

# 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.

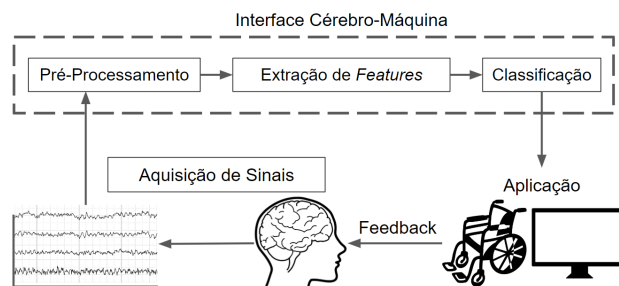


Fig. 1. Fluxograma de uma ICM-IM. Nas etapas de aquisição e pré-processamento são empregadas ferramentas que possibilitam o controle de aplicações baseadas no reconhecimento de padrões em sinais de EEG.

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  $(.)^T$  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,

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

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*  $\hat{\mathbf{f}}_i \in \mathcal{R}^{N_c \times 1}$  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.

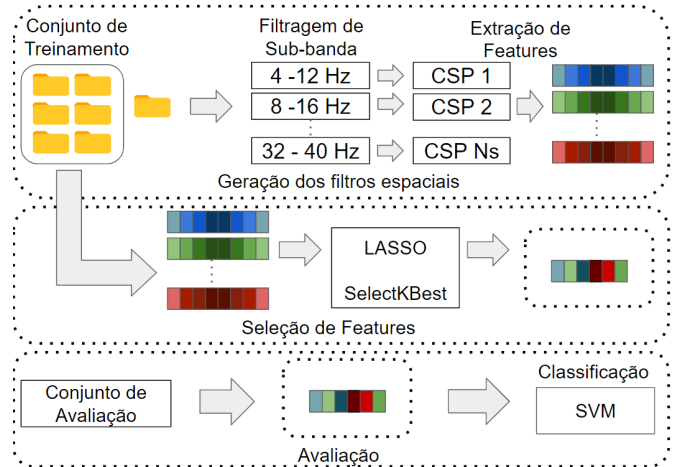


Fig. 2. Configuração proposta da ICM-IM utilizando a geração de filtros espaciais variados.

## 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				$\pm 13,70$	$\pm 12,69$	$\pm 13,10$

## 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. Schlöogl, 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 tsgl-eegnet for motor imagery eeg-based brain-computer interfaces," *IEEE Access*, vol. 9, pp. 25118–25130, 2021.