma técnica de aprendizado em conjunto que consiste em gerar múltiplos subconjuntos de dados por meio de amostragem com ou sem reposição, sobre os quais diferentes modelos são treinados [6]. As predições desses modelos são então combinadas, geralmente por votação majoritária ou *Soft Voting*. A Fig. 2 apresenta a estratégia.

Essa abordagem é especialmente eficaz para algoritmos suscetíveis ao *overfitting*, pois promove a diversidade entre os modelos e melhora a capacidade de generalização. No contexto de ICMs, o *Bagging* permite treinar classificadores com diferentes subconjuntos dos sinais de EEG, capturando variações relevantes e gerando uma classificação final mais robusta e precisa.

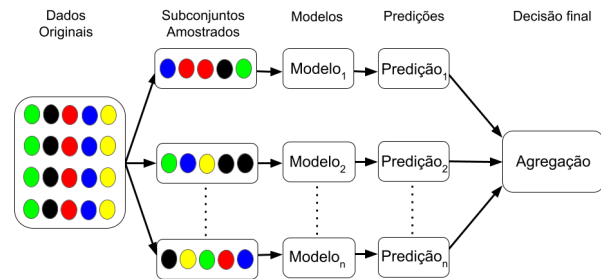


Fig. 2. Etapas do método de aprendizado em conjunto *Bagging*.

2) *Soft Voting*: O *Soft Voting* é uma técnica de combinação de classificadores em que cada modelo fornece uma probabilidade associada a cada classe. As probabilidades são então agregadas por média, e a classe com maior valor médio é selecionada como a predição final [9]. Também é possível aplicar pesos aos modelos, atribuindo maior influência às predições dos classificadores mais confiáveis.

G. Métricas de Avaliação

Para avaliar o desempenho dos modelos propostos na classificação da imagética motora, esta pesquisa adota o índice *Kappa*, que avalia a concordância entre as predições de um classificador e os rótulos reais, sendo também útil para verificar se o desempenho observado é estatisticamente superior ao que seria esperado por mero acaso [10]. O índice *Kappa* é definido por (5):

$$\kappa = \frac{P_o - P_e}{1 - P_e}, \quad (5)$$

onde  $P_o$  representa a proporção de concordância observada, ou seja, a frequência com que os classificadores produzem a mesma predição. O indicador  $P_e$  corresponde à concordância esperada, considerando que as classificações coincidam apenas por acaso. O índice *Kappa* varia de  $-1$  a  $1$ , onde  $\kappa = 1$  indica concordância perfeita,  $\kappa = 0$  reflete concordância ao acaso e  $\kappa < 0$  sugere desacordo maior do que o esperado por acaso, indicando possível inconsistência nas predições.

Além de quantificar a precisão do modelo, o *Kappa* permite testes estatísticos de significância entre diferentes classificadores. Apresentada em (6), a estatística  $z_\kappa$ , obtida a partir da diferença entre dois coeficientes  $\kappa_1$  e  $\kappa_2$  e de suas variâncias  $\sigma_{\kappa_1}^2$  e  $\sigma_{\kappa_2}^2$ , segue uma distribuição normal padrão:

$$z_\kappa = \frac{|k_1 - k_2|}{\sqrt{\sigma_{\kappa_1}^2 - \sigma_{\kappa_2}^2}}, \quad (6)$$

com isso, é possível calcular o p-valor, que indica se a diferença entre os desempenhos é estatisticamente significativa. Para um nível de significância  $\alpha = 0,05$ , valores de  $z_\kappa > 1,96$  indicam significância estatística com 95% de confiança. Caso contrário, não há diferença significativa entre os classificadores.

III. METODOLOGIA

A. Ambiente Experimental e Conjunto de Dados

Os experimentos desta pesquisa foram realizados na plataforma *Google Colaboratory*, que permite o uso de *notebooks Jupyter* em nuvem com recursos computacionais escaláveis. Essa infraestrutura facilita o desenvolvimento, a reprodutibilidade e o compartilhamento das análises. O código foi implementado em *Python*, utilizando as bibliotecas *scikit-learn*, voltada ao aprendizado de máquina [21], e *MNE-Python*, especializada na manipulação e visualização de sinais EEG [22]. Essa configuração viabilizou a execução de todas as etapas do pipeline da ICM-IM, desde o pré-processamento até a classificação.

O conjunto de dados utilizado foi o *dataset 2a* da Competição IV de ICM, uma base pública amplamente reconhecida na área [11]. Ele é disponibilizado em dois subconjuntos, cada um correspondente a uma sessão de coleta. O *dataset* contém sinais de EEG de nove sujeitos em tarefas de imagética motora com quatro classes: mão esquerda (ME), mão direita (MD), pés (PE) e língua (LG), totalizando 288 tentativas por sujeito em cada sessão. Neste estudo, as quatro classes

foram combinadas em pares binários segundo a estratégia um contra um (*one-vs-one*), gerando seis combinações distintas. Os sinais foram registrados por meio de 22 canais com taxa de amostragem de 250 Hz. Cada tentativa, com duração de 4 segundos, foi segmentada em épocas de 2 segundos, produzindo 500 amostras por segmento.

B. Ajuste e Seleção de Modelos ICMs Personalizados

Para cada sujeito da base de dados, os modelos de ICM-IM foram ajustados individualmente com o objetivo de maximizar a acurácia de classificação. Cada etapa da cadeia de processamento definida em II possui múltiplos algoritmos e hiperparâmetros. A Fig. 3 ilustra o processo de ajuste da ICM por etapa, com destaque para os hiperparâmetros considerados. Neste trabalho, esse ajuste foi realizado por meio da plataforma *AutoBCI*, desenvolvida por [4], que automatiza a seleção por meio de otimização Bayesiana. Essa abordagem explora eficientemente o espaço de busca, reduzindo o custo computacional e adaptando-se às variações interindividuais.

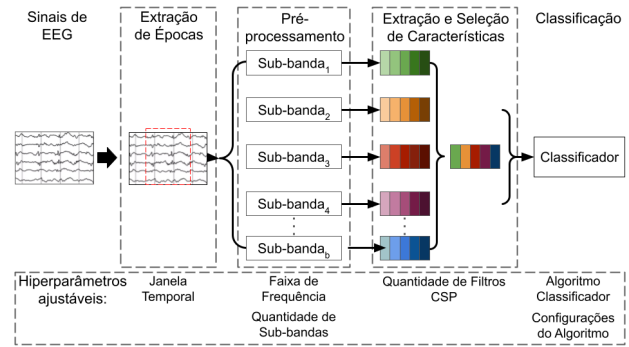


Fig. 3. Ajuste de uma interface cérebro-máquina de sub-bandas.

O espaço de busca da otimização Bayesiana é apresentado na Tabela I. A largura da janela ( $t_{max} - t_{min}$ ) foi fixada em 2 segundos, enquanto sua posição temporal variou. A banda de filtragem ( $f_h - f_l$ ) abrangeu frequências associadas à imaginação motora, com sub-bandas configuradas em janelas de 4 Hz e sobreposição de 50%, respeitando os limites da banda total. O número de componentes CSP ( $n_{csp}$ ) foi limitado ao total de eletrodos disponíveis. Dentre os classificadores testados, cujos hiperparâmetros também foram incluídos no processo de ajuste, a LDA e a LR utilizaram modelos padrões da biblioteca *scikit-learn*.

TABELA I  
HIPERPARÂMETROS CONFIGURÁVEIS EM CADA ETAPA DA ICM.

Hiperparâmetro	Variável	Domínio	Etapa da ICM
Janela temporal	$t_{min}$ $t_{max}$	[0, 0.5, ..., 2.5, 3] [1, 1.5, ..., 3.5, 4]	Extração de épocas
Faixa de frequência	$f_l$ $f_h$	[0.5, 1, ..., 29, 30] [8, 9, ..., 48, 49]	Pré-processamento
Qtd. de sub-bandas	$n_{band}$	[1, 2, ..., 23, 24]	Pré-processamento
Qtd. filtros CSPs	$n_{csp}$	[2, 4, ..., 21, 22]	Pré-processamento
Classificador	$clf$	[LDA, LR, SVM, KNN, MLP]	Classificação

O ajuste foi realizado para cada um dos seis pares de classes possíveis: {ME, MD}, {ME, PE}, {ME, LG}, {MD, PE}, {MD, LG} e {PE, LG}, conforme a estratégia um contra um. Para cada par, foram extraídas 144 épocas da primeira sessão de cada sujeito. A geração dos modelos de ICM-IM ocorreu por meio de 100 iterações da busca Bayesiana. A cada iteração, avaliou-se o desempenho de classificação médio por meio de validação cruzada  $5 \times 5$ , com subconjuntos de treino (ET) e de validação (EV) utilizados para estimar a acurácia de generalização em cada configuração.

Ao término das 100 iterações, foram selecionados os três conjuntos de hiperparâmetros com maior acurácia média durante a validação cruzada. Esse processo foi repetido para todos os pares de classes e sujeitos, totalizando 162 ajustes distintos (9 sujeitos  $\times$  6 pares  $\times$  3 modelos ótimos). Os três modelos selecionados para cada par de classes compõem a base para a aplicação das estratégias de aprendizado em conjunto descritas a seguir.

### C. Aplicação de Técnicas de Aprendizado em Conjunto

Neste trabalho, a técnica de aprendizado em conjunto adotada é o *Bagging*, aplicada ao contexto da classificação binária de sinais de imagética motora. O objetivo é combinar modelos ICM-IM personalizados, previamente otimizados, a fim de melhorar o desempenho de classificação. Para cada sujeito, são selecionados os três modelos com maior acurácia para cada par de classes, totalizando dezoito modelos. Todos utilizam uma janela temporal padronizada de dois segundos (s), garantindo consistência no número de épocas utilizadas. Porém, com posições temporais distintas entre si (Ex.: {[0 - 2], [0, 5 - 2, 5], ..., [2 - 4]} s), gerando diversidade nos dados de ajuste.

O *Bagging* combina classificadores treinados de forma independente para reduzir a variância do modelo final. Nesta abordagem, os três melhores modelos por par de classes são combinados visando a melhor generalização na distinção entre duas classes específicas. A técnica de decisão utilizada é o *Soft Voting*, em que as saídas probabilísticas dos classificadores são agregadas por média, e a classe com maior valor médio é selecionada como predição final. Essa abordagem favorece a robustez frente a variações individuais nos sinais de EEG.

Para o treinamento dos modelos personalizados, o conjunto total de dados de cada sujeito (incluindo ambas as sessões) foi reamostrado sem reposição para formar quatro subconjuntos de mesma quantidade. Três subconjuntos foram utilizados para o treinamento e o quarto para validação. Cada modelo foi treinado com um subconjunto distinto, promovendo diversidade entre eles. A validação foi realizada com um conjunto comum, composto por amostras extraídas do intervalo de EEG [0, 5 - 2, 5 s]. As saídas dos modelos foram então combinadas por *Soft Voting*. A Fig. 4 ilustra a metodologia aplicada.

Esse processo foi repetido quatro vezes, trocando o subconjunto de validação a cada repetição. Assim, cada partição atuou uma vez como conjunto de teste, permitindo uma avaliação mais robusta. A metodologia foi aplicada separadamente para cada um dos seis pares de classes, com modelos combinados

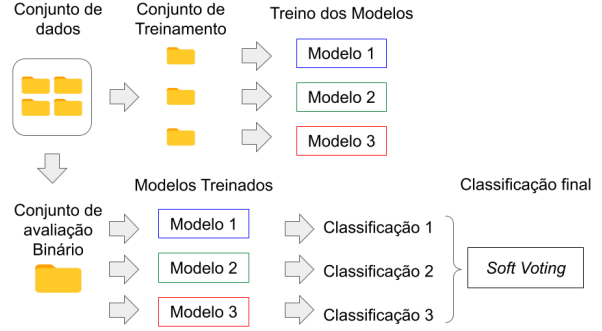


Fig. 4. Esquema de aplicação do algoritmo *Bagging* para a classificação binária.

específicos para cada caso. Espera-se que os classificadores individuais, ao errarem em instâncias distintas, se complementem, potencializando o desempenho do modelo combinado final em comparação aos modelos base.

### D. Procedimento de Avaliação dos Modelos

Na abordagem de classificação binária adotada neste trabalho, a avaliação dos modelos é realizada com base em duas métricas principais: o índice  $\kappa$  médio ( $\kappa_m$ ) e o índice estatístico  $z_\kappa$ . Essas métricas são aplicadas tanto aos modelos individuais quanto aos modelos combinados obtidos por aprendizado em conjunto, com o objetivo de verificar a confiabilidade estatística dos classificadores e a significância dos ganhos obtidos pela combinação.

O índice  $\kappa_m$  é calculado a partir das matrizes de confusão geradas nas quatro execuções da metodologia para cada par de classes e sujeito. As matrizes são somadas e divididas por quatro, gerando uma matriz média a partir da qual se calcula o coeficiente  $\kappa_m$ . Esse índice avalia a concordância entre as predições e os rótulos reais, desconsiderando os acertos esperados ao acaso, sendo especialmente útil em cenários com múltiplas execuções e variação interindividual.

A significância estatística da diferença entre os desempenhos dos modelos é verificada por meio do índice  $z_\kappa$ , conforme detalhado na Seção II-G. Esse índice é calculado com base na comparação entre o melhor modelo individual e o melhor modelo combinado, considerando o  $\kappa_m$  obtido.

## IV. RESULTADOS

Na Seção IV, são apresentados os resultados obtidos tanto para os modelos individuais selecionados, conforme descrito na Seção III-B, quanto para a abordagem de aprendizado em conjunto discutida na Seção III-C.

A Tabela II apresenta os desempenhos dos modelos individuais, bem como os resultados da aplicação do algoritmo *Bagging* para a classificação binária entre os diferentes pares de classes do conjunto de dados utilizado. Os nove sujeitos avaliados, descritos na Seção III-A, estão representados na tabela pela numeração de 1 a 9. Os modelos combinados de cada sujeito, visualizados em [23] e indicados na coluna **Mod**,

TABELA II  
RESULTADOS OBTIDOS PARA OS MODELOS INDIVIDUAIS E EM CONJUNTO AVALIADOS.

Suj	Mod	ME x MD			ME x PE			ME x LG			MD x PE			MD x LG			PE x LG			
		$\kappa_m$ (%)		$z_{\kappa}$	$\kappa_m$ (%)		$z_{\kappa}$	$\kappa_m$ (%)		$z_{\kappa}$	$\kappa_m$ (%)		$z_{\kappa}$	$\kappa_m$ (%)		$z_{\kappa}$	$\kappa_m$ (%)		$z_{\kappa}$	
		Mod	SV		Mod	SV		Mod	SV		Mod	SV		Mod	SV		Mod	SV		Mod
1	M1	72,92		0,12	93,75		0	95,14		0,14	91,67		0,15	98,61		0	43,75		0,23	
	M2	44,44	<b>74,31</b>		90,28	<b>94,44</b>		93,75	<b>95,83</b>		95,83	<b>96,53</b>		98,61	<b>99,31</b>		<b>99,31</b>	36,11		<b>47,22</b>
	M3	60,42			<b>94,44</b>			90,97			93,06			<b>99,31</b>			34,72			
2	M1	5,56		0	61,11		0,43	31,94		0,6	61,11		0,83	37,5		0,5	70,83		0,06	
	M2	<b>13,19</b>	<b>13,19</b>		60,42	<b>66,67</b>		36,81	<b>45,83</b>		65,28	<b>75,69</b>		33,33	<b>45,14</b>		62,5	<b>71,53</b>		
	M3	5,56			58,33			35,42			65,97			35,42			56,94			
3	M1	<b>93,06</b>		0,11	<b>93,06</b>		0,22	92,36		0,37	92,36		0,11	93,75		0,13	50,69		0,64	
	M2	84,03	92,36		80,56	91,67		93,06	<b>95,14</b>		91,67	<b>93,06</b>		89,58	<b>95,14</b>		52,08	<b>66,67</b>		
	M3	90,28			79,86			87,5			60,42			94,44			58,33			
4	M1	34,72		0,59	50		0,77	65,97		0,34	63,19		0,6	48,61		0,48	45,14		1,17	
	M2	30,56	<b>43,75</b>		65,97	<b>75</b>		57,64	<b>70,14</b>		68,06	<b>75</b>		59,03	<b>65,28</b>		47,22	<b>64,58</b>		
	M3	28,17			64,58			63,19			54,17			51,39			48,61			
5	M1	61,81		1,03	56,25		0,52	44,44		0,56	54,17		0,62	69,44		0,82	34,03		0,28	
	M2	64,58	<b>78,47</b>		47,22	<b>63,19</b>		51,39	<b>60,42</b>		51,39	<b>62,5</b>		70,14	<b>79,17</b>		18,06	<b>44,44</b>		
	M3	66,67			33,33			52,78			40,28			65,28			40,28			
6	M1	13,89		0,39	34,72		1,26	25,69		0,46	37,5		0,81	23,61		0,04	49,31		0,14	
	M2	24,31	<b>30,56</b>		43,75	<b>53,47</b>		<b>40,28</b>	37,5		41,67	<b>54,17</b>		34,03	<b>34,72</b>		47,92	<b>51,39</b>		
	M3	24,31			31,25			22,22			42,36			12,5			18,06			
7	M1	66,67		0,66	97,92		0,71	97,22		0,41	99,31		0,5	<b>97,22</b>		0	<b>77,78</b>		0,26	
	M2	60,42	<b>81,25</b>		98,61	<b>100</b>		95,14	<b>98,61</b>		98,61	<b>100</b>		95,83	<b>97,22</b>		54,17	75		
	M3	74,31			97,92			93,75			98,61			95,83			63,89			
8	M1	84,72		0	77,08		0,55	90,28		0,46	77,78		0,5	77,78		0,14	74,31		0,74	
	M2	<b>90,97</b>	<b>90,97</b>		75	<b>82,64</b>		81,25	<b>91,67</b>		76,39	<b>79,17</b>		77,08	80,56		73,61	<b>81,94</b>		
	M3	87,5			65,28			85,42			48,61			<b>81,94</b>			64,58			
9	M1	<b>88,19</b>		0	<b>88,89</b>		0	93,06		0,29	68,75		0,67	85,42		0,16	77,78		0,15	
	M2	87,5	<b>88,19</b>		88,19	<b>88,89</b>		95,14	95,14		64,58	<b>76,39</b>		81,25	<b>86,81</b>		76,39	<b>83,33</b>		
	M3	83,33			87,5			<b>96,53</b>			68,06			80,56			81,94			
$\mu$		64,58	78,47		75	82,64		85,42	91,67		65,97	76,39		77,78	80,56		52,08	66,67		
$\sigma$		29,59	29,17		21,48	15,88		26,26	23,84		19,93	15,28		26,44	23,13		17,54	14,55		

são identificados como M1, M2 e M3. As siglas ME, MD, PE e LG referem-se, respectivamente, às classes mão esquerda, mão direita, pés e língua.

Ainda na Tabela II, a sigla SV representa o algoritmo *Soft Voting*, utilizado como técnica de agregação no *Bagging*. O índice  $z_{\kappa}$  foi calculado a partir da comparação entre o modelo combinado por *Soft Voting* e o modelo individual com maior valor de  $\kappa_m$ . Na parte inferior da tabela, o indicador  $\mu$  representa a média dos valores de  $\kappa_m$  para os modelos individuais e combinados considerando todos os sujeitos. O símbolo  $\sigma$  representa o desvio padrão correspondente.

A análise dos resultados com base no índice  $\kappa_m$  revela que a combinação de modelos proporcionou melhorias em diferentes pares de classes e sujeitos. Para o par (ME x MD), os sujeitos 1, 4, 5, 6 e 7 apresentaram desempenho superior com o modelo combinado. No caso de (ME x PE), os ganhos foram observados nos sujeitos 2, 4, 5, 6, 7 e 8, enquanto para (ME x LG), os sujeitos 1, 2, 3, 4, 5, 7 e 8 se beneficiaram da combinação. Já para o par (MD x PE), todos os sujeitos mostraram melhora com o modelo combinado. Nos pares (MD x LG) e (PE x LG), os ganhos foram verificados, respectivamente, nos sujeitos 2, 3, 4, 5, 6 e 9, e nos sujeitos 1, 2, 3, 4, 5, 6, 8 e 9.

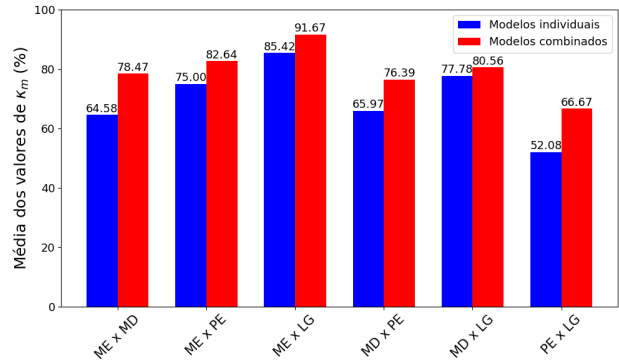


Fig. 5. Média dos  $\kappa_m$  individuais e combinados considerando os pares de classes e sujeitos avaliados.

A Fig. 5 apresenta a comparação do desempenho médio de  $\kappa_m$  entre pares de classes, considerando modelos individuais e combinados. O par (ME x LG) obteve os maiores valores em ambas as abordagens, destacando-se como o mais discriminável. Por outro lado, o par (PE x LG) apresentou os menores desempenhos médios, indicando maior dificuldade

de separação entre essas classes. No geral, todos os pares de classes se beneficiaram da combinação de modelos, com destaque também para os pares (ME × MD) e (ME × PE), que apresentaram ganhos expressivos em relação aos modelos individuais. De forma quantitativa, a estratégia de *Bagging* resultou em um aumento médio de aproximadamente 14,4% nos valores de  $\kappa_m$  quando comparada aos modelos individuais.

Esses resultados indicam que, para a maioria dos pares de classes e sujeitos, a estratégia de aprendizado em conjunto proporciona um aumento na confiabilidade estatística da classificação, refletido por maiores valores de  $\kappa_m$ . No entanto, ao avaliar o índice  $z_\kappa$  entre os modelos combinados e os melhores modelos individuais, observa-se que, em nenhum dos casos, a diferença foi estatisticamente significativa. Isso sugere que, embora o aprendizado em conjunto possa melhorar o desempenho médio, os ganhos não são, em geral, estatisticamente robustos frente à variabilidade individual presente nos sinais de EEG.

### V. CONCLUSÃO

Este trabalho avaliou a combinação de ICMs personalizadas por meio da técnica de *Bagging* na classificação binária de sinais de imagética motora. A análise quantitativa indicou que a abordagem combinada elevou o índice  $\kappa_m$  em média 14,4% em relação aos modelos individuais otimizados por meio da otimização Bayesiana. Entretanto, os ganhos não foram estatisticamente significativos segundo a estatística  $z_\kappa$ , evidenciando limitações na robustez da metodologia aplicada.

Entre as restrições do estudo, destacam-se o uso de apenas um banco de dados público, a limitação à classificação binária e a ausência de experimentos em protótipos reais. Esses fatores reduzem a generalização dos resultados e indicam a necessidade de cautela quanto à adoção imediata da técnica. Apesar disso, os achados reforçam o potencial do aprendizado em conjunto para lidar com a variabilidade interindividual presente nos sinais de EEG.

Como perspectivas futuras, sugere-se explorar outras estratégias de *ensemble* (como *Boosting* e *Stacking*), ampliar a validação em diferentes bases de dados e implementar a abordagem em protótipos de sistemas práticos. Tais avanços podem ampliar a aplicabilidade do método em cenários reais, como reabilitação motora e sistemas de mobilidade assistida, tornando-o mais consistente e relevante para a comunidade científica e para aplicações clínicas.

### REFERÊNCIAS

- [1] F. Lotte, "A tutorial on EEG signal-processing techniques for mental state recognition in brain-computer interfaces," in *Guide to Brain-Computer Music Interfacing*, Londres: Springer, 2014, pp. 133–161. [Online]. Disponível em: <https://hal.inria.fr/hal-01055103>.
- [2] K.-T. Kim, T. Carlson e S.-W. Lee, "Design of a robotic wheelchair with a motor imagery based brain-computer interface," em *Anais do 2013 International Winter Workshop on Brain-Computer Interface (BCI)*, Gangwon, Coreia do Sul, 2013, pp. 46–48, doi: 10.1109/IWW-BCI.2013.6506625.
- [3] G. Papanastasiou *et al.*, "Brain computer interface based applications for training and rehabilitation of students with neurodevelopmental disorders: A literature review," *Heliyon*, vol. 6, n. 9, p. e04250, 2020. [Online]. Disponível em: <https://www.sciencedirect.com/science/article/pii/S240584402031094X>.
- [4] V. Vilas-Boas e C. D. Silva, "Configuração automática de hiperparâmetros em sistemas BCI-MI de sub-bandas," em *Anais do XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBRt)*, Florianópolis, SC, Brasil, 2020, p. 5. [Online]. Disponível em: <https://biblioteca.sbrt.org.br/articles/2245>.
- [5] H. Bashashati, R. K. Ward e A. Bashashati, "Bayesian optimization of BCI parameters," em *Anais do 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2016, pp. 1–5.
- [6] M. Rahimi *et al.*, "Ensemble methods combination for motor imagery tasks in brain computer interface," em *Anais do 2016 23rd Iranian Conference on Biomedical Engineering e 1st International Iranian Conference on Biomedical Engineering (ICBME)*, 2016, pp. 336–340.
- [7] X. Wei, E. Dong e L. Zhu, "Multi-class MI-EEG classification: Using FBCSP and ensemble learning based on majority voting," em *Anais do 2021 China Automation Congress (CAC)*, 2021, pp. 872–876.
- [8] C. Du, C. Shi, H. Huang e X. Wu, "The motor imagery EEG classification method combining common spatial pattern and ensemble learning," em *Anais do 2021 6th International Conference on Communication, Image and Signal Processing (CCISP)*, Chengdu, China, 2021, pp. 361–366, doi: 10.1109/CCISP52774.2021.9639289.
- [9] K. Sen e B. Verma, "Heart disease prediction using a soft voting ensemble of gradient boosting models, RandomForest, and Gaussian Naive Bayes," em *Anais do 2023 4th International Conference for Emerging Technology (INCET)*, Belgaum, Índia, 2023, pp. 1–7, doi: 10.1109/INCET57972.2023.10170399.
- [10] D. Chicco, M. J. Warrens e G. Jurman, "The Matthews correlation coefficient (MCC) is more informative than Cohen's kappa and Brier score in binary classification assessment," *IEEE Access*, vol. 9, pp. 78368–78381, 2021.
- [11] C. Brunner *et al.*, "BCI competition 2008 – Graz data set A," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces)*, Graz University of Technology, vol. 16, 2008. [Online]. Disponível em: <http://www.bbci.de/competition/iv/desc2a.pdf>
- [12] Q. Novi *et al.*, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," em *Anais do 2007 3rd International IEEE/EMBS Conference on Neural Engineering*, 2007, pp. 204–207.
- [13] T. de J. Lopes, *Estudo, Implementação e Comparação de Diferentes Tipos de Pré-Processamento Aplicados a um Sistema Interface Cérebro-Computador*, Trabalho de Conclusão de Curso (Bacharelado em Engenharia Elétrica), TUCURUÍ, 2019.
- [14] Z. Tang *et al.*, "Single-trial classification of different movements on one arm based on ERD/ERS and corticomuscular coherence," *IEEE Access*, vol. 7, pp. 128185–128197, 2019.
- [15] J. R. Wolpaw e E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*, Oxford University Press, 2012.
- [16] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," *Journal of Neural Engineering*, vol. 15, n. 3, p. 031005, abr. 2018. [Online]. Disponível em: <https://dx.doi.org/10.1088/1741-2552/aab2f2>.
- [17] R. Bhuvana, S. Maheshwari e S. Sasikala, "Predict the heart disease using a logistic regression classifier algorithm," em *Anais do 2023 12th International Conference on System Modeling Advancement in Research Trends (SMART)*, 2023, pp. 649–652.
- [18] R. Negri, *Reconhecimento de Padrões: Um Estudo Dirigido*, [S.l.]: [s.n.], 2021. ISBN: 9786555061635.
- [19] J. Z. Esquivel, J. A. C. Vargas e P. Lopez-Meyer, "Fractional adaptation of activation functions in neural networks," em *Anais do 2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 7544–7550.
- [20] T. Mladenova e I. Valova, "Comparative analysis between the traditional k-nearest neighbor and modifications with weight-calculation," em *Anais do 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2022, pp. 961–965.
- [21] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [22] A. Gramfort *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, n. 267, pp. 1–13, 2013.
- [23] V. da S. Jorge, *Aprendizado em conjunto aplicado à classificação da imagética motora*, Dissertação de Mestrado, Núcleo de Desenvolvimento Amazônico em Engenharia, Universidade Federal do Pará, TUCURUÍ, 2025. [Online]. Disponível em: <https://repositorio.ufpa.br/jspui/handle/2011/17393>.

## 4 CONCLUSÃO FINAL

Os dois artigos que compõem este trabalho abordam estratégias distintas, porém complementares, com o objetivo comum de melhorar o desempenho de Interfaces Cérebro-Máquina baseadas em Imagética Motora.

O primeiro artigo concentrou-se na etapa de extração e seleção de características dos sinais de EEG, investigando o uso de métodos de seleção de *features* para aumentar a capacidade de discriminação entre classes motoras. Os resultados demonstraram que técnicas como LASSO e *SelectKBest* possibilitaram melhorias no desempenho dos sistemas de classificação, além de evidenciar a importância da qualidade das características extraídas para o funcionamento das ICMs. Em especial, o método LASSO destacou-se pela capacidade de reduzir significativamente a dimensionalidade dos dados, selecionando um número menor de características sem perda relevante de desempenho.

O segundo artigo direcionou sua análise para a etapa de classificação, explorando a combinação de modelos personalizados por meio de técnicas de aprendizado em conjunto. Nesse contexto, foi proposta uma abordagem baseada em *Bagging* e *Soft Voting*, associada à otimização Bayesiana para ajuste automático dos hiperparâmetros dos modelos. Os resultados mostraram ganhos médios nos valores do índice *Kappa* médio em comparação aos modelos individuais, indicando potencial da estratégia proposta para aumentar a robustez da classificação. Entretanto, as análises estatísticas mostraram que esses ganhos não foram estatisticamente significativos, evidenciando que a elevada variabilidade dos sinais de EEG ainda representa um desafio importante para a generalização dos modelos.

De forma geral, os resultados reforçam que o desempenho de sistemas de Imagética Motora depende da integração entre diferentes etapas da cadeia de processamento de uma ICM. Enquanto o primeiro trabalho evidencia a importância da qualidade e representatividade das *features* extraídas, o segundo destaca a influência da robustez e diversidade dos modelos utilizados na etapa de classificação.

Como perspectivas futuras, destaca-se a possibilidade de integrar as abordagens propostas nos dois artigos, combinando técnicas de seleção de características com estratégias de aprendizado em conjunto. Além disso, também se mostra promissor o refinamento das estratégias de combinação de modelos e o desenvolvimento de métodos capazes de reduzir os impactos da variabilidade intersujeitos dos sinais de EEG.

Por fim, este trabalho contribui para o avanço metodológico de sistemas de Interfaces Cérebro-Máquina baseadas em Imagética Motora, apresentando resultados experimentais relevantes que podem servir como base para pesquisas futuras na área.

## REFERÊNCIAS

- BASHASHATI, H.; WARD, R. K.; BASHASHATI, A. Bayesian optimization of BCI parameters. In: **Anais do 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)**. [S.l.]: IEEE, 2016. p. 1–5.
- BHUVANA, R.; MAHESHWARI, S.; SASIKALA, S. Predict the heart disease using a logistic regression classifier algorithm. In: **Anais do 2023 12th International Conference on System Modeling Advancement in Research Trends (SMART)**. [S.l.]: IEEE, 2023. p. 649–652.
- BRUNNER, C. et al. BCI competition 2008 – Graz data set A. **Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces)**, Graz University of Technology, v. 16, 2008. Disponível em: <<http://www.bbci.de/competition/iv/desc2a.pdf>>. Acesso em: 28 mai. 2026.
- CHICCO, D.; WARRENS, M. J.; JURMAN, G. The Matthews correlation coefficient (MCC) is more informative than Cohen’s kappa and Brier score in binary classification assessment. **IEEE Access**, v. 9, p. 78368–78381, 2021.
- DENG, X. et al. Advanced ts-gl eegnet for motor imagery eeg-based brain-computer interfaces. **IEEE Access**, v. 9, p. 25118–25130, 2021.
- DU, C. et al. The motor imagery EEG classification method combining common spatial pattern and ensemble learning. In: **Anais do 2021 6th International Conference on Communication, Image and Signal Processing (CCISP)**. Chengdu, China: IEEE, 2021. p. 361–366. doi: 10.1109/CCISP52774.2021.9639289.
- DUEÑAS, G.; JIMENEZ, S.; BAQUERO, J. Automatic prediction of item difficulty for short-answer questions. In: **Anais do 2015 10th Computing Colombian Conference (10CCC)**. [S.l.]: IEEE, 2015. p. 478–485.
- ESQUIVEL, J. Z.; VARGAS, J. A. C.; LOPEZ-MEYER, P. Fractional adaptation of activation functions in neural networks. In: **Anais do 2020 25th International Conference on Pattern Recognition (ICPR)**. [S.l.]: IEEE, 2021. p. 7544–7550.
- GRAMFORT, A. et al. MEG and EEG data analysis with MNE-Python. **Frontiers in Neuroscience**, v. 7, n. 267, p. 1–13, 2013.
- JORGE, V. da S. **Aprendizado em conjunto aplicado à classificação da imagética motora**. 2025. Dissertação (Mestrado em Engenharia) – Núcleo de Desenvolvimento Amazônico em Engenharia, Universidade Federal do Pará, Tucuruí, 2025. Disponível em: <<https://repositorio.ufpa.br/jspui/handle/2011/17393>>. Acesso em: 28 mai. 2026.

KIM, K.-T.; CARLSON, T.; LEE, S.-W. Design of a robotic wheelchair with a motor imagery based brain-computer interface. In: **Anais do 2013 International Winter Workshop on Brain-Computer Interface (BCI)**. Gangwon, Coreia do Sul: IEEE, 2013. p. 46–48. doi: 10.1109/IWWBCI.2013.6506625.

KUMAR, J. S.; BHUVANESWARI, P. Analysis of electroencephalography (eeg) signals and its categorization – a study. **Procedia Engineering**, v. 38, p. 2525–2536, 2012. International Conference on Modelling Optimization and Computing.

LINDIG, C.; BOUGRAIN, L. Comparison of sensorimotor rhythms in EEG signals during simple and combined motor imageries over the contra and ipsilateral hemispheres. In: **Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. Milão, Itália: IEEE, 2015. p. 3953–3956.

LOPES, T. de J. **Estudo, Implementação e Comparação de Diferentes Tipos de Pré-Processamento Aplicados a um Sistema Interface Cérebro Computador**. 2019. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Elétrica) – Universidade Federal do Pará, Tucuruí, 2019.

LOTTE, F. A tutorial on EEG signal-processing techniques for mental state recognition in brain-computer interfaces. In: **Guide to Brain-Computer Music Interfacing**. Londres: Springer, 2014. p. 133–161. Disponível em: <<https://hal.inria.fr/hal-01055103>>. Acesso em: 28 mai. 2026.

LOTTE, F. et al. A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. **Journal of Neural Engineering**, v. 15, n. 3, p. 031005, abr. 2018. Disponível em: <<https://dx.doi.org/10.1088/1741-2552/aab2f2>>. Acesso em: 28 mai. 2026.

MIAO, Y. et al. Learning common time-frequency-spatial patterns for motor imagery classification. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, v. 29, p. 699–707, 2021.

MLADENOVA, T.; VALOVA, I. Comparative analysis between the traditional k-nearest neighbor and modifications with weight-calculation. In: **Anais do 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)**. [S.l.]: IEEE, 2022. p. 961–965.

NEGRI, R. **Reconhecimento de Padrões: Um Estudo Dirigido**. [S.l.]: [s.n.], 2021. ISBN: 9786555061635.

NOVI, Q. et al. Sub-band common spatial pattern (SBCSP) for brain-computer interface. In: **Anais do 2007 3rd International IEEE/EMBS Conference on Neural Engineering**. [S.l.]: IEEE, 2007. p. 204–207.

PAPANASTASIOU, G. et al. Brain computer interface based applications for training and rehabilitation of students with neurodevelopmental disorders: A literature review. **Heliyon**, v. 6, n. 9, p. e04250, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S240584402031094X>>. Acesso em: 28 mai. 2026.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

RAHIMI, M. et al. Ensemble methods combination for motor imagery tasks in brain computer interface. In: **Anais do 2016 23rd Iranian Conference on Biomedical Engineering e 1st International Iranian Conference on Biomedical Engineering (ICBME)**. [S.l.]: IEEE, 2016. p. 336–340.

SEN, K.; VERMA, B. Heart disease prediction using a soft voting ensemble of gradient boosting models, RandomForest, and Gaussian Naive Bayes. In: **Anais do 2023 4th International Conference for Emerging Technology (INCET)**. Belgaum, Índia: IEEE, 2023. p. 1–7. doi: 10.1109/INCET57972.2023.10170399.

SILVA, C.; DUARTE, R.; TROFINO, A. Feature extraction improvements using an lmi approach and riemannian geometry tools: An application to bci. In: **Anais do 2016 IEEE Conference on Control Applications (CCA)**. [S.l.]: IEEE, 2016. p. 966–971.

TANG, Z. et al. Single-trial classification of different movements on one arm based on ERD/ERS and corticomuscular coherence. **IEEE Access**, v. 7, p. 128185–128197, 2019.

VILAS-BOAS, v.; SILVA, C. D. Configuração automática de hiperparâmetros em sistemas BCI-MI de sub-bandas. In: **Anais do XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)**. Florianópolis, SC, Brasil: SBrT, 2020. p. 5. Disponível em: <<https://biblioteca.sbrt.org.br/articles/2245>>. Acesso em: 28 mai. 2026.

WEI, X.; E, Dong; ZHU, L. Multi-class MI-EEG classification: Using FBCSP and ensemble learning based on majority voting. In: **Anais do 2021 China Automation Congress (CAC)**. [S.l.]: IEEE, 2021. p. 872–876.

WOLPAW, J. R.; WOLPAW, E. W. **Brain-Computer Interfaces: Principles and Practice**. [S.l.]: Oxford University Press, 2012.

ZHANG, S. et al. Overall optimization of csp based on ensemble learning for motor imagery eeg decoding. **Biomedical Signal Processing and Control**, v. 77, p. 103825, 2022.

ZHENG, L. et al. Ensemble learning method based on temporal, spatial features with multi-scale filter banks for motor imagery eeg classification. **Biomedical Signal Processing and Control**, v. 76, p. 103634, 2022.