



UFOP

Universidade Federal
de Ouro Preto

**Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Engenharia Elétrica**

Trabalho de Conclusão de Curso

**Análise do desempenho de técnicas
multiclasse de extração de
características em interfaces cérebro
máquina ativadas por imaginação de
movimento**

Jefferson da Silva Costa

**João Monlevade
2022**

Jefferson da Silva Costa

**Análise do desempenho de técnicas multiclasse
de extração de características em interfaces
cérebro máquina ativadas por imaginação de
movimento**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal de Ouro Preto como parte dos
requisitos para obtenção do Título de Bacharel em
Engenharia Elétrica do Instituto de Ciências Exa-
tas e Aplicadas, da Universidade Federal de Ouro Preto

Orientador: Fabrício Erazo Costa

João Monlevade

2022

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

C837a Costa, Jefferson da Silva.

Análise do desempenho de técnicas multiclasse de extração de características em interfaces cérebro máquina ativadas por imaginação de movimento. [manuscrito] / Jefferson da Silva Costa. - 2022.
67 f.

Orientador: Prof. Me. Fabrício Javier Erazo Costa.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia Elétrica .

1. Interface cérebro-computador. 2. Imaginação - Movimentos. 3. Modelagem. I. Costa, Fabrício Javier Erazo. II. Universidade Federal de Ouro Preto. III. Título.

CDU 004.5

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431



FOLHA DE APROVAÇÃO

Jefferson da Silva Costa

Análise do desempenho de técnicas multiclasse de extração de características em interfaces cérebro máquina ativadas por imaginação de movimento

Monografia apresentada ao Curso de Engenharia Elétrica da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheiro Eletricista

Aprovada em 10 de junho de 2022

Membros da banca

MSc - Fabricio Javier Erazo Costa - Orientador - Universidade Federal de Ouro Preto
Dra - Sarah Negreiros de Carvalho Leite - Universidade Federal de Ouro Preto
Dr - Glauco Ferreira Gazel Yared - Universidade Federal de Ouro Preto

Fabricio Javier Erazo Costa, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 21/06/2022



Documento assinado eletronicamente por **Fabricio Javier Erazo Costa**, PROFESSOR DE MAGISTERIO SUPERIOR, em 21/06/2022, às 12:57, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0348233** e o código CRC **E61BAE6B**.

Agradecimentos

Agradeço imensamente à minha família, pelo apoio e amor incondicionais. Especialmente à minha mãe Geralda e minha avó Dona Glória por não me deixarem desistir frente às dificuldades, que foram e ainda são o alicerce de tudo que sou, ao meu Pai Reginaldo e meu irmão Lucas, obrigado pela parceria, ao Ilklyn, obrigado pelo companheirismo e apoio de anos. Sem vocês nada disso teria sido possível, vocês são meu maior orgulho.

Resumo

Interfaces Cérebro Máquina (ICMs) são sistemas que permitem o controle de dispositivos externos à partir de sinais adquiridos da atividade cerebral. Neste trabalho, foram avaliados os impactos da etapa de seleção de características e de quatro abordagens de extração de características pelo método de Padrões Espaciais Comuns com aplicação de Banco de Filtros (FBCSP do inglês *Filter Bank Common Spatial Patterns*), no desempenho da discriminação entre quatro classes de Imaginação de Movimento (IM) de 9 Indivíduos. Os conjuntos de características para abordagens FBCSP por Diagonalização Simultânea (FBCSP DS), FBCSP Dividir e Conquistar (FBCSP DC), FBCSP Par a Par (FBCSP PW do inglês *Pairwise*) e FBCSP Um versus Resto (FBCSP OxR do inglês *One versus Rest*) foram extraídos de trechos segmentados de 4.5 s e filtrados na faixa de 4-40 Hz por filtros Butterworth de ordem 4 com 4 Hz de largura de banda. Para os métodos FBCSP DC, PW e OxR, é proposta a concatenação das características binárias (modelos independentes) extraídas das combinações específicas de classes propostas em cada abordagem para a criação de um modelo único. O método de seleção de características por Mínima Redundância e Máxima Relevância foi utilizado para selecionar conjuntos ótimos de atributos. Os coeficientes *kappa* para os métodos clássicos FBCSP DC, PW e OxR foram 0.52, 0.572 e 0.569, respectivamente, enquanto que para as abordagens propostas FBCSP DS, DC, PW e OxR foram de 0.5761, 0.6161, 0.6489 e 0.6295, respectivamente. Os métodos propostos apresentaram uma significativa redução no conjunto de características utilizadas para modelagem selecionando, em média, 46% destas. Os resultados deste trabalho, sugerem um aumento no desempenho (coeficiente *kappa*), e uma diminuição no conjunto de características utilizadas. Portanto, as técnicas apresentadas poderiam ser utilizadas para implementar ICMs multiclasse.

Palavras-chaves: Interface Cérebro Máquina, EEG, Imaginação de Movimento, CSP, FBCSP, CSP multiclasse, Informação Mútua, mRMR, SVM.

Abstract

Brain Computer Interfaces (BCI) are systems that allow users to control external devices through signals acquired from brain activity. In this work, we evaluated the impact of the feature selection step and four different approaches to the Filter Bank Common Spatial Pattern method (FBCSP) on the BCI system's performance of 9 subjects to discriminate between four Motor Imagery (MI) classes. The feature sets for the FBCSP Joint Diagonalization (FBCSP DS), FBCSP Divide and Conquer (FBCSP DC), FBCSP Pairwise (FBCSP PW) and FBCSP One versus Rest (FBCSP OxR) approaches were extracted from segments with 4.5 s duration and filtered with Butterworth filters (4 – 40 Hz with order 4 and 4 Hz bandwidth). For the FBCSP DC, PW and OxR approaches, we propose to concatenate the binary feature sets (independent models) extracted from the particular combination of classes, resulting in a single model for each method. The feature selection method Minimum Redundancy and Maximal Relevancy (mRMR) was applied to select an optimal attribute subset. The kappa coefficients for the classical methods FBCSP DC, PW and OxR were 0.52, 0.572 and 0.569, respectively, and for the proposed methods FBCSP DS, DC, PW and OxR resulted in 0.5761, 0.6161, 0.6489 and 0.629, respectively. The feature selection step significantly reduced the number of features used to estimate the models, selecting on average 46% of attributes. This study suggests an improvement in performance (kappa coefficient) and a reduction of the feature subsets. Therefore, the presented methods could be applied to a multiclass BCI implementation.

Keywords Brain Computer Interfaces, EEG, Motor Imagery, CSP, FBCSP, multiclass CSP, Mutual Information, mRMR, SVM.

Lista de ilustrações

Figura 1 – Mapeamento das áreas corticais: Áreas 4 e 6 são as áreas de interesse para atividades motoras sendo área motora primária e motora suplementar, respectivamente	18
Figura 2 – Homúnculo Motor de Penfield - Representação das áreas somatotópicas na região sensório-motora central e suas relações com membros do corpo humano	19
Figura 3 – (a) Padrão internacional de posicionamento de eletrodos 10-20 (b) Padrão internacional de posicionamento de eletrodos 10-10	20
Figura 4 – Representação Esquemática das principais etapas de uma ICM baseada em EEG: I-Aquisição de sinais elétricos de EEG, II-Pré-processamento, condicionamento dos sinais e remoção de artefatos, III-Extração de características, IV-Seleção de características e V-Aplicação de um algoritmo de classificação para identificação da ação	22
Figura 5 – Esquema temporal do paradigma do experimento. De $t = -1s$ à $t = 0s$ o indivíduo recebe a indicação visual do movimento que deve ser imaginado, de $t = 0s$ à $t = 3s$ deve sustentar a imaginação do movimento (BRUNNER et al., 2008)	25
Figura 6 – Arranjo dos eletrodos utilizado na aquisição dos sinais de EEG da base de dados. Os eletrodos são numerados de 1 à 25, sendo os 22 primeiros referentes aos sinais de EEG e os três últimos aos sinais de eletro-oculogramas.	26
Figura 7 – Trecho de evento de IM MD para indivíduo 1 no canal C3 segmentado. Em destaque, observa-se o intervalo onde há IM MD.	27
Figura 8 – Média das observações de cada classe para os eletrodos de EEG. Destacados em roxo, estão os intervalos onde ocorrem os eventos de IM.	28
Figura 9 – Média das observações de cada classe para os eletrodos de EEG após a aplicação do referenciamento por CAR. Destacados em roxo, estão os intervalos onde ocorrem os eventos de IM.	29
Figura 10 – (a) Estimativa PLV global calculado através do PLV das combinações de pares dos eletrodos FZ, C3, CZ, C4 e PZ. (b) Estimativa do PLV local calculado a partir da média dos PLVs dos eletrodos na vizinhança de C3.	30
Figura 11 – Metodologia de extração de características por FBCSP. O sinal de EEG de entrada é filtrado de 4-40Hz em bandas de 4Hz sem sobreposição, em seguida, é realizada a extração de características por CSP multiclasse para a saída de cada filtro.	33

Figura 12 – Banco de Filtros aplicado aos sinais de EEG. O banco é composto por filtros Butterworth de ordem 4, de 4-40Hz com bandas de 4Hz sem sobreposição.	34
Figura 13 – Modelo CSP DC. Dos sinais de EEG de entrada S são extraídas características CSP binárias através de $C-1$ modelos, onde C é o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.	35
Figura 14 – Modelo CSP OxR. Dos sinais de EEG de entrada S são extraídas características binárias através de C modelos, sendo C o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.	36
Figura 15 – Modelo CSP Par a Par. Dos sinais de EEG de entrada S são extraídas características binárias através de C modelos, sendo C o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.	36
Figura 16 – Modelo CSP DS. As características são extraídas através de um único modelo do qual se estima uma matriz de transformação W que, aplicada aos dados de EEG de entrada resultando em um conjunto de características Z	37
Figura 17 – Relação entre Informação Mútua e Entropia: O diagrama representa a relação entre medidas de informação associadas às variáveis X e Y . A área de intercessão entre os diagramas representa a informação mútua $I(X,Y)$ entre as variáveis X e Y . Os círculos representam, individualmente, as medidas de entropia de cada variável $H(X)$ e $H(Y)$	40
Figura 18 – Método de seleção de características: Ao conjunto de características, inicialmente vazio, são adicionadas de forma sequencial as características uma a uma, cada subconjunto é avaliado de acordo com os critérios de mRMR. Caso um subconjunto seja melhor avaliado em relação ao anterior, define-se um novo subconjunto. O procedimento se repete até que o critério de parada seja atingido e um subconjunto ótimo seja definido.	40

Figura 19 – Etapas da ICM implementada: <i>I</i> - Pré-processamento, <i>II</i> - Extração de características pelos métodos CSP multiclasse, <i>III</i> - Seleção de características por mRMR baseado em Informação Mútua <i>IV</i> - Classificação, com a realização pelo algoritmo <i>SVM Par a Par</i> com <i>kernel</i> gaussiano.	46
Figura 20 – Taxas de acerto para FBCSP, CSP,PLV, Welch: Média das taxas de acerto de 6 modelos binários extraídos da combinação de pares de classes (4 classes).	48
Figura 21 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das amostras de modelagem para o método FBCSP DS	50
Figura 22 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP DS	52
Figura 23 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP DS	53
Figura 24 – Desempenho dos sistemas (coeficiente <i>kappa</i>) dos quatro métodos de extração de características, para os 9 indivíduos.	54
Figura 25 – Matrizes de Confusão das quatro abordagens FBCSP DS, DC, PW e OxR derivadas da classificação das 2292 observações de todos os 9 indivíduos.	55
Figura 26 – Distribuição das características selecionadas pelo método <i>mRMR</i> no tempo [-0.5 à 4 s] e na frequência [4 à 40 Hz].	56
Figura 27 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP DC	63
Figura 28 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem para o método FBCSP DC	64
Figura 29 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP DC	64
Figura 30 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP PW	65
Figura 31 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem do método FBCSP PW	66

Figura 32 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP PW	66
Figura 33 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP OxR	67
Figura 34 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem para o método FBCSP OxR	68
Figura 35 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP OxR	68

Lista de tabelas

Tabela 1	– Bandas de frequência de EEG, suas respectivas amplitudes médias e o tipo de atividade às quais se relacionam (BAMDAD; ZARSHENAS; AUAIS, 2015)	21
Tabela 2	– Número de observações após normalização de classes. Foram removidos os trechos sinalizados como contendo artefatos no arquivo de informações da base de dados. Desta forma, para cada indivíduo, são consideradas o mesmo número de observações para todas as quatro classes de IM nos dados de treinamento e de teste.	27
Tabela 3	– Taxas de acerto para os métodos binários de extração de características CSP, PLV e Welch: Foram extraídas características das combinações das classes par a par. Os valores representam a média da taxa de acerto para a classificação de seis modelos (4 classes). Em negrito estão destacadas as maiores taxas de desempenho por indivíduo e o maior valor médio para a taxa de acerto.	47
Tabela 4	– Distribuição dos modelos para os métodos: FBCSP DS - um modelo para discriminação das quatro classes de maneira simultânea. FBCSP DC - três modelos binários para discriminação. FBCSP PW - seis modelos binários para discriminação. FBCSP OxR - quatro modelos binários para discriminação	48
Tabela 5	– Número de características por Modelo após a concatenação das características extraídas de cada método	49
Tabela 6	– Coeficiente Kappa após a validação cruzada com 5x10-fold para os quatro métodos propostos	49
Tabela 7	– Porcentagem, em relação ao conjunto total, de características selecionadas pelo método <i>mRMR</i> para cada um dos métodos propostos e cada um dos indivíduos	50
Tabela 8	– Desempenho dos sistemas utilizando as amostras cegas para os quatro métodos. É realizada uma comparação entre os modelos que utilizam 100% das características e os que utilizam os subconjuntos escolhidos pelo método de seleção. Em negrito estão destacados os maiores valores de desempenho em relação ao conjunto de características utilizado e sublinhado destacam-se os maiores valores de desempenho para cada indivíduo e da média geral.	51

Tabela 9 – Comparação do desempenho dos modelos propostos e modelos da literatura. A nomenclatura '-C' identifica os métodos onde é proposta a concatenação das características. Em negrito, estão destacados os maiores valores de desempenho em relação à metodologia proposta e sublinhado estão os maiores valores de desempenho para cada indivíduo.	55
Tabela 10 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método CSP. As características foram extraídas do intervalo de $t=-0.5s$ à $t=3s$ após filtragem de 8-30Hz	62
Tabela 11 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método PLV. As características foram extraídas pelo método de PLV Global dos sinais de EEG segmentados de $t=0s$ à $t=3s$ filtrados de 8-30Hz	62
Tabela 12 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método do Periodograma de Welch. As características foram extraídas do sinal segmentado de $t=0s$ à $t=3s$ em trechos de 1s com sobreposição de 500ms	63

Lista de abreviaturas e siglas

PLV	<i>Phase Locking Value</i>
CSP	<i>Common Spatial Pattern</i>
FBCSP	<i>Filter Bank Common Spatial Pattern</i>
DS	Diagonalização Simultânea
DC	Dividir e Conquistar
PP	Par a Par
PW	<i>Pair Wise</i>
OxR	<i>One vs Rest</i>
SVM	<i>Support Vector Machine</i>
ICM	Interface Cérebro-Máquina
IM	Imaginação do Movimento
ME	Mão Esquerda
MD	Mão Direita
PÉ	Ambos os Pés
LI	Lingua
EEG	Eletroencefalograma
EOG	Eletro-Oculograma
CAR	<i>Common Average Reference</i>
mRMR	Máxima Relevância Mínima Redundância

Sumário

1	INTRODUÇÃO	16
1.1	Objetivos	17
1.1.1	Objetivos Específicos	17
2	REVISÃO BIBLIOGRÁFICA	18
2.1	Região sensório motora	18
2.2	Aquisição do Eletroencefalograma	19
2.2.1	Padrões Internacionais de Posicionamento de Eletrodos	20
2.2.2	Bandas de frequência do EEG	20
2.2.3	Respostas motoras relacionadas a eventos motores	21
2.3	Interfaces cérebro-máquinas	22
2.3.1	Tipo de sinal utilizado	23
2.3.2	Adequação do Sinal	23
2.3.3	Extração de Características	23
2.3.4	Seleção de Características	24
2.3.5	Classificação	24
3	MATERIAIS E MÉTODOS	25
3.1	Base de dados	25
3.2	Pré-Processamento	26
3.2.1	Segmentação dos Dados	28
3.2.2	Filtragem Espacial	28
3.3	Extração de Características	29
3.3.1	<i>Phase Locking Value</i>	29
3.3.2	Periodograma de Welch	30
3.3.3	Padrões Espaciais Comuns	31
3.3.4	<i>Filter Bank Common Spatial Pattern</i>	33
3.4	Estratégias CSP Multiclasse	34
3.4.1	CSP Dividir e Conquistar	34
3.4.2	CSP Um versus Resto	35
3.4.3	CSP Par a Par	36
3.4.4	CSP Diagonalização Simultânea	37
3.5	Seleção de Características	38
3.5.1	Máxima relevância e Mínima redundância	38
3.6	Classificação	41
3.6.1	<i>Support Vector Machine</i>	41

3.6.2	Dados linearmente separáveis com tolerância de violação de margem	43
3.6.3	Dados não linearmente separáveis	44
3.7	Metodologia de Processamento	45
4	RESULTADOS E DISCUSSÃO	47
4.1	Comparação FBCSP, CSP, PLV e Welch	47
4.2	FBCSP	48
4.3	FBCSP multiclasse com seleção de características	50
4.3.1	Teste dos modelos com amostras cegas	51
4.3.2	Comparação com modelos de votação	55
4.3.3	Padrões Tempo x Frequência das Características Seleccionadas	56
5	CONCLUSÃO	57
	REFERÊNCIAS	58
6	MATERIAIS ELABORADOS PELO AUTOR	62

1 Introdução

Interfaces cérebro-máquina (ICM) por imaginação do movimento (IM), são sistemas que realizam a comunicação direta entre humano e computador, através da aquisição de sinais decorrentes da atividade neural gerada por eventos de imagética motora que, após processados, são traduzidos em comandos (ALLISON; WOLPAW; WOLPAW, 2007). Esses sistemas são amplamente aplicados na comunicação, reabilitação e melhoria da qualidade de vida de pacientes com movimentos parcial ou totalmente comprometidos.

Dentre os principais desafios na implementação de ICMs, está a etapa de extração de características que podem ser de potência, computadas, por exemplo, pelo método de estimação espectral de potência de Welch (WELCH, 1967), características de fase, como *Phase Locking Value* (PLV) que estima a sincronização entre canais de EEG (WANG et al., 2006) ou estatísticos, como *Common Spatial Patterns* (CSP), um método de filtragem espacial que tem como objetivo encontrar projeções do conjunto de observações de EEG que maximizam a variância para as observações de uma classe e minimizam a variância para as observações da outra, obtendo bom desempenho na discriminação entre duas classes (BLANKERTZ et al., 2007). A extração de características por CSP é, normalmente, realizada na banda de frequência de EEG de 8-30Hz. A abordagem *Filter Bank Common Spatial Patterns* (FBCSP) propõe a aplicação de um banco de filtros afim de selecionar canais com maior capacidade de discriminação entre classes para cada uma das bandas (ANG et al., 2008)

Dentre as extensões multiclasse do método FBCSP, existem as que buscam definir modelos que discriminem as classes simultaneamente, como o método FBCSP por diagonalização simultânea (LIYANAGE et al., 2010), e abordagens que buscam dividir o problema na estimação de múltiplos modelos binários, como os métodos FBCSP *Ove vs Rest* (OxR), *Pairwise* (PW) e *Dividir e conquistar* (DC) (ANG et al., 2012) que, devido o modo de organização das características extraídas, necessitam de múltiplos modelos de classificadores afim de realizar a discriminação entre classes.

Desta forma, neste trabalho é proposta uma abordagem onde as características extraídas pelos métodos PW, DC e OxR sejam concatenadas em um único vetor de características para cada observação, afim de estimar, para cada um dos métodos, um único modelo de classificador. Será avaliado se abordagem proposta, aliada a uma seleção de características por informação mutua (PENG; LONG; DING, 2005), e o método de classificação por *Support Vector Machine* (SVM) com um *kernel* Gaussiano é capaz de melhorar, de forma significativa, o desempenho dos métodos propostos pela literatura.

Na primeira sessão do trabalho, discute-se sobre as principais características fi-

siológicas das regiões corticais de interesse, alguns aspectos importantes das respostas observadas durante eventos de IM, bem como estrutura geral e etapas de uma ICM. Em seguida, são discutidos os principais materiais e métodos utilizados nas etapas de pré-processamento, extração de características, seleção de características e classificação. Por fim, são apresentados e discutidos os resultados obtidos e realizadas algumas comparações entre esses resultados e resultados da literatura.

1.1 Objetivos

Este trabalho tem por objetivo a implementação de diferentes técnicas de extração de características de métodos FBCSP multiclasse, e a avaliação do impacto desses métodos no desempenho de uma Interface Cérebro Máquina *single trial*.

1.1.1 Objetivos Específicos

- Comparar o desempenho de sistemas com extração de características de fase, potência e espaciais.
- Avaliar o impacto de adaptações em técnicas da literatura no desempenho dos métodos de extração de características.
- Avaliar o impacto da etapa de seleção de características e determinar a existência de padrões nos conjuntos de características selecionados.

2 Revisão bibliográfica

2.1 Região sensório motora

A realização de tarefas motoras voluntárias segue uma hierarquia funcional que pode ser dividida em *estratégia*, *tática* e *execução* (BEAR; CONNORS; PARADISO, 2008). A estratégia, realizada no neocórtex e nos núcleos basais, determina a partir de informações sensoriais, a relação entre o corpo e o ambiente. A tática, realizada no córtex central e no cerebelo, planeja, espacial e temporalmente, sequências de músculos a serem contraídos para a realização da tarefa. A execução, envolvendo tronco encefálico e conjuntos neuronais originados principalmente no córtex motor descendentes na medula, é responsável pela ativação dos neurônios que geram o movimento relacionado à tarefa (BEAR; CONNORS; PARADISO, 2008). Na Figura 1 é possível observar as principais estruturas responsáveis pelo planejamento e execução de tarefas motoras: o córtex motor primário na área 4, e o córtex motor suplementar na área 6, juntas, essas áreas estão envolvidas no controle, planejamento e direcionamento de movimentos voluntários.

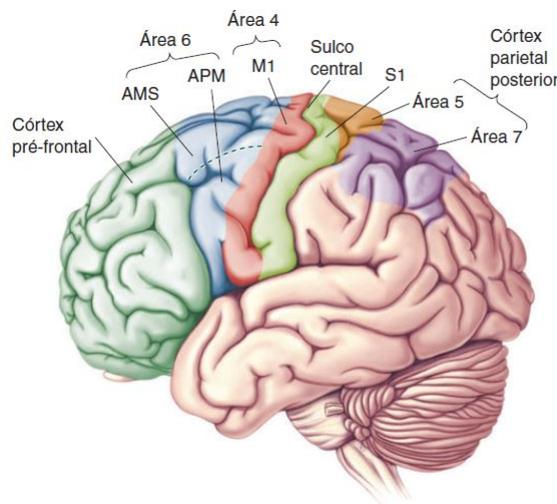


Figura 1 – Mapeamento das áreas corticais: Áreas 4 e 6 são as áreas de interesse para atividades motoras sendo área motora primária e motora suplementar, respectivamente

Entre 1930 e 1950, o neurocirurgião Wilder Graves Penfield, realizou um mapeamento da ativação das áreas corticais relacionadas à funções motoras e sensoriais de diferentes partes do corpo humano (BEAR; CONNORS; PARADISO, 2008). Desses estudos, Penfield concluiu que o giro pré-central do lobo frontal está fortemente relacionado com respostas motoras e somatossensórias (PENFIELD; BOLDREY, 1937). O resultado desse mapeamento ficou conhecido como homúnculo de Penfield, como se mostra na Figura

2, (BEAR; CONNORS; PARADISO, 2008).

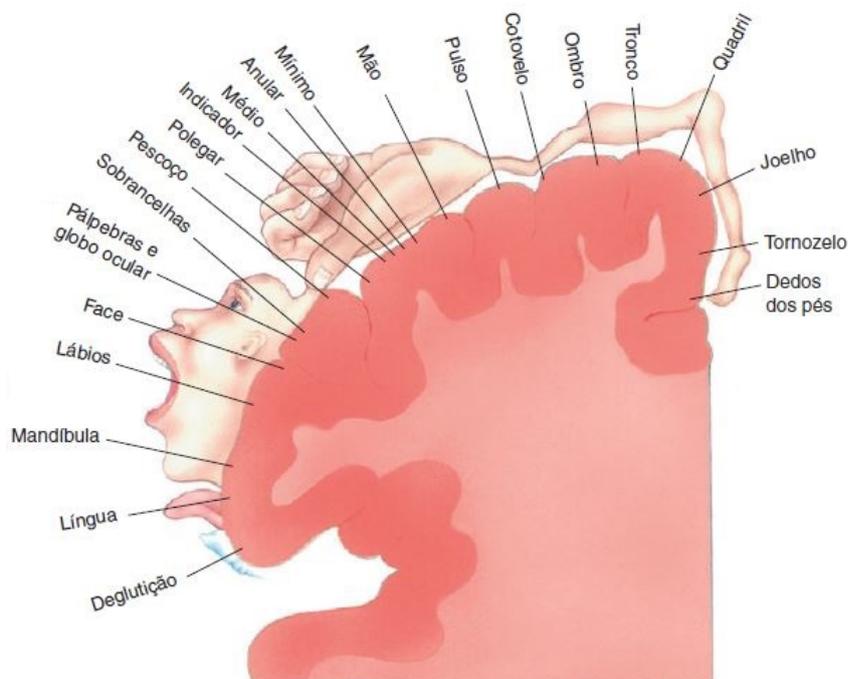


Figura 2 – Homúnculo Motor de Penfield - Representação das áreas somatotópicas na região sensório-motora central e suas relações com membros do corpo humano

Do córtex motor primário, suplementar e pré-motor, se originam dois terços das fibras das vias laterais, especialmente o trato córtico-espinal, por onde se propagam as informações de controle da musculatura distal durante movimentos voluntários (BEAR; CONNORS; PARADISO, 2008). Os axônios desse trato são majoritariamente contralaterais, significando que o lado direito do córtex motor controla a musculatura do lado esquerdo do corpo e vice-versa.

2.2 Aquisição do Eletroencefalograma

Normalmente, a amplitude dos sinais de EEG é de 5 a 200 μV e precisa ser amplificado antes que possa ser gravado (BAHRI,2014). Estes sinais podem ser corrompidos por diferentes fontes de ruído como artefatos fisiológicos, EMG, ECG, EOG, movimentos do corpo, suor ou artefatos de instrumentação, tais como o ruído de 60 Hz da rede elétrica, ou a diminuição da amplitude pela própria impedância do eletrodo ou dos fios (RAMPIL, 1998). Esses problemas são comumente solucionados com a adição de etapas de pré-amplificação, filtragem e melhoria da relação sinal ruído nos circuitos de coleta (ZHANG et al., 2017).

2.2.1 Padrões Internacionais de Posicionamento de Eletrodos

O registro dos sinais de eletroencefalograma (EEG) é realizado com o posicionamento de eletrodos no escalpo. A distribuição dos eletrodos se baseia em pontos de referência padrão, e deve ser feita de maneira proporcional ao tamanho do crânio (KLEM *et al.*, 1999). A superfície do escalpo deve estar adequadamente coberta pelos eletrodos, e a nomenclatura alfanumérica é determinada de forma a identificar determinadas regiões do córtex. Na Figura 3 é possível ver as distribuições de eletrodos e nomenclaturas para os padrões internacionais 10-20 e 10-10. Nos padrões de posicionamento mais utilizados, a distribuição dos eletrodos é feita com base nas medidas entre os pontos náseo e ínio, com o posicionamento dos eletrodos em intervalos de 10% dessa mesma medida, realizando o mesmo processo com as referências auriculares A1 e A2.

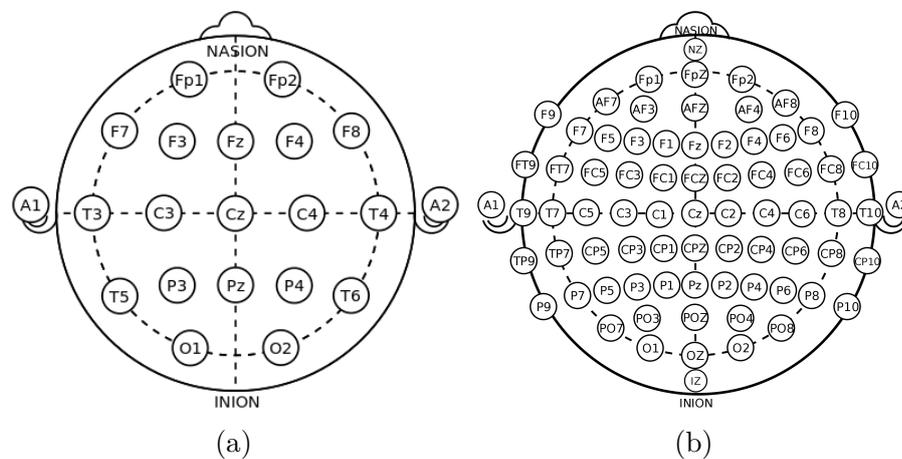


Figura 3 – (a) Padrão internacional de posicionamento de eletrodos 10-20 (b) Padrão internacional de posicionamento de eletrodos 10-10

2.2.2 Bandas de frequência do EEG

A amplitude dos sinais de EEG está relacionada com o número de neurônios ativados de forma síncrona, um maior número de neurônios síncronos indica uma maior amplitude, enquanto um menor número resulta em menor amplitude (C-H, 2018). Em geral, a frequência de oscilações cerebrais é inversamente relacionada com a sua amplitude. Por exemplo, observações do ritmo mu, frequências entre 8 e 13 Hz, tem maior amplitude que o ritmo beta central com frequências em torno de 20 Hz, por sua vez, o ritmo beta tem uma amplitude maior que oscilações em torno de 40 Hz (PFURTSCHELLER; LOPES, 1999). Assim, faixas de frequência mais baixa correspondem à ativação de grupos maiores de neurônios, em contraponto, em frequências mais altas, observa-se um número menor de neurônios sendo ativados. Na Tabela 1 é possível observar uma estimativa das relações entre amplitude do sinal de EEG e as principais bandas de frequência observadas nesses sinais.

Banda	Frequências (Hz)	Amplitude(μ V)	Atividade Correspondente
Delta	0.5-4	20 - 200	Sono
Theta	4 -7	< 20	Meditação
Alpha	8 -12	30 - 50	Relaxamento
Beta	13 - 30	5 - 10	Concentração
Gama	>30	5 - 10	Percepção consciente

Tabela 1 – Bandas de frequência de EEG, suas respectivas amplitudes médias e o tipo de atividade às quais se relacionam (BAMDAD; ZARSHENAS; AUAIS, 2015)

2.2.3 Respostas motoras relacionadas a eventos motores

As respostas a estímulos internos ou externos criam dois tipos de sinais registrados no EEG, o primeiro conhecido como potencial relacionado ao evento (ERP do inglês *Event Related Potential*) e a sincronização/dessincronização relacionada ao evento (ERD/ERDE do inglês *Event Related Desynchronization/Synchronization*) (PFURTSCHELLER; LOPES, 1999).

Respostas do tipo ERP são chamadas de síncronas em tempo e frequência, (*time-locked*), pois seu atraso em relação ao estímulo externo é considerado constante, assim, a detecção desses potenciais pode ser feita a través de técnicas lineares como a estimação da média de n trechos (PFURTSCHELLER; LOPES, 1999).

Com gravação de sinais de EEG no córtex motor primário durante tarefas motoras, é percebida uma redução do potencial cerca de 2 segundos antes do início da imaginação ou execução do movimento, esse comportamento é chamado de potencial relacionado ao evento (MRCP do inglês *Movement-related cortical potential*) (DEECKE; SCHEID; KORNHUBER, 1969). Este potencial é dividido em três componentes divide os MRCPs em três componentes: A primeira conhecida como potencial de prontidão (Readiness Potential), que é caracterizado por um leve aumento da amplitude do sinal cerca de 1s antes da execução da tarefa, seguido do potencial pré-motor, um aumento mais rápido na amplitude, e por fim, o potencial motor, que apresenta um aumento na negatividade do sinal.

Em contraponto, existem respostas temporalmente atreladas ao evento mas que não apresentam a mesma fase ao início de cada segmento, denominadas de fase bloqueadas (*phase-locked*) (PFURTSCHELLER; LOPES, 1999), que apresentam características não-lineares. ERD/ERS são respostas caracterizadas por mudanças na potência de faixas de frequência específicas do EEG, especialmente na faixa de 8 a 13 Hz (banda *mu*), associadas aos ritmos sensoriomotores e observadas principalmente em áreas contralaterais ao movimento (PFURTSCHELLER; LOPES, 1999). ERD é o resultado de uma diminuição na sincronização de neurônios, o que causa diminuição de potência em bandas de frequências relacionadas ao EEG basal, ao contrário, ERS é o resultado de um aumento na sincronização de neurônios, o que produz um aumento na potência relacionado ao EEG basal (HE et al.,

2013).

A análise da ERD/ERS em áreas sensório-motoras pode ser usada para identificar e classificar estados de planejamento e imaginação de diferentes classes de movimento (HE et al., 2013), isso aliado a possibilidade de indivíduos aprenderem a controlar a amplitude ritmos sensório-motores através da imagética motora (WOLPAW et al., 1991) viabiliza a implementação de sistemas de interface cérebro máquina, que interligam diretamente o sistema nervoso central e aplicações computacionais externas.

A IM pode ser entendida com o processo no qual o indivíduo imagina a sensação do movimento sem a ativação das vias eferentes (músculos) (DECETY, 1996). Esse tipo de resposta não depende de estímulos externos, uma vez que a imagem do movimento se forma através de modelos já definidos por parâmetros memorizados da sequencia de músculos a serem ativados (GUILLOT; COLLET, 2010). Apesar da IM e a execução do movimento serem diferentes, possuem áreas de ativação cortical similares principalmente nas regiões pré-motora e motora suplementar (JEANNEROD, 1994), com correspondências na região somato-sensitiva, ressaltando paralelos fisiológicos entre imaginação, planejamento e execução (DECETY, 1996). Desta forma, a prática mental de IM possui aplicações importantes em técnicas de neuroreabilitação em indivíduos com algum tipo de comprometimento motor (BAMDAD; ZARSHENAS; AUAIS, 2015).

2.3 Interfaces cérebro-máquinas

Uma ICM converte a atividade cerebral em uma saída artificial que substitui, restaura, complementa ou melhora saída natural, desse modo, altera a evolução das interações entre o sistema nervoso central e meio ambiente (WOLPAW, 2013). Assim, uma ICM cria uma interação em tempo real entre seu usuário e o ambiente sem a necessidade de controle ou ação muscular.

Esse tipo de sistema é comumente implementado em quatro etapas principais: aquisição dos sinais, extração de características, classificação e tradução em comandos (HE et al., 2013).

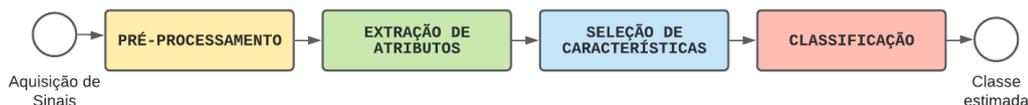


Figura 4 – Representação Esquemática das principais etapas de uma ICM baseada em EEG: I-Aquisição de sinais elétricos de EEG, II-Pré-processamento, condicionamento dos sinais e remoção de artefatos, III-Extração de características, IV-Seleção de características e V-Aplicação de um algoritmo de classificação para identificação da ação

Dentre as diversas categorias de ICMs, destacam-se as ICM baseadas em EEG por sua alta resolução temporal, portabilidade, uso relativamente simples e o fato de a aquisição do sinal ser feita de maneira não invasiva.

2.3.1 Tipo de sinal utilizado

Uma vez que as ICMs baseadas em EEG buscam interpretar e traduzir sinais eletrofisiológicos em respostas somatossensoria ou motoras, existem diversas abordagens a serem usadas no que diz respeito ao tipo de sinal utilizado. (WOLPAW et al., 2002) faz uma diferenciação entre algumas classes de ICMs de acordo com os sinais utilizados dentre os quais, destacam-se *Movement-related cortical potentials* e ritmos sensoriomotores.

Segundo (WOLPAW, 2013), ICMs que utilizam ritmos sensoriomotores podem apresentar melhor desempenho na realização de tarefas cognitivas mais complexas. Assim, a utilização da banda μ em aplicações desse tipo apresentam uma série de vantagens uma vez que percebe-se uma forte associação entre essa faixa de frequências e as áreas mais diretamente responsáveis por respostas motoras e a análise é feita principalmente com base na ERD/ERS dessa banda (WOLPAW et al., 1991) e (WOLPAW et al., 2002). Uma característica relevante do uso desse tipo de sinais em aplicações ICM, é que não há a necessidade da realização do movimento, possibilitando assim, a implementação de sistemas baseados em imagética motora(WOLPAW, 2013).

2.3.2 Adequação do Sinal

Esta etapa visa a adequação do sinal para as etapas posteriores. Filtragem digital, segmentação, janelamento podem ser utilizadas no pré-processamento (SARMA et al., 2016). A implementação dessa etapa visa a melhoria do desempenho geral do sistema através da atenuação de ruídos e remoção de artefatos (SARMA et al., 2016).

2.3.3 Extração de Características

A extração de características tem como objetivo a obtenção de atributos compactos que melhor representem um determinado sinal. Tratando-se de ICMs baseadas em EEG, as características podem ser extraídas de ritmos sensório motores, ou sinais induzidos por estímulos visuais, táteis ou auditivos (DORNHEGE et al., 2007). Podem ser representadas informações de fase/sincronização, potência ou informações espaciais destes sinais. Porém, a etapa extração de características tem, por objetivo final, a extração de atributos que facilitem a interpretação dos sinais (WOLPAW, 2013).

2.3.4 Seleção de Características

Esta etapa é opcional e tem como objetivo encontrar um subconjunto de características que possa representar melhor a informação quando comparado com o conjunto total. Os conjuntos de características analisadas estão distribuídos em espaços de alta dimensionalidade. Os algoritmos de seleção realizam a redução do número de características de forma que com o menor subconjunto possível, obtenha-se um desempenho igual ou superior ao sistema quando utilizando o conjunto completo de características (DASH; LIU, 1997). Comumente, a seleção de características é realizada em quatro etapas: Busca heurística por um subconjunto de amostras, avaliação do subconjunto, critério de parada e validação dos resultados. Sequencialmente, escolhe-se um subconjunto de características de acordo com algum método de busca, esse subconjunto é então, avaliado por um critério pré-definido, as características melhor avaliadas são, por fim, validadas (TANG; ALELYANI; LIU, 2014). Métodos que performam seleção de características podem ser implementados de maneira independente de algoritmos de classificação, como os chamados filtros, ou utilizar a otimização de parâmetros de avaliação do subconjunto ótimo de características, como é o caso dos *wrappers* (TANG; ALELYANI; LIU, 2014).

Os métodos de filtros para seleção de características, tipicamente, tem dois passos. No primeiro passo, é gerado um subconjunto de características e no segundo passo as características são avaliadas seguindo um critério pré estabelecido. Posteriormente, as características são *rankeadas* e o subconjunto com características melhor avaliado é selecionado como o subconjunto ótimo (TANG; ALELYANI; LIU, 2014).

2.3.5 Classificação

Nesta etapa, um modelo matemático definido a partir das características extraídas dos sinais de EEG é utilizado para realizar a discriminação entre as classes em análise, traduzindo as características extraídas em sinais de comando (HE et al., 2013). Dentre os métodos mais utilizados para a classificação de características em ICMs estão os algoritmos de redes neurais artificiais (PFURTSCHELLER et al., 1997) e *Support Vector Machines* (VALLABHANENI; HE, 2004).

3 Materiais e Métodos

Neste Capítulo são apresentados os dados utilizados e a metodologia proposta para a implementação das etapas de uma interface cérebro máquina.

3.1 Base de dados

A base de dados analisada é da competição de ICMs denominada a *BCI competition* de 2008 (BRUNNER et al., 2008). Os sinais foram registrados com a configuração de eletrodos no padrão internacional 10-20, e consiste na IM de quatro tarefas motoras distintas: IM mão esquerda (IM ME)(1), IM mão direita (IM MD)(2), IM de ambos os pés (IM PÉ)(3) e IM da língua (IM LI)(4). Os dados de 9 indivíduos foram gravados em sessões com um dia de separação, e cada tarefa foi repetida 6 vezes por sessão.

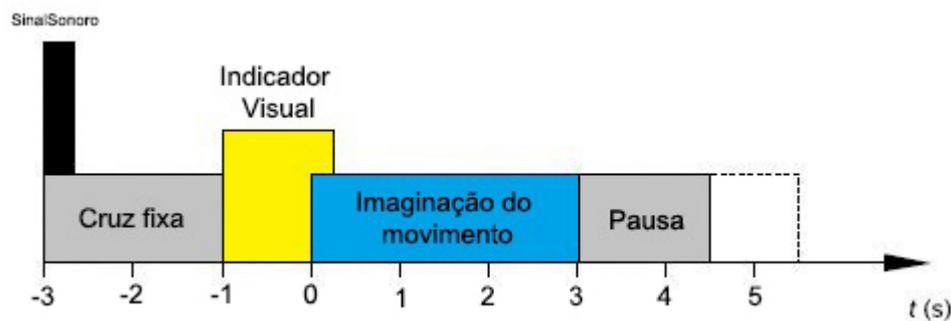


Figura 5 – Esquema temporal do paradigma do experimento. De $t = -1s$ à $t = 0s$ o indivíduo recebe a indicação visual do movimento que deve ser imaginado, de $t = 0s$ à $t = 3s$ deve sustentar a imaginação do movimento (BRUNNER et al., 2008)

O protocolo de temporização para realização das IMs se inicia com o indivíduo sentado em posição confortável, foi exibida uma cruz fixa por 2 s, em seguida um indicador visual de preparação é mostrado e alerta ao usuário que a tarefa deve ser realizada, com a indicação o tipo de IM. O sujeito é instruído a sustentar a IM por 3 s, seguidos por um período de repouso de 1.25 s entre cada ciclo. A Figura 5 apresenta um esquemático do protocolo, considerando o tempo $t = 0$ o momento em que se inicia a imaginação da tarefa.

A frequência de amostragem do sinal de EEG foi de 250 Hz, assim, o condicionamento do sinal é feito com a aplicação de um filtro passa-faixas de 0.5-100 Hz, além da aplicação de um filtro *notch* para a retirada do ruído da rede elétrica de 50 Hz.

Os eletrodos estão dispostos de acordo com o padrão internacional 10-20, e estão relacionados com os canais de saída numerados de 1 à 25 como mostra a Figura 6. Sendo

os primeiros 22 eletrodos correspondentes aos sinais de EEG coletados, e os três últimos contendo os sinais de eletro-oculograma. Os eletrodos C3, CZ, C4 estão relacionados a região sensório-motora do córtex correspondentes aos movimentos imaginados.

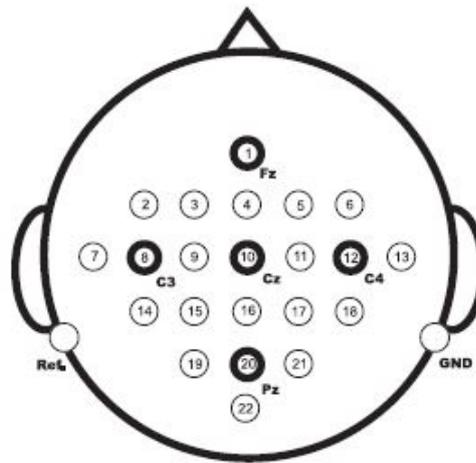


Figura 6 – Arranjo dos eletrodos utilizado na aquisição dos sinais de EEG da base de dados. Os eletrodos são numerados de 1 à 25, sendo os 22 primeiros referentes aos sinais de EEG e os três últimos aos sinais de eletro-oculogramas.

3.2 Pré-Processamento

Cada trecho de IM possui aproximadamente 7.5 s de duração. Os eventos de imaginação de movimento tem início no instante $t=0$ s e fim em $t=3$ s. Na Figura 7 é possível observar a média dos trechos de EEG do canal C3 para o indivíduo 1, onde se destaca o intervalo no qual há IM da mão direita.

Na base de dados utilizada neste trabalho, foi conduzida uma inspeção visual feita pela equipe pesquisadora e os trechos com artefatos foram marcados no arquivo de informações. Uma das etapas de pré-processamento deste trabalho foi a realização uma normalização no número para que todas as classes tivessem o mesmo número de observações, para tanto, foram removidos os trechos marcados como contendo artefatos de EOG. O número de observações por classe podem ser vistos na Tabela 2.

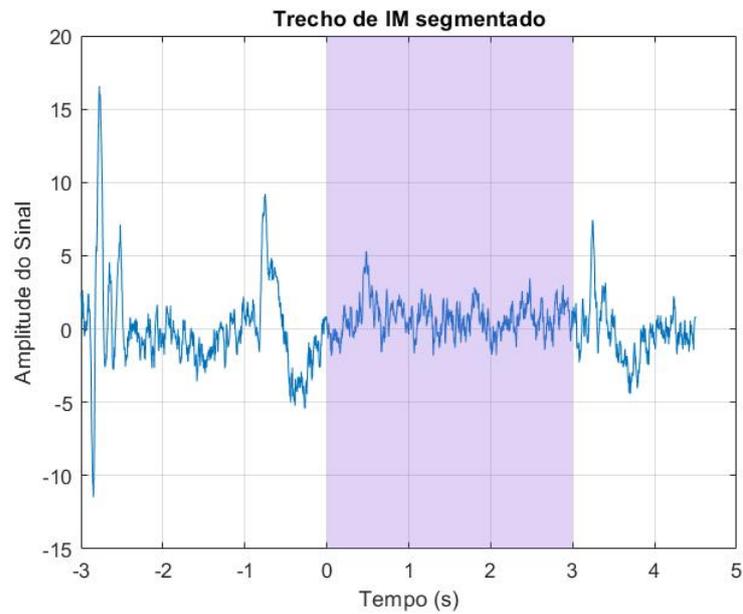


Figura 7 – Trecho de evento de IM MD para indivíduo 1 no canal C3 segmentado. Em destaque, observa-se o intervalo onde há IM MD.

Tabela 2 – Número de observações após normalização de classes. Foram removidos os trechos sinalizados como contendo artefatos no arquivo de informações da base de dados. Desta forma, para cada indivíduo, são consideradas o mesmo número de observações para todas as quatro classes de IM nos dados de treinamento e de teste.

Ind	IM ME		IM MD		IM PÉ		IM LI	
	Trein.	Teste	Trein.	Teste	Trein.	Teste	Trein.	Teste
1	67	69	67	69	67	69	67	69
2	66	69	66	69	66	69	66	69
3	66	67	66	67	66	67	66	67
4	62	53	62	53	62	53	62	53
5	63	65	63	65	63	65	63	65
6	49	53	49	53	49	53	49	53
7	66	66	66	66	66	66	66	66
8	64	66	64	66	64	66	64	66
9	53	65	53	65	53	65	53	65

3.2.1 Segmentação dos Dados

A Segmentação dos sinais é feita de $t=-3$ s à $t=4.5$ s. O intervalo tem um total de 1875 amostras, e a IM ocorre no intervalo de $t=0$ s à $t=3$ s. Na Figura 8, observa-se a média das observações de cada um dos 22 eletrodos de EEG para cada uma das quatro classes.

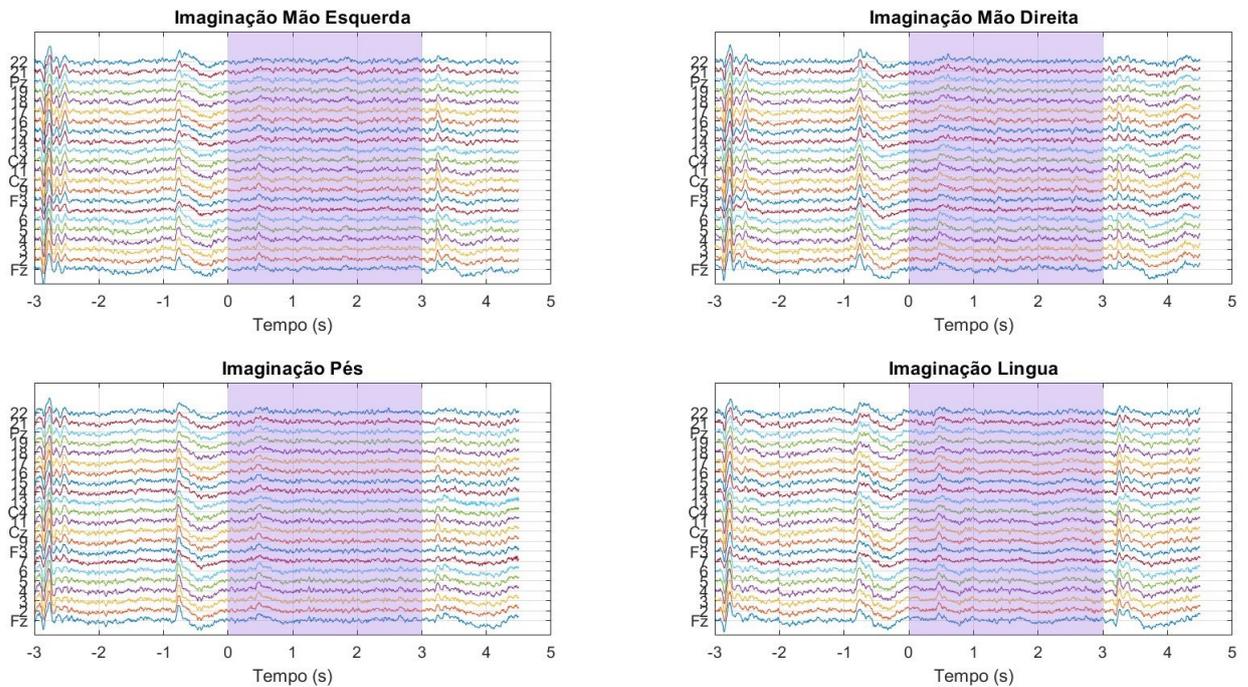


Figura 8 – Média das observações de cada classe para os eletrodos de EEG. Destacados em roxo, estão os intervalos onde ocorrem os eventos de IM.

3.2.2 Filtragem Espacial

Como filtragem espacial foi utilizada a técnica de promediação de referência comum (CAR do inglês *Common Average Reference*) é um método de filtragem espacial utilizado para o referenciamento dos sinais. Sua aplicação consiste em subtrair do sinal de um eletrodo S_i , onde i representa a i -ésima posição do eletrodo, a estimação da média de dos sinais registrados, exceto do eletrodo em questão. (SYAM et al., 2017).

$$S_i^{CAR}(t) = S_i(t) - \frac{1}{M-1} \sum_{t \neq i}^M s(t) \quad (3.1)$$

Onde M é o número de canais de EEG do arranjo. Para casos nos quais os eletrodos estão regularmente espaçados, a aplicação desse método gera uma distribuição de potencial de média zero, além de aumentar a relação sinal ruído e melhorar respostas locais (LUDWIG et al., 2009). Na Figura 9 são mostradas as médias dos sinais dos eletrodos após a aplicação do algoritmo, para cada uma das classes de evento de IM.

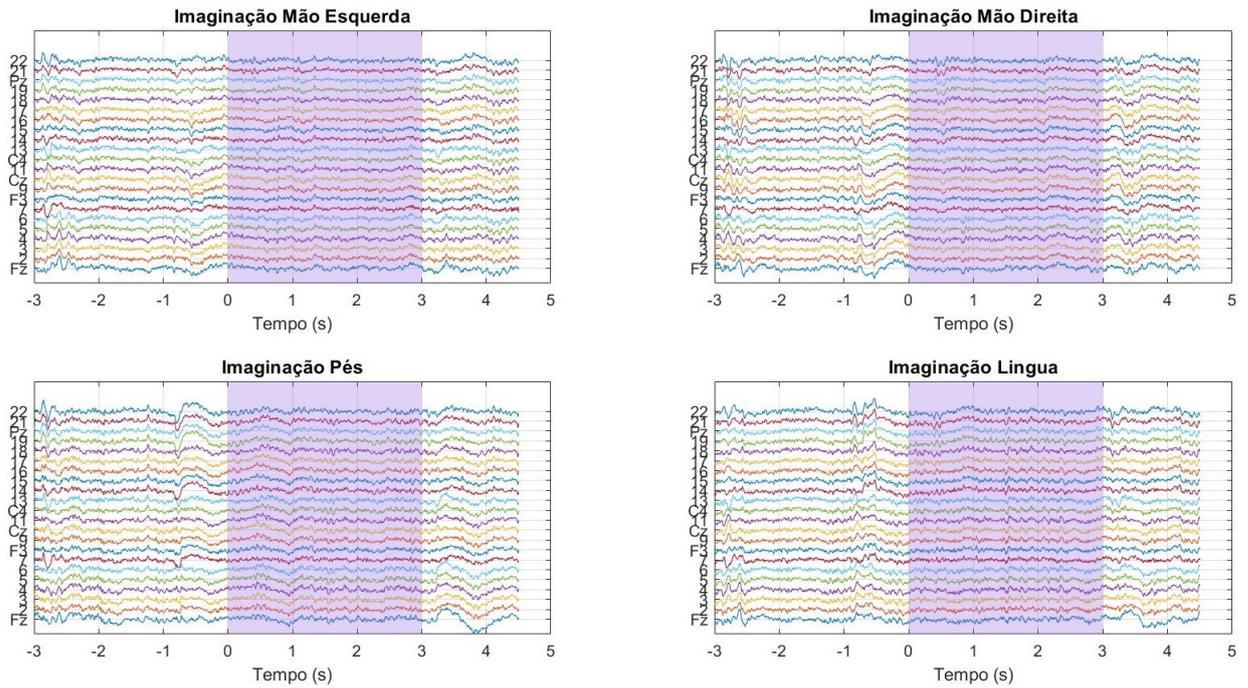


Figura 9 – Média das observações de cada classe para os eletrodos de EEG após a aplicação do referenciamento por CAR. Destacados em roxo, estão os intervalos onde ocorrem os eventos de IM.

3.3 Extração de Características

Na etapa de extração de características são computadas, a partir dos padrões nos ritmos cerebrais causados pelos eventos motores, atributos de fase, potência, ou estatísticos.

3.3.1 *Phase Locking Value*

Sincronização é a medida da relação entre sinais no domínio do tempo, sem levar em consideração suas amplitudes. Dois sinais são ditos síncronos se seus ritmos coincidem (VARELA et al., 2001). Uma das formas utilizadas para quantificação da sincronização em ICMs é o valor de bloqueio de fase (PLV, do inglês *Phase Locking Value*). Dados os sinais de dois eletrodos $S_1(t)$ e $S_2(t)$, é possível calcular a relação entre as fases dos sinais adquiridos e determinar, numericamente, o acoplamento entre suas fases (WANG et al., 2006).

Para tanto, utiliza-se a Transformada Hilbert, $H\{\}$, definida como a convolução do sinal $S_i(t)$ com o filtro de resposta ao impulso $1/\pi t$ (KSCHISCHANG, 2006).

$$H\{S(t)\} = \int_{-\infty}^{\infty} \frac{S(\tau)}{\pi(\tau - t)} \quad (3.2)$$

Assim, define-se o sinal analítico:

$$H\{S_i(t)\} = S_i(t) + jH\{S_i(t)\} \quad (3.3)$$

Considerando-se o sinal $S_i(t)$ segmentado em trechos com L amostras e sobreposição s e uma janela $w[n]$, tem-se

$$x_0[n] = x[0 : L - 1]w[0 : L - 1] \quad (3.7)$$

$$x_1[n] = x[L - s : 2L - 1 - s]w[L - s : 2L - 1 - s] \quad (3.8)$$

$$x_k[n] = x[k(L - s) : (k + 1)L - 1 - ks]w[k(L - s) : (k + 1)L - 1 - ks] \quad (3.9)$$

Em seguida, calcula-se a transformada rápida de Fourier dos sinais janelados:

$$X_k(j\omega) = \frac{1}{L} \sum_{n=0}^{L-1} x_k[n]w_k[n]e^{-j2\pi\omega n} \quad (3.10)$$

Por fim, calcula-se a média da potência dos K trechos de $X_k(j\omega)$:

$$P_{xx} = \frac{L}{UK} \sum_{k=0}^K |X_k(j\omega)|^2 \quad (3.11)$$

Onde U é a potência da janela aplicada a cada trecho, de valor constante.

3.3.3 Padrões Espaciais Comuns

Padrões Espaciais Comuns (CSP do inglês *Common Spatial Patterns*) é um método de filtragem espacial utilizado em ICMs que realiza uma transformação que maximiza a diferença de variância entre duas classes (MEISHERI; RAMRAO; MITRA, 2016).

Dado um sinal de EEG S com dimensões $M \times N \times T$, onde M é o número de canais, T o número de amostras dos sinais e N o número de observações. Calcula-se, para cada classe C , a matriz de covariância normalizada X_c , com $c \in C$, de dimensões $M \times M$, à partir dos sinais S_c correspondentes a cada classe.

$$X_c = \frac{S_c S_c'}{\text{traço}(S_c S_c')} \quad (3.12)$$

A função *traço*() calcula a soma dos elementos da diagonal principal da matriz. Calcula-se as matrizes \overline{X}_c , que são as médias de todas as observações para cada classe.

As transformações lineares do método CSP são baseadas na diagonalização simultânea das matrizes \overline{X}_c (WANG; GAO; GAO, 2006).

$$X = \overline{X}_1 + \overline{X}_2 = U_0 \Sigma U_0' \quad (3.13)$$

onde U_0 é a matriz de autovetores e Σ é a matriz cujos valores da diagonal principal são os autovalores de X .

Desta forma, encontra-se uma matriz P com a qual é possível transformar os dados na direção da matriz média, seguido de um escalonamento nos valores das observação. A matriz P é chamada de matriz de branqueamento. (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999).

$$P = \sqrt{\Sigma}U_0' \quad (3.14)$$

Quando a matriz de branqueamento P é aplicada às matrizes de covariância médias de cada classe separadamente, ambas as classes terão os mesmos eixos principais e seus autovalores terão soma unitária, Assim, quando:

$$M_2 = U\Sigma_2U' \quad (3.15)$$

$$M_2 = U\Sigma_2U' \quad (3.16)$$

Portanto:

$$\Sigma_1 + \Sigma_2 = I \quad (3.17)$$

Consequentemente, as direções com maiores autovalores e variância para uma classe, correspondem às direções de menores autovalores e variância para a outra classe e vice-versa. Desta forma, pode-se escolher os m maiores e m menores autovalores, pois essas direções apresentam maior discriminação entre tarefas (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999). Para transformação dos dados de EEG, S , utiliza-se a matriz de transformação W :

$$W = U'P \quad (3.18)$$

Para a filtragem espacial, tem-se:

$$Z = WS \quad (3.19)$$

Onde Z é a matriz $N \times T$ após a filtragem espacial, W é a matriz de transformação $A \times A$ e S é a matriz de dados brutos $N \times T$.

Como características para a discriminação, utiliza-se a variância normalizada das m primeiras e m últimas colunas de Z , sendo comumente $m = 3$, assim:

$$f_p = \log \left(\frac{\text{var}(Z(p))}{\sum_{p=1}^{2m} \text{var}(Z(p))} \right) \quad (3.20)$$

Onde f_p tem dimensões $1 \times 2m$, e a transformação logarítmica é aplicada para que a distribuição de características f_p seja normal.

3.3.4 Filter Bank Common Spatial Pattern

O método Filter Bank Common Spatial Pattern (FBCSP) visto na Figura 11, é composto de duas etapas: Filtragem na frequência e Filtragem espacial pelo algoritmo CSP. Esse método é apresentado como uma alternativa para para a otimização de seleção de banda para cada indivíduo.

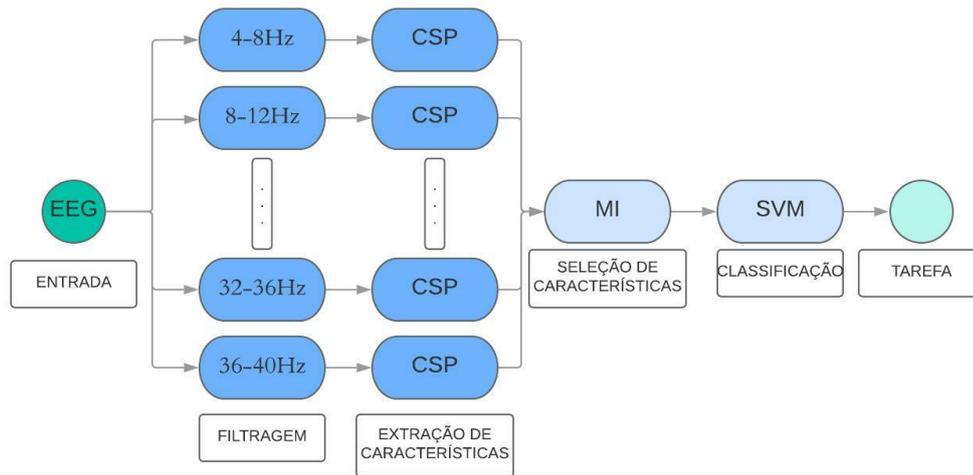


Figura 11 – Metodologia de extração de características por FBCSP. O sinal de EEG de entrada é filtrado de 4-40Hz em bandas de 4Hz sem sobreposição, em seguida, é realizada a extração de características por CSP multiclasse para a saída de cada filtro.

A etapa de extração de características é composta de uma série de filtros passa banda Butterworth de ordem 4, com 4 Hz de banda, que filtram o sinal de entrada de 4 à 40 Hz na primeira etapa, Figura 12. Em seguida, o algoritmo de CSP é aplicado à saída de cada filtro.

O conjunto de características resultante f_p tem dimensões $N \times \sigma 2m$, onde N é o número de observações, σ é o número de filtros e m é o número de características utilizadas da filtragem espacial.

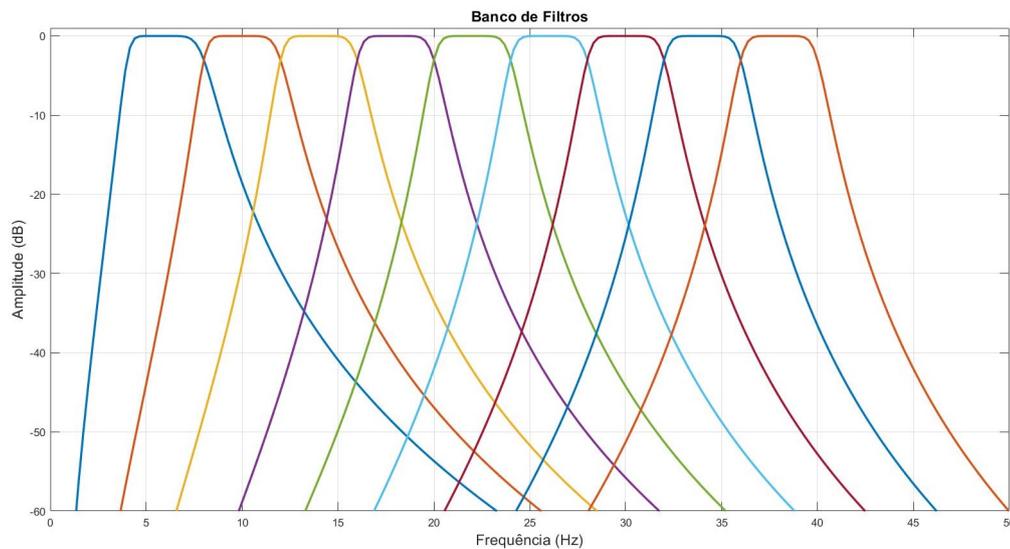


Figura 12 – Banco de Filtros aplicado aos sinais de EEG. O banco é composto por filtros Butterworth de ordem 4, de 4-40Hz com bandas de 4Hz sem sobreposição.

3.4 Estratégias CSP Multiclasse

A principal limitação do método CSP é o fato de que só se pode garantir uma matriz de transformação W que maximize a discriminação entre classes para dados binários, onde só existem duas classes. Portanto, são necessárias adaptações para uma discriminação multiclasse, alguns métodos são sugeridos na literatura como Diagonalização Simultânea (GROSSE-WENTRUP; BUSS, 2008), Par a Par (PW do inglês *Pairwise*), (CHIN et al., 2009), Um versus Resto (OxR do inglês *One versus Rest*), e Dividir e Conquistar (ANG et al., 2012) (GROSSE-WENTRUP; BUSS, 2008).

3.4.1 CSP Dividir e Conquistar

No método CSP Dividir e Conquistar (CSP DC), para o presente trabalho, tem-se um conjunto de classes $\omega = \{1, 2, 3 \text{ e } 4\}$, sendo as tarefas de IM da ME, MD, PÉ e LI, respectivamente. Para as quatro classes da base de dados, são necessários 3 modelos para a discriminação multiclasse ou, $C - 1$ modelos, sendo C o número de classes.

Os modelos são definidos aplicando-se o método CSP binário entre uma classe e as demais. Desta forma, é criado um conjunto $\omega' = \{\alpha_1, \beta_1\}$, onde, por exemplo, $\alpha_1 = \{1\}$ e $\beta_1 = \{2, 3, 4\}$. Sendo a classe representada por α_1 , retirada da estimação do próximo modelo. Assim, para o primeiro filtro espacial tem-se $\omega'_1 = \{\alpha_1, \beta_1\}$ com $\alpha_1 = \{1\}$ e $\beta_1 = \{2, 3, 4\}$, para o segundo, $\omega'_2 = \{\alpha_2, \beta_2\}$ com $\alpha_2 = \{2\}$ e $\beta_2 = \{3, 4\}$ e, por fim, para o terceiro, $\omega'_3 = \{\alpha_3, \beta_3\}$ sendo $\alpha_3 = \{3\}$ e $\beta_3 = \{4\}$. O processo é mostrado na Figura 13.

Por fim, as etapas de modelagem para ω'_1 , ω'_2 e ω'_3 , resultam em um conjunto de

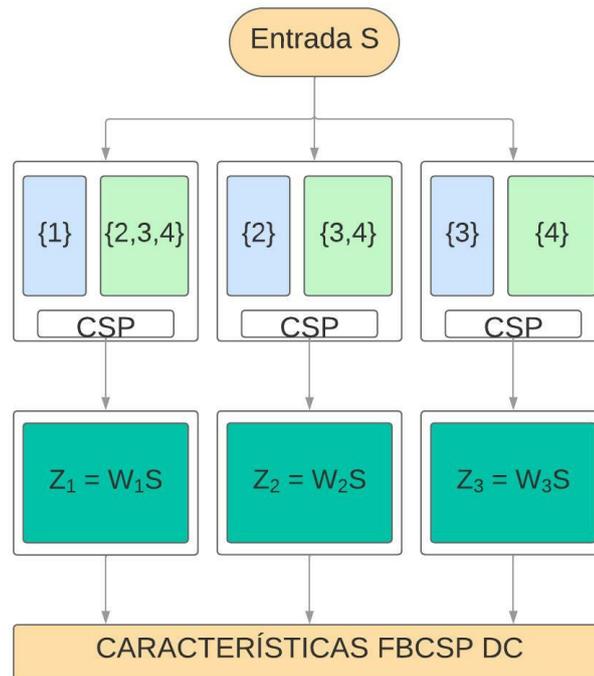


Figura 13 – Modelo CSP DC. Dos sinais de EEG de entrada S são extraídas características CSP binárias através de C-1 modelos, onde C é o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.

matrizes de transformação $W = \{w_1, w_2, w_3\}$, que são aplicadas ao conjunto de dados de EEG S , resultando em conjuntos Z_1, Z_2 e Z_4 que são, por fim, concatenados em um único vetor de características por observação.

3.4.2 CSP Um versus Resto

Na abordagem CSP OxR são estimados, pelo método CSP binário, modelos para discriminação entre uma classe e todas as outras (ANG et al., 2012). Assim, para o modelo onde $\omega'_1 = \{\alpha_1, \beta_1\}$ com a classe $\alpha_1 = \{1\}$, tem-se $\beta_1 = \{2, 3, 4\}$, $\omega'_2 = \{\alpha_2, \beta_2\}$ com $\alpha_2 = \{2\}$ e $\beta_2 = \{1, 3, 4\}$, $\omega'_3 = \{\alpha_3, \beta_3\}$ com $\alpha_3 = \{3\}$, tem-se $\beta_3 = \{1, 3, 4\}$ e $\omega'_4 = \{\alpha_4, \beta_4\}$ com a classe $\alpha_4 = \{4\}$ e $\beta_4 = \{1, 2, 3\}$. Desta forma, são necessários um modelo para cada classe analisada. Na Figura 14, observa-se a estrutura do método de extração de características CSP OxR para as quatro classes utilizadas neste trabalho.

Por fim, cada um dos modelos $\omega'_1, \omega'_2, \omega'_3$ e ω'_4 resulta em uma matriz de transformação $W = \{w_1, w_2, w_3, w_5\}$ que são aplicadas ao conjunto de dados de EEG S , resultando em conjuntos de características Z_1, Z_2, Z_3 e Z_4 que são, por fim, concatenados em um único vetor de características para cada observação.

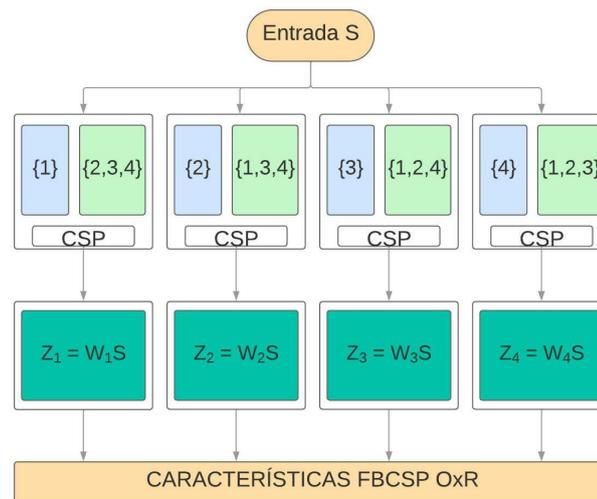


Figura 14 – Modelo CSP OXR. Dos sinais de EEG de entrada S são extraídas características binárias através de C modelos, sendo C o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.

3.4.3 CSP Par a Par

Para o método CSP Par a Par (CSP PP), ou *CSP Pair Wise* (CSP PW), são estimados, através da aplicação de CSP binário para cada combinação de duas classes. Desta forma, para um conjunto com C classes, tem-se $C!/2!(C-2)!$ modelos. Desta forma, para as quatro classes analisadas, devem ser estimados 6 modelos, como visto na Figura 15.

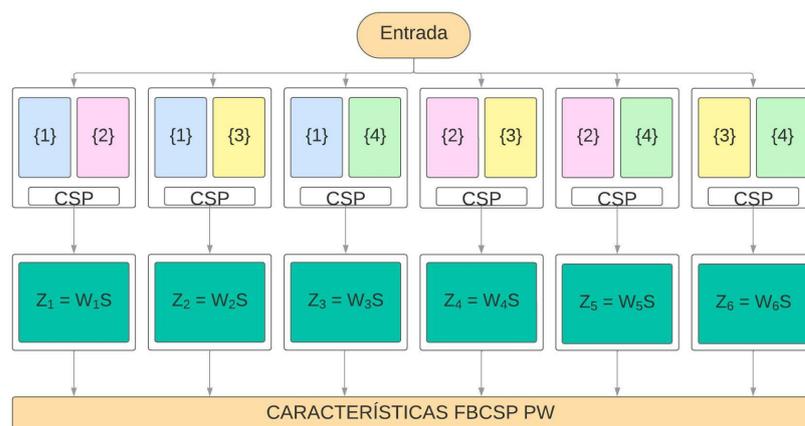


Figura 15 – Modelo CSP Par a Par. Dos sinais de EEG de entrada S são extraídas características binárias através de C modelos, sendo C o número de classes. Cada modelo resulta em uma matriz de transformação w que, aplicada aos dados de EEG S , gerando um conjunto de características Z . Por fim, as características dos modelos são concatenadas em um único vetor de características.

É estimada uma matriz de transformação w para cada modelo de CSP binário que, aplicadas aos dados de EEG de entrada S resultam em um conjunto de características Z para cada par de classes. Essas características são, por fim, concatenadas em um único vetor de características para cada observação.

3.4.4 CSP Diagonalização Simultânea

O algoritmo CSP clássico é capaz de maximizar a variância entre os dados de uma classe em relação à outra, para o caso de classificação binária. Entretanto, a diagonalização simultânea das matrizes de covariância não é aplicável para mais de duas classes. Uma possível abordagem para este problema é a aproximação da diagonalização simultânea por métodos de otimização (LIYANAGE et al., 2010).

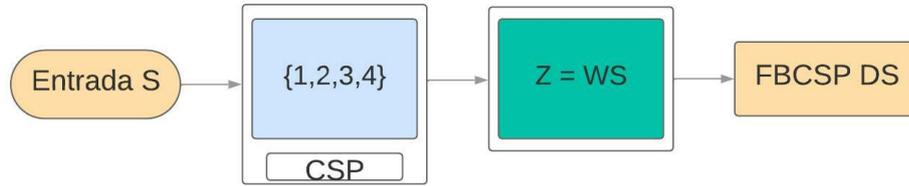


Figura 16 – Modelo CSP DS. As características são extraídas através de um único modelo do qual se estima uma matriz de transformação W que, aplicada aos dados de EEG de entrada resultando em um conjunto de características Z .

Sendo o conjunto $X_C = \{X_1, X_2, \dots, X_c\}$, composto por C matrizes de covariância $M \times M$ e $\mathbf{off}()$ a soma dos elementos fora da diagonal principal das matrizes do conjunto.

$$\mathbf{off}(X_C) = \sum_{i \neq j} |X_{1i,j}|^2 + \sum_{i \neq j} |X_{2i,j}|^2 + \dots + \sum_{i \neq j} |X_{ci,j}|^2 \quad (3.21)$$

O procedimento de diagonalização simultânea visa minimizar a soma dos elementos fora da diagonal principal das matrizes das matrizes $U \Sigma_c U^T$. Afim de construir uma matriz U , o método de Jacobi aplica uma série de rotação de planos no conjunto de matrizes X_C , através da multiplicação por uma matriz $R(i, j, c, s)$ (CARDOSO; SOULOUMIAC, 1996).

$$R(i, j, c, s) = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \quad (3.22)$$

Onde $c \in R^+$, $s \in C$ e $|c|^2 + |s|^2 = 1$. Sendo esta última condição, importante para a convergência do método para a solução ótima. Desta forma, o método iterativo busca minimizar a função $O(c, s)$ tal que:

$$O(c, s) = \sum_{k=1:K} \mathbf{off} \left(R(i, j, c, s) X_c R(i, j, c, s)^T \right) \quad (3.23)$$

Para quaisquer valores de (i, j) , pode-se definir uma matriz $G(X_c)$ tal que:

$$G(X_c) = \mathbf{Real} \left(\sum_{c=1, C} h^T(X_c)h(X_c) \right) \quad (3.24)$$

$$h(X_c) = [x_{cii} - x_{cjj}, x_{cij} + x_{cji}, i(x_{cji} - x_{cij})] \quad (3.25)$$

Para o caso da diagonalização simultânea de matrizes reais e simétricas, como é o caso das matrizes de covariância, os parâmetros c e s são ambos reais. Assim, o termo complexo da matriz $h(X_c)$ é igual a zero, e G tem dimensões 2x2 (LIYANAGE et al., 2010).

Pode-se parametrizar $c = \cos(\theta)$ e $s = e^{i\gamma \text{sen}(\theta)}$ sendo $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ (BUNSEGERSTNER; BYERS; MEHRMANN, 1993). Porém, para casos em que as matrizes X_c são reais, $\gamma = 0$ e $s = \text{sen}(\theta)$, sendo assim, um problema de otimização de uma variável apenas.

O parâmetro θ é calculado através do autovetor $[x, y]$ associado ao maior autovalor de G :

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{x}{y + \sqrt{x^2 + y^2}} \right) \quad (3.26)$$

O método iterativo termina quando o valor de θ atinge um limiar pré determinado.

3.5 Seleção de Características

Uma vez realizada a extração das características, uma maneira de diminuir o custo computacional da classificação e/ou aumentar o desempenho na discriminação entre classes utilizando técnicas de seleção de características. Para este trabalho, serão utilizadas as técnicas de filtros com critérios de Máxima Relevância e Mínima Redundância baseados em Informação Mútua (MI, do inglês *Mutual Information*).

3.5.1 Máxima relevância e Mínima redundância

Dadas duas variáveis aleatórias X e Y , com distribuição de probabilidade conjunta $p(x, y)$ e distribuições marginais $p(x)$ e $p(y)$, pode-se definir como entropia $H(x)$ e $H(y)$ como a incerteza relacionada a essas variáveis (BERAHA et al., 2019). Informação mútua pode ser interpretada como a redução da incerteza de uma variável X quando observada

uma variável Y ao considerar-se entropia relativa entre as probabilidades conjunta e marginais dessas distribuições. (CHOW; HUANG, 2005), assim:

$$H(X) = - \sum_x p(x) \log(p(x)) \quad (3.27)$$

$$H(X, Y) = - \sum_{x,y} p(x, y) \log(p(x|y)) \quad (3.28)$$

Onde $H(x)$ é a entropia de x com $x \in X$, e $H(x|y)$ é a entropia conjunta de x e y com $x \in X$ e $y \in Y$.

Portanto, pode-se definir Informação Mútua em termos da diferença entre a entropia marginal e a entropia conjunta das variáveis:

$$I(X; Y) = H(x) - H(x|y) \quad (3.29)$$

Ou então por:

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (3.30)$$

O diagrama da Figura 17, fornece uma representação gráfica dessa relação.

A principal ideia por trás dos métodos de máxima relevância e mínima redundância, (*mRMR*), é busca por um subconjunto ótimo ξ onde as características sejam pouco relacionadas entre si, ou seja, que carreguem pouca informação mútua em relação às outras características do subconjunto, e que ao mesmo tempo, sejam fortemente relacionadas com os rótulos das classes, ou seja, tendo altos valores de informação mútua em relação aos rótulos. Desta forma, redundância pode ser definida como o somatório das informações mútuas entre as características que compõe esse conjunto ξ :

$$R = \frac{1}{|\xi|^2} \sum_{x_i x_j \in S} I(x_i, x_j) \quad (3.31)$$

Ao realizar a minimização do valor da redundância $R(\xi)$, tem-se o subconjunto no qual as características são mais independentes entre si e com a maximização da medida global da relevância $D(\xi)$, obtem-se o subconjunto mais relacionado com os rótulos de classes c .

$$D = \frac{1}{|\xi|} \sum_{x_i \in S} I(x_i, c) \quad (3.32)$$

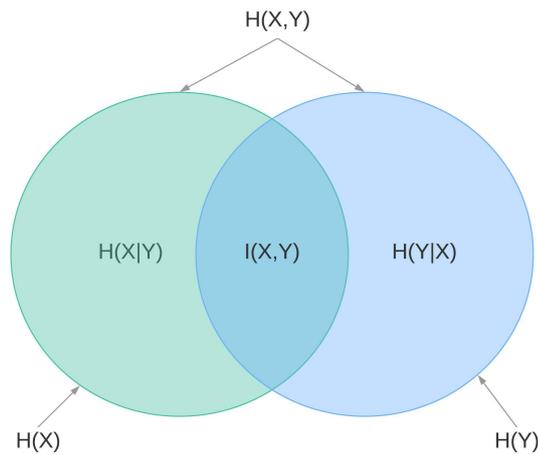


Figura 17 – Relação entre Informação Mútua e Entropia: O diagrama representa a relação entre medidas de informação associadas às variáveis X e Y. A área de intercessão entre os diagramas representa a informação mútua $I(X,Y)$ entre as variáveis X e Y. Os círculos representam, individualmente, as medidas de entropia de cada variável $H(X)$ e $H(Y)$.

Por fim, o resultado do método *mRMR*, assinalado por ψ , consiste em maximizar a diferença entre relevância D e redundância R , afim de se selecionar o subconjunto ótimo de características (PENG; LONG; DING, 2005).

$$\max \psi(D, R) = \max(D - R) \tag{3.33}$$

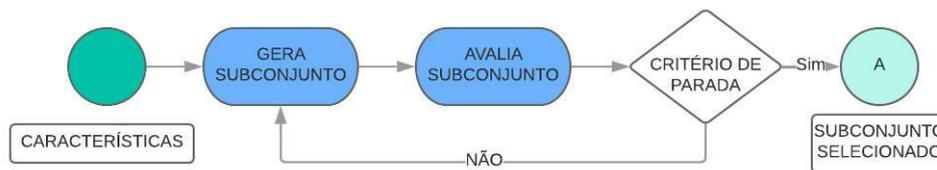


Figura 18 – Método de seleção de características: Ao conjunto de características, inicialmente vazio, são adicionadas de forma sequencial as características uma a uma, cada subconjunto é avaliado de acordo com os critérios de mRMR. Caso um subconjunto seja melhor avaliado em relação ao anterior, define-se um novo subconjunto. O procedimento se repete até que o critério de parada seja atingido e um subconjunto ótimo seja definido.

Desta forma, foi implementado um filtro para seleção de características, Figura 18, utilizando como critérios de avaliação a redundância R e a relevância D de um subconjunto ξ aleatório de características. Por fim, encontrar o melhor subconjunto ξ consiste em um problema de otimização para maximizar o valor de $\psi = D - R$.

O algoritmo implementado Algoritmo 1, calcula a relevância R_i de todas as características de um subconjunto ξ através da informação mútua entre as características e os rótulos de classe. Uma busca sequencial direta é realizada, onde as características são adicionadas uma a uma. Calcula-se a informação mútua entre os pares de características do subconjunto ξ afim de se determinar quais características possuem mais informação mútua entre si. Por fim, calcula-se o valor de ψ para o subconjunto avaliado, caso seja maior que o anterior, um novo subconjunto ótimo é selecionado.

Algorithm 1 Busca de Subconjunto de Características

Require: Dados, labels

Ensure: *CaracSel*

$S \leftarrow \emptyset$

contador $\leftarrow 0$

CaracSel $\leftarrow \emptyset$

numCaracterísticas $\leftarrow \text{Tamanho}(\text{Dados})$

$\psi \leftarrow 0$

enquanto *contador* < *numCaracterísticas* **faça**

$S \leftarrow \text{concatena}(\text{CaracSel}, \text{contador})$

$D \leftarrow \sum_{i \in S} \frac{\text{calculaInformaçãoMutua}(S_i, \text{labels})}{|S|}$

$R \leftarrow \sum_{i, j \in S} \frac{\text{calculaInformaçãoMutua}(S_i, S_j)}{|S|^2}$

$\psi_{aux} \leftarrow D - R$

se $\psi_{aux} > \psi$ **então**

$\text{CaracSel} \leftarrow S$

senão

CaracSel não muda

fim se

$\psi \leftarrow \psi_{aux}$

contador $\leftarrow \text{contador} + 1$

fim enquanto

3.6 Classificação

Nesta etapa, um modelo matemático é utilizado para realizar a discriminação entre as classes analisadas, produzindo como resultado, uma estimativa da classe à qual pertence uma observação de entrada. Neste trabalho é implementado um algoritmo de Máquina de Vetores de Suporte (SVM, do inglês *Support Vector Machine*), associado a um *kernel* gaussiano para separação não linear dos dados.

3.6.1 *Support Vector Machine*

Considerando as amostras de treinamento $\{s_i, c_i\}$ em que s_i é a i -ésima observação de entrada e c_i é a classe à qual essa amostra pertence com $c_i \in \{-1, 1\}$ ou seja, os padrões

são linearmente separáveis, é possível definir uma superfície de separação através da seguinte equação:

$$\omega^T s_i + b = 0 \quad (3.34)$$

Nessa equação s_i representa o vetor de características e ω é um vetor de pesos ajustáveis. Essa Superfície de separação é chamada de *hiperplano*, e sua definição mais geral para um espaço L -dimensional é uma superfície de $L - 1$ dimensões que divide o espaço em duas partes (HAYKIN, 1994).

Para (HAYKIN, 1994), supondo-se a existência de um hiperplano de separação, pode-se implementar um classificador que atribui a uma classe a uma observação, baseada em qual lado do hiperplano a amostra se encontra através da avaliação do sinal da equação $f(s_i) = \omega^T s_i^* + b$. Caso positivo, é atribuída à observação a classe 1, caso contrário, é atribuída a classe -1. A formulação dessa suposição é a seguinte:

$$\omega^T s_i + b \geq +1 \text{ para } c_i = +1 \quad (3.35)$$

$$\omega^T s_i + b \leq -1 \text{ para } c_i = -1 \quad (3.36)$$

Pode-se ainda, simplificar as equações 3.35 e 3.36 combinando-as da seguinte forma:

$$c_i(\omega^T s_i) - 1 \geq 0 \quad \forall i \quad (3.37)$$

É possível ainda, utilizar-se da amplitude de $f(s_i)$ na estimativa da certeza da classificação, uma vez que valores maiores de $f(s_i)$ indicam uma maior distância entre a observação e o hiperplano e, conseqüentemente uma maior certeza na determinação atribuída à observação, enquanto valores menores, especialmente próximos a zero, indicam uma maior proximidade ao plano de separação, portanto menos certeza no resultado da classificação (JAMES et al., 2013).

Partindo ainda da suposição de que os dados podem ser perfeitamente separados por um hiperplano, deve-se determinar qual dos possíveis hiperplanos infinitos caracteriza a melhor superfície de separação. Esse problema pode ser resolvido com a escolha do hiperplano de margem máxima, ou hiperplano ótimo, no qual a superfície de decisão está o mais distante possível das observações de treinamento. A margem é definida pela menor distância entre o hiperplano e as observações de treinamento mais próximas, chamadas vetores de suporte. A definição do hiperplano ótimo se mostra muito sensível aos vetores de suporte, uma vez que a adição ou retirada de uma dessas observações poderia resultar em uma superfície de separação completamente diferente (JAMES et al., 2013), essa mesma sensibilidade não é observada em relação as outras amostras.

Para os casos de igualdade nas equações 3.35 e 3.36 os pontos estão exatamente sobre as margens de separação e são, portanto, vetores de suporte. As distâncias entre esses pontos e o hiperplanos são dadas por:

$$D_1 = D_2 = \frac{|1 - b|}{|\omega|} = \frac{|-1 - b|}{|\omega|} \quad (3.38)$$

Somando ambas as distâncias, tem-se o valor da margem de separação.

$$F = D_1 + D_2 = \frac{2}{|\omega|} \quad (3.39)$$

Dessa forma, o problema de maximização da margem se torna um problema de otimização quadrática para a minimização do valor de $|\omega|^2$ (BURGES, 1998). A satisfação desse critério garante que todas as amostras estarão, não só do lado correto do plano de decisão, mas também a pelo menos uma distância D_1 e D_2 , com uma separação F entre as fronteiras.

3.6.2 Dados linearmente separáveis com tolerância de violação de margem

O classificador de minimização da margem apresenta alto desempenho em casos separáveis porém, não tem bom desempenho para dados não separáveis (HAYKIN, 1994). Desta forma, segundo (JAMES et al., 2013), para esses casos deve-se relaxar as condições das equações 3.35 e 3.36 através da definição de parâmetros de tolerância de violação da margem $\chi_i \geq 0$ que permitem amostras individuais estarem do lado errado da margem. Assim, define-se um hiperplano que não separa duas classes perfeitamente, afim de se obter uma melhor classificação para a maior parte dos dados a custo de se classificar erroneamente algumas amostras. Desta forma, a equação 3.37 pode ser escrita para o caso não separável da seguinte forma:

$$c_i(\omega^T s_i) - 1 + \chi_i \geq 0 \quad \forall i \quad (3.40)$$

São utilizados multiplicadores de Lagrange para a resolução do problema de otimização, assim surge o parâmetro L_p que deve ser minimizado.

$$L_p = \sum_{i=1}^n \epsilon - \frac{1}{2} \sum_{i,j=1}^n \epsilon_i \epsilon_j c_i c_j s_i s_j \quad (3.41)$$

Assim, surgem as condições para os valores de α_i :

$$\sum_{i=1}^n \epsilon_i c_i = 0 \quad (3.42)$$

$$0 \leq \epsilon_i \leq V \quad (3.43)$$

Define-se também, um parâmetro $V = \sum_{i=1}^n \chi_i$ que representa a violação total aceitável para a determinação do hiperplano. É evidente que quando $V = 0$, não há violação da margem e portanto, tem-se um conjunto de dados linearmente separável.

A partir dos valores dos multiplicadores de Lagrange ϵ_i , são calculados os valores do vetor de pesos ω , e de bias, b .

$$\omega = \sum_{i=1}^n \epsilon_i c_i s_i \quad (3.44)$$

$$b = 1 - \omega^T s \quad (3.45)$$

3.6.3 Dados não linearmente separáveis

Os métodos de SVM são uma extensão dos métodos de classificação de dados não separáveis que realiza o aumento de dimensão das características, afim de se encontrar um hiperplano de separação ótima nos moldes das equações 3.40, 3.42 e 3.43.

O aumento da dimensionalidade é feito através de uma operação de produto interno, que realiza o mapeamento das amostras do espaço de características de entrada m_0 para o espaço aumentado m_1 . Uma generalização dessa operação é:

$$K(s, s_i) = \Theta(s)\Theta(s_i) \quad (3.46)$$

Onde $\Theta(\cdot)$ denota o tipo de mapeamento a ser realizado, podendo ser por exemplo, linear, quadrático, radial, gaussiano etc. O mapeamento comumente é o que dá nome ao *kernel*.

Com a adição da transformação pelo *kernel* de produto interno, as equações 3.34, 3.44 e 3.45 devem ser reescritas de forma que utilizem o novo espaço de características.

$$\sum \omega_i \Theta(s_i) = 0 \quad (3.47)$$

$$\omega = \sum \epsilon_i c_i \Theta(s_i) \quad (3.48)$$

$$\sum \epsilon_i c_i \Theta^T(s_i) = 0 \quad (3.49)$$

As condições para a determinação dos coeficientes multiplicadores de Lagrange permanecem como das equações 3.42 e 3.43. A determinação do hiperplano de separação ótimo deve ser feita no espaço aumentado de características e por fim, as transformações devem ser feitas de maneira inversa, afim de se classificar os dados no espaço dimensional original.

3.7 Metodologia de Processamento

Inicialmente foram analisados métodos de extração de características binárias: FBCSP, CSP, PLV e Welch. Para a extração de características por CSP e FBCSP, os trechos de observação de IM foram segmentados, dentro do trecho de -0.5 à 4.5 s, em intervalos com duração de 2 s e sobreposição de 50%. Para o método CSP, os sinais foram filtrados na faixa de frequência de 8 à 30 Hz. Para a extração de características por FBCSP foi utilizada a mesma segmentação temporal, porém, aplica-se um banco de filtros de 4-40Hz, com bandas de 4 Hz sem sobreposição entre as bandas. Para cada filtro do banco de filtros, são escolhidos 2 m canais, com $m = 3$, representando 6 canais com maior capacidade de discriminação entre classes.

NA extração de características feita pelo método PLV, utilizam-se observações de IM segmentadas de 0 à 3 s em uma faixa de frequências de 8 à 30 Hz estimando-se a sincronização do eletrodos da região motora [C3, 9, Cz, 11, C4] com seus eletrodos adjacentes.

As características de potência, extraídas através do periodograma de *Welch*, são geradas à partir de trechos de IM segmentados de 0 à 3 s, com uma janela de *Hamming* de 1 s (250 amostras), e sobreposição de 500 ms (125 amostras).

Por fim, o sistema ICM implementado é composto por quatro etapas:

- *I* - Pré-processamento, composta pela remoção de artefatos, segmentação e referenciamento do arranjo por CAR,
- *II* - Extração de características, onde são aplicados os métodos CSP multiclasse descritos na Seção 3.4,
- *III* - Seleção de características, onde são gerados, avaliados e *rankeados* os subconjuntos de características e escolhido o subconjunto ótimo pelo critério mRMR
- *IV* - Classificação realizada pelo algoritmo *SVM Par a Par* com *kernel* gaussiano para determinar a qual classe uma observação de entrada pertence.

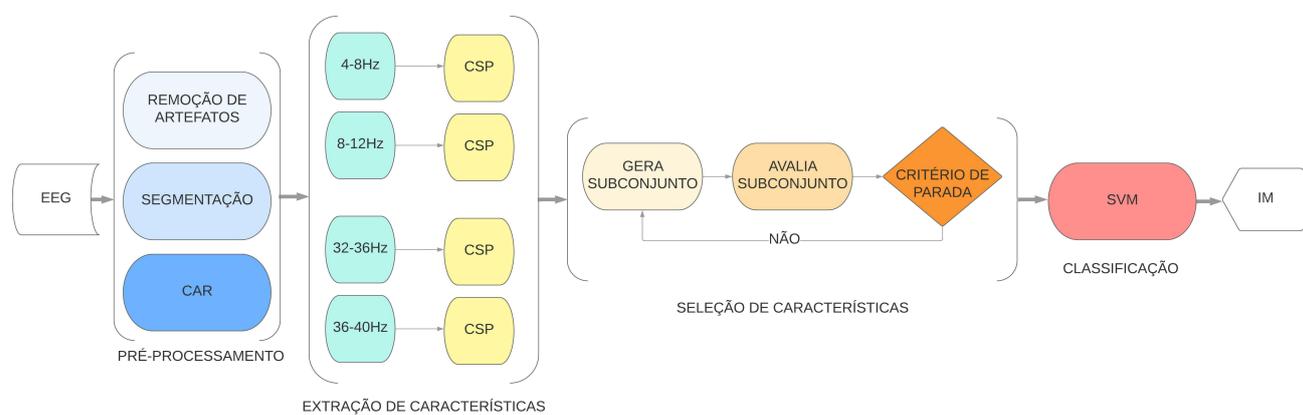


Figura 19 – Etapas da ICM implementada: *I* - Pré-processamento, *II* - Extração de características pelos métodos CSP multiclasse, *III* - Seleção de características por mRMR baseado em Informação Mútua *IV* - Classificação, com a realização pelo algoritmo *SVM Par a Par* com *kernel* gaussiano.

4 Resultados e Discussão

Neste Capítulo serão apresentados os resultados obtidos a partir da implementação das etapas anteriormente discutidas. A estimação dos desempenhos dos modelos foi realizada por uma validação cruzada de de 10-*fold* repetida 5 vezes (5x10-*fold*), para diminuir a chance de *overfitting*. Os modelos multi-dimensionais extraídos do conjunto de observações são visualizados em duas dimensões com o auxílio do método estatístico t-SNE.

4.1 Comparação FBCSP, CSP, PLV e Welch

A Tabela 3 apresenta os resultados da média das taxas de acerto por indivíduo para todas as combinações binárias: ME vs MD, ME vs PÉ, ME vs LI, MD vs PÉ, MD vs LI e PÉ vs LI. Este resultado mostra valores superiores para as abordagens CSP e FBCSP com taxas superiores a 90% para os indivíduos com melhor desempenho (1, 2, 7 e 8), e taxas superiores a 60% para os indivíduos com menor desempenho para o método CSP e 70% para o método FBCSP.

Tabela 3 – Taxas de acerto para os métodos binários de extração de características CSP, PLV e Welch: Foram extraídas características das combinações das classes par a par. Os valores representam a média da taxa de acerto para a classificação de seis modelos (4 classes). Em negrito estão destacadas as maiores taxas de desempenho por indivíduo e o maior valor médio para a taxa de acerto.

Taxas de acerto Classificação binária (%)				
Indivíduo	FBCSP	CSP	PLV	Welch
1	92.30	92.27	81.41	58.69
2	81.40	78.74	52.90	50.97
3	91.17	90.29	78.11	57.39
4	82.23	80.66	63.21	59.75
5	78.33	60.97	51.15	50.13
6	71.70	69.97	59.91	53.15
7	92.05	90.28	61.36	49.37
8	93.99	90.28	73.23	53.16
9	88.21	87.95	74.23	69.32
Média	84.93	82.71	66.17	55.75

Como se mostra na Figura 20, a média das taxas de todos os indivíduos da Tabela 3, existe uma superioridade para o método FBCSP e por esse motivo, as abordagens multiclasse serão realizadas utilizando este método, com a aplicação do banco de filtros.

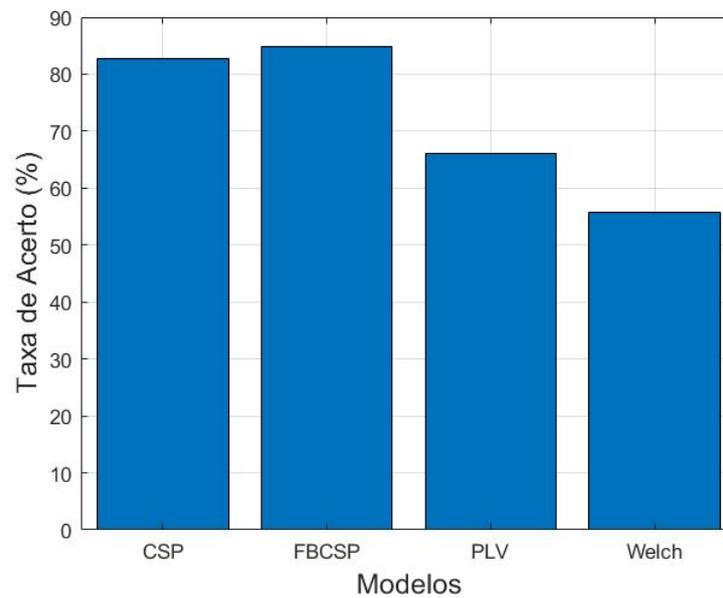


Figura 20 – Taxas de acerto para FBCSP, CSP, PLV, Welch: Média das taxas de acerto de 6 modelos binários extraídos da combinação de pares de classes (4 classes).

4.2 FBCSP

A implementação do algoritmo de extração de características por FBCSP multi classe foi realizada de quatro formas: DS, DC, PW e OxR.

Para o método FBCSP DS, as características para determinação do modelo são extraídas para todas as quatro classes de IM através da diagonalização simultânea das matrizes de covariância das classes. Para os métodos FBCSP DC, as características são extraídas em três etapas, nas quais são estimadas as matrizes de transformação W para cada etapa. As observações de entrada são transformadas por cada uma dessas matrizes. No método PW, as matrizes de transformação são extraídas de seis etapas de discriminação de duplas de características. Por fim, na abordagem OxR uma classe foi discriminada das três classes restantes, e as quatro matrizes de transformação foram aplicadas aos dados de entrada.

Tabela 4 – Distribuição dos modelos para os métodos: FBCSP DS - um modelo para discriminação das quatro classes de maneira simultânea. FBCSP DC - três modelos binários para discriminação. FBCSP PW - seis modelos binários para discriminação. FBCSP OxR - quatro modelos binários para discriminação

Modelos de Extração de Características	
DS	$\{ME, MD, PÉ, LI\}$
DC	$\{ME\}_x\{MD, PÉ, LI\}, \{MD\}_x\{PÉ, LI\}, \{PÉ\}_x\{LI\}$
PW	$\{ME\}_x\{MD\}, \{ME\}_x\{PÉ\}, \{ME\}_x\{LI\}, \{MD\}_x\{PÉ\}, \{MD\}_x\{LI\}, \{PÉ\}_x\{LI\}$
OxR	$\{ME\}_x\{MD, PÉ, LI\}, \{MD\}_x\{ME, PÉ, LI\}, \{PÉ\}_x\{ME, MD, LI\}, \{LI\}_x\{MD, ME, PÉ\}$

Desta forma, o conjunto de características no método FBCSP DS consiste em

216 características por observação. Para o método FBCSP DC, com três etapas, têm-se $216 \times 3 = 648$ características por observação. Para o método FBCSP PW têm-se 1296 características por observação e para o método FBCSO OxR têm-se 864 características por observação.

Tabela 5 – Número de características por Modelo após a concatenação das características extraídas de cada método

Número de Características por Modelo				
	FBCSP DS	FBCSP DC	FBCSP PW	FBCSP OxR
Nº características	216	648	1296	864

O desempenho de cada implementação, Tabela 6, foi estimado pela média dos coeficientes *kappa* após uma validação cruzada de 5×10 -fold. Todas as implementações tiveram desempenhos médios acima de 0.9700 para modelagem com 80% das amostras e teste com os 20% restantes, com destaque para o método DC que teve um desempenho médio $k = 0.9873$.

Tabela 6 – Coeficiente Kappa após a validação cruzada com 5×10 -fold para os quatro métodos propostos

Desempenho do Sistema (Kappa)				
Ind.	FBCSP DS	FBCSP DC	FBCSP PW	FBCSP OxR
1	0.9799	0.9950	0.9799	0.9801
2	0.9747	0.9899	0.9697	0.9546
3	0.9747	0.9848	0.9799	0.9648
4	0.9892	0.9946	0.9955	0.9784
5	0.9841	0.9870	0.9871	0.9947
6	0.9867	0.9933	0.9932	0.9795
7	0.9899	0.9801	0.9881	0.9949
8	0.9740	0.9793	0.9844	0.9843
9	0.9620	0.9814	0.9811	0.9873
Média	0.9795	0.9873	0.9843	0.9798

Como visto pelas taxas de desempenho, todos os modelos são eficientes na discriminação entre os diferentes tipos de eventos de IM analisados. Na Figura 21 observa-se a discriminação entre as classes utilizando o método FBCSP DS, com as observações de treinamento. Para a visualização de dados multidimensionais em duas dimensões foi utilizado o método *tsne*. Pode-se observar que existe variabilidade de resposta entre indivíduos, porém, é notável a possibilidade de discriminação entre as classes do problema.

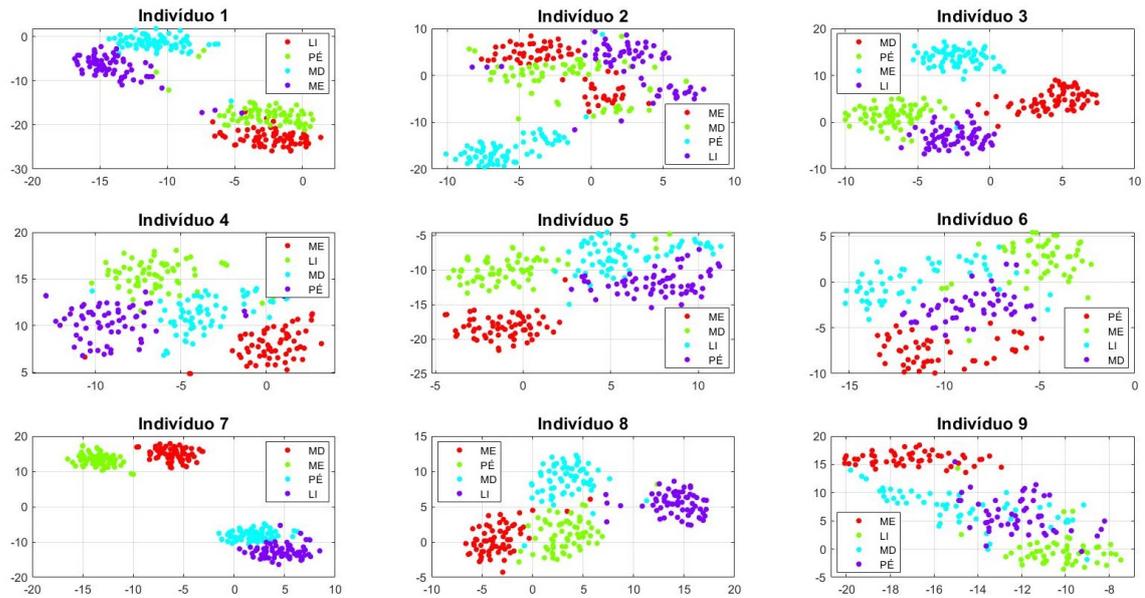


Figura 21 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das amostras de modelagem para o método FBCSP DS

4.3 FBCSP multiclasse com seleção de características

O algoritmo de seleção de características foi aplicado ao conjunto de dados para modelagem. Utilizando um algoritmo de busca sequencial, foram selecionadas as características mais relevantes. A Tabela 7 mostra o percentual de características selecionadas em relação ao conjunto total de características. Estes valores variam entre 43 e 50% com uma média de 46% de características selecionadas para cada modelo.

Tabela 7 – Porcentagem, em relação ao conjunto total, de características selecionadas pelo método *mRMR* para cada um dos métodos propostos e cada um dos indivíduos

Ind.	Características Selecionadas (%)			
	FBCSP DS	FBCSP O _x R	FBCSP PW	FBCSP DC
1	44.44	46.30	47.22	46.30
2	49.54	48.77	46.60	48.77
3	45.37	45.83	45.60	45.83
4	44.91	46.91	47.07	46.91
5	46.76	46.30	45.29	46.30
6	41.67	48.15	45.68	48.15
7	44.91	46.30	47.69	46.30
8	45.37	44.29	46.45	44.26
9	43.98	43.98	50.00	43.98
Média	45.21	46.31	46.84	46.31

4.3.1 Teste dos modelos com amostras cegas

Após a aplicação do método de seleção de características, os modelos são avaliados por seu desempenho na identificação de IMs com as observações de teste da base de dados. As observações passam pelo mesmo pré-processamento e são transformadas pelas matrizes derivadas da extração de características das observações de treinamento.

O desempenho do sistema foi avaliado com 100% das características e com os subconjuntos selecionados pelo método *mRMR*. A Tabela 8 mostra o desempenho dos modelos utilizando o coeficiente *kappa*. O desempenho dos métodos foram, em média, de 0.6489 para FBCSP PW, 0.6295 para o FBCSP OxR, 0.6161 para FBCSP DC e 0.5761 para FBCSP DS. Os modelos estimados a partir dos subconjuntos selecionados pelo método de filtragem para seleção de características *mRMR* apresentaram uma pequena melhoria de desempenho em relação aos modelos que utilizam o conjunto de 100% das características extraídas. Em negrito, na Tabela 8, observam-se os maiores valores de kappa entre os modelos com 100% das características e os modelos definidos a partir dos subconjuntos selecionados.

Tabela 8 – Desempenho dos sistemas utilizando as amostras cegas para os quatro métodos. É realizada uma comparação entre os modelos que utilizam 100% das características e os que utilizam os subconjuntos escolhidos pelo método de seleção. Em negrito estão destacados os maiores valores de desempenho em relação ao conjunto de características utilizado e sublinhado destacam-se os maiores valores de desempenho para cada indivíduo e da média geral.

Ind	Desempenho do Sistema (Kappa)							
	FBCSP DS		FBCSP PW		FBCSP OxR		FBCSP DC	
	Selec. Car.	100% Car.	Selec. Car.	100% Car.	Selec. Car.	100% Car.	Selec. Car.	100% Car.
1	0.7729	0.7391	0.7681	0.7440	0.7729	0.7295	0.7874	0.7536
2	0.5169	0.4686	0.4831	0.4831	0.4541	0.4203	0.5362	0.4396
3	0.7960	0.7214	0.7960	0.7711	0.7761	0.7463	0.8308	0.7910
4	0.2767	0.3208	0.5660	0.5409	0.5346	0.4717	0.4591	0.4088
5	0.4821	0.4615	0.5590	0.5897	0.5538	0.4769	0.5077	0.5590
6	0.2767	0.2893	0.2327	0.3145	0.3522	0.3396	0.2579	0.3208
7	0.7879	0.8081	0.8990	0.8535	0.8636	0.8586	0.8232	0.8182
8	0.7273	0.7374	0.8333	0.8131	0.7172	0.7475	0.7677	0.7273
9	0.5487	0.5436	0.7026	0.6923	0.6410	0.6154	0.5744	0.5795
Média	0.5761	0.5655	0.6489	0.6440	0.6295	0.6006	0.6161	0.5998

Na Tabela 8 estão sublinhados os maiores valores de kappa por indivíduo. Observa-se que para cinco dos nove indivíduos, o método FBCSP PW apresentou as maiores taxas de desempenho. Para para os indivíduos 4, 7, 8 e 9 o modelo FBCSP PW com o subconjunto de características selecionadas pelo método *mRMR* apresentou valores de kappa de 0.5660, 0.8990, 0.8333 e 0.7026, respectivamente e kappa igual a 0.5897 para o modelo FBCSP PW com 100% das características. Para os indivíduos 1, 2 e 3, o modelo

FBCSP DC com seleção de características apresentou melhor desempenho, sendo kappa igual a 0.7874, 0.5362 e 0.8308, respectivamente.

Na Figura 22 é possível observar uma representação em duas dimensões do conjunto de características multidimensionais extraídas pelo método FBCSP DS, após a aplicação do método de seleção de características. É possível observar como a discriminação é menor quando comparada às características extraídas das observações de treinamento, utilizadas para estimação dos modelos.

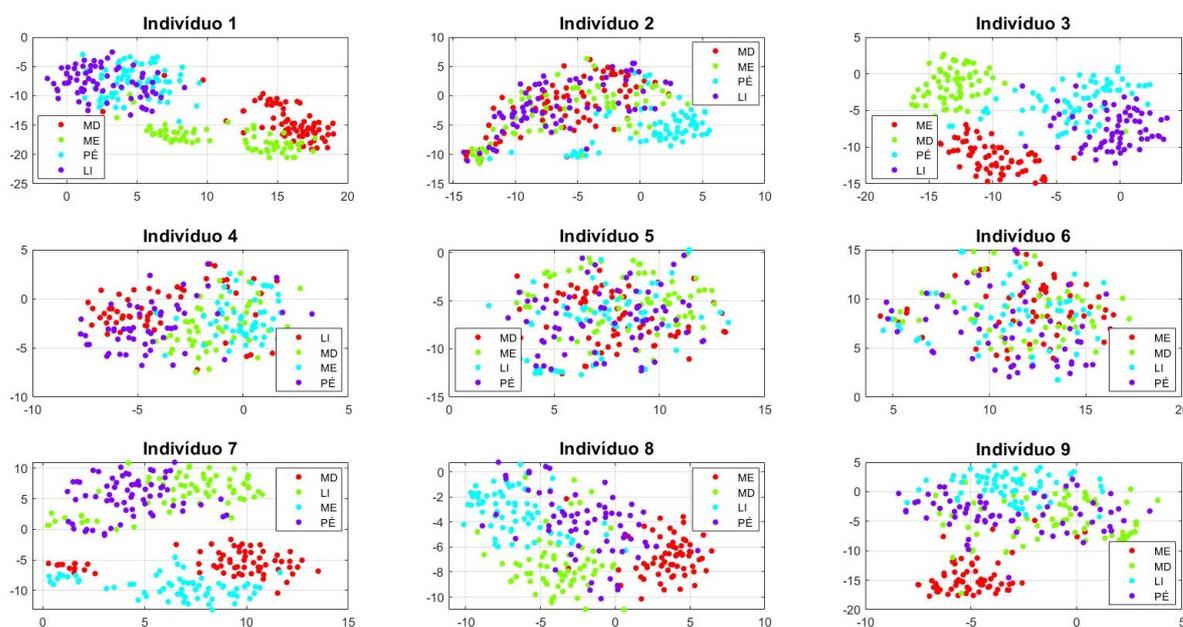


Figura 22 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP DS

Na Figura 23, têm-se as matrizes de confusão para cada indivíduo para método FBCSP DS. Para os indivíduos 1, 3, 7, 8 e 9, é possível ver que as taxas de acerto, diagonal principal das matrizes, é consideravelmente superior aos resultados falso positivo e falso negativo (outras posições das matrizes). Também para esses indivíduos, é possível perceber que há pares de classes de IM para as quais o sistema tem melhor desempenho, por exemplo, para o Indivíduo 1, o sistema é melhor em discriminar IM ME e IM PÉ, para o indivíduo 3, IM LI e IM MD, para o indivíduo 8 IM ME e IM MD e para o indivíduo 9 IM ME e IM PÉ. Para os Indivíduos 4, 5 e 6, há, para cada Indivíduo, uma classe em específico na qual é possível observar um melhor desempenho, sendo essas as classes LI, ME e LI, respectivamente.

Nos padrões de desempenho dos métodos FBCSP DC, PW e OxR, é possível notar que há uma alta capacidade de discriminação entre 3 classes para os indivíduos 1, 3, 7, 8 e

9 e entre 2 classes para os indivíduos 2, 4, 5 e 6. As matrizes de confusão dos sistemas podem ser vistas no material Anexo 30, 27 e 33.

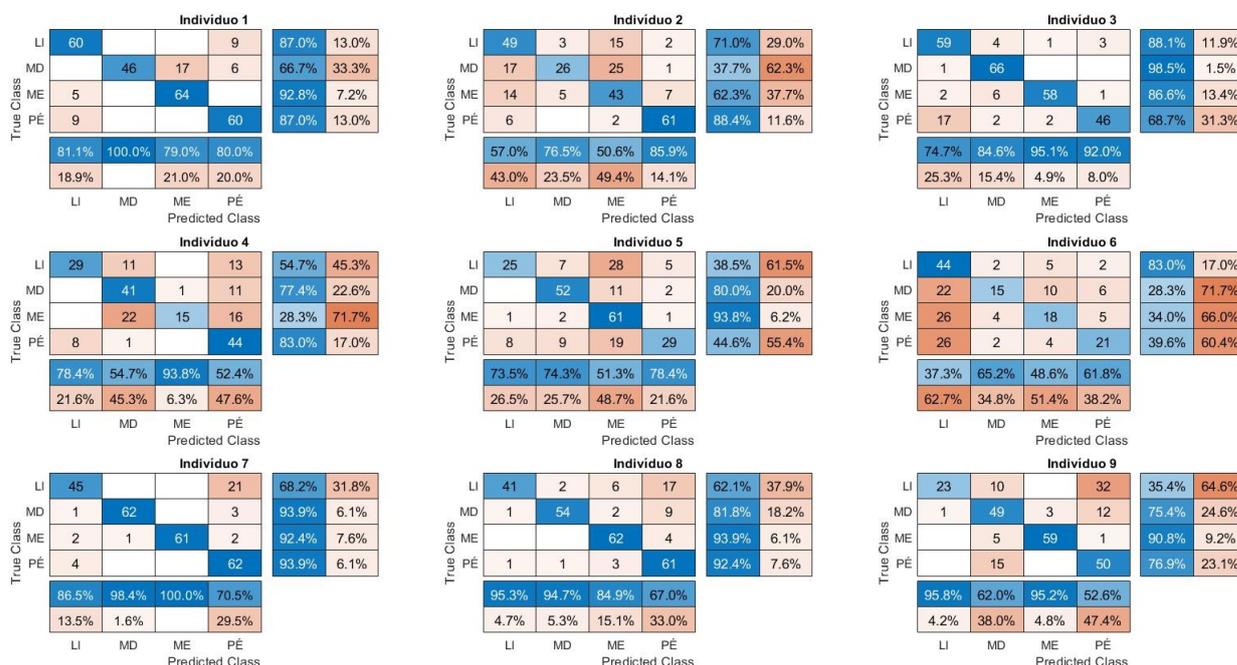


Figura 23 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP DS

Na Figura 24 é possível observar que, apesar de apresentar desempenho consistente, superiores à 0.7273 para os melhores indivíduos (1, 3, 7 e 8), o método FBCSP DS tem uma performance média inferior aos outros métodos propostos com kappa médio de 0.5761 e com valores de $kappa = 0.2767$ para os piores indivíduos (4 e 6). Em especial, quando comparado com o método FBCSP PW, que apresenta valores de kappa de 0.7711 à 0.8990 para os indivíduos 1, 3, 7 e 8.

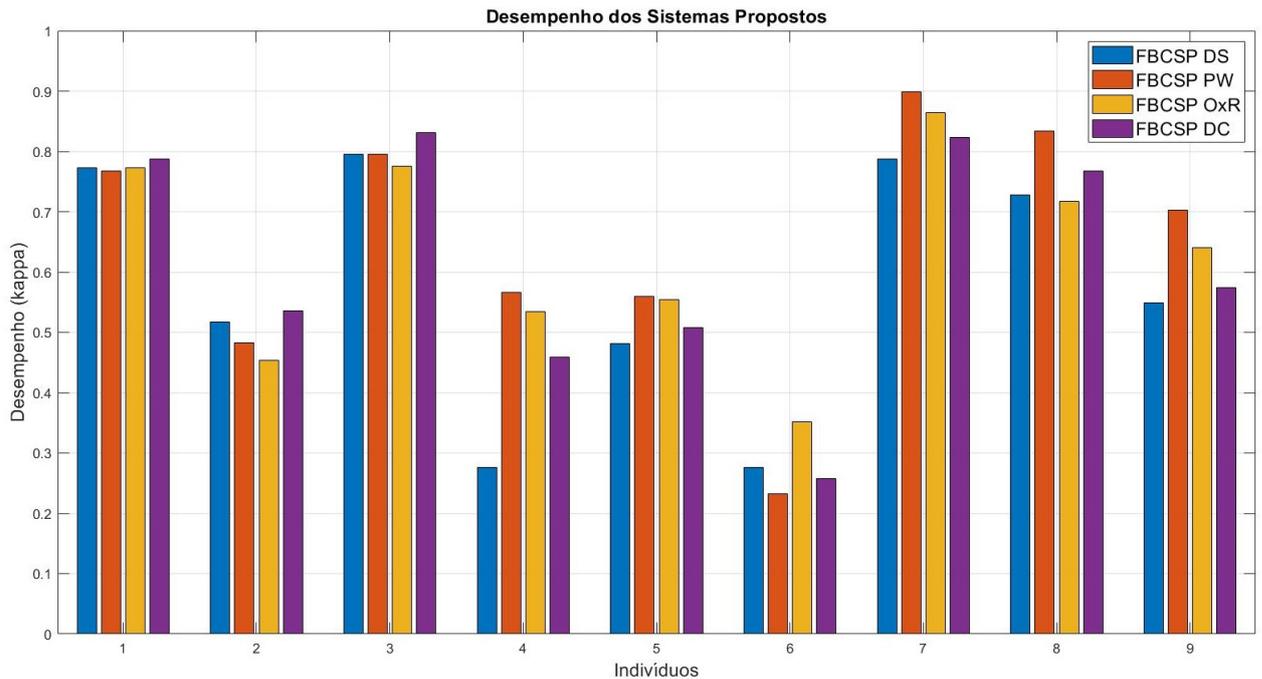


Figura 24 – Desempenho dos sistemas (coeficiente $kappa$) dos quatro métodos de extração de características, para os 9 indivíduos.

Foram obtidas as taxas de acerto de 67.31% para FBCSP DC, 71.64% para FBCSP DC, 72.49% para FBCSP PW e 71.88% para FBCSP OxR. Para todas as abordagens, o desempenho da discriminação entre PÉ e LI são, de forma geral, os mais baixos, 10.12% para FBCSP DS, 6.41% para FBCSP PW, 7.8% para FBCSP DC e 10.69% para FBCSP OxR, considerando-se os falsos positivos e falsos negativos para ambas as classes. Na Figura 25 observam-se as matrizes de confusão para os quatro métodos, considerando os resultados da classificação das 2292 observações dos 9 indivíduos.

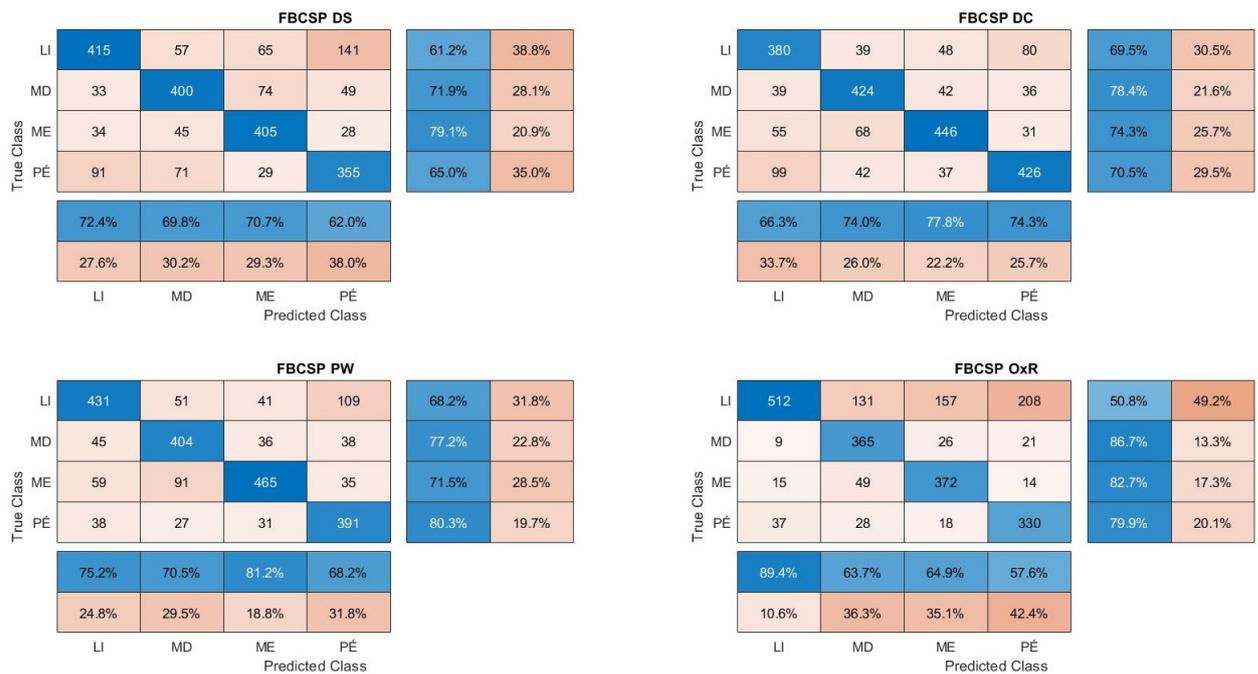


Figura 25 – Matrizes de Confusão das quatro abordagens FBCSP DS, DC, PW e OxR derivadas da classificação das 2292 observações de todos os 9 indivíduos.

4.3.2 Comparação com modelos de votação

Em comparação com os métodos propostos pela literatura, (ANG et al., 2012), é possível notar uma evidente melhora no desempenho dos sistemas quanto aplicadas as etapas propostas de pré-processamento, seleção de características e concatenação das características extraídas, bem como a relevância da extração de características por FBCSP DS. De forma geral, os modelos DC, PW e OxR tem melhoras de desempenho em comparação aos modelos propostos pela literatura, com aumentos no coeficiente *kappa* de 0.096, 0.077 e 0.061, respectivamente.

Tabela 9 – Comparação do desempenho dos modelos propostos e modelos da literatura. A nomenclatura '-C' identifica os métodos onde é proposta a concatenação das características. Em negrito, estão destacados os maiores valores de desempenho em relação à metodologia proposta e sublinhado estão os maiores valores de desempenho para cada indivíduo.

Ind.	Desempenho do Sistema (Kappa)						
	FBCSO DS	FBCSP DC-C	FBCSP DC	FBCSP PW-C	FBCSOP PW	FBCSO OxR-C	FBCSP OxR
1	0.7729	0.7874	0.708	0.7681	0.782	0.7729	0.676
2	0.5169	0.5362	0.370	0.4831	0.407	0.4541	0.417
3	0.7960	0.8308	0.657	0.7960	0.755	0.7761	0.745
4	0.2767	0.4591	0.472	0.5660	0.528	0.5346	0.481
5	0.4821	0.5077	0.407	0.5590	0.417	0.5535	0.398
6	0.2767	0.2579	0.264	0.2327	0.185	0.3522	0.273
7	0.7879	0.8232	0.727	0.8990	0.796	0.8586	0.773
8	0.7273	0.7677	0.579	0.8333	0.741	0.7172	0.755
9	0.5487	0.5744	0.495	0.7026	0.537	0.6410	0.606
Média	0.5761	0.6161	0.520	0.6489	0.572	0.6295	0.569

Considerando-se apenas os quatro indivíduos com melhor desempenho médio em todos os métodos (indivíduos 1, 3, 7 e 8) as abordagens DC, PW e OxR tem melhoras de desempenho em relação aos modelos propostos pela literatura com aumento de $kappa$ de 0.1345, 0.056 e 0.0313, respectivamente. O método FBCSP DS nas mesmas condições, tem melhoria de desempenho em relação ao coeficiente $kappa$ de 0.1032, 0.0025 e 0.0338 em relação aos métodos FBCSP DC, PW e OxR propostos pela literatura, respectivamente.

4.3.3 Padrões Tempo x Frequência das Características Seleccionadas

É possível ver na Figura 26 a relação entre tempo e frequência das características seleccionadas para cada modelo. O eixo do tempo indica qual o instante de início do trecho do sinal do qual se extraiu a característica, considerando-se a duração de 2 s, desta forma, uma característica que se inicia em -0.5 s foi extraída do trecho de -0.5 à 1.5 s. O eixo de frequências indica a banda da qual se extraiu a característica. De forma geral, as características seleccionadas como mais significativas em todos os quatro métodos são extraídas dos primeiros segundos do evento, entre 0 e 2.5 s, e das bandas mu e $beta$ central (8-28 Hz), um resultado já esperado para eventos motores.

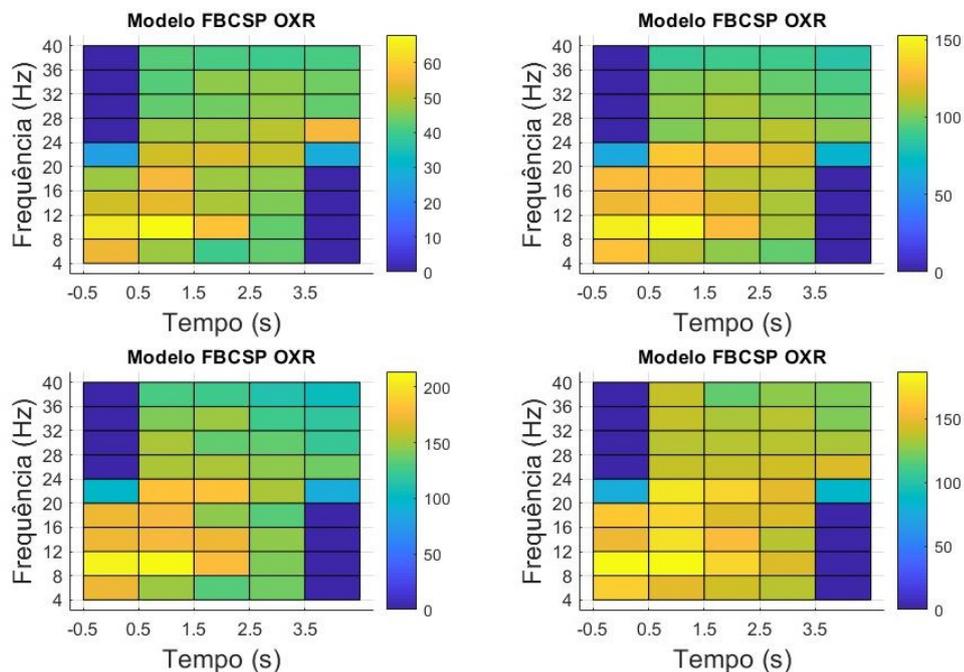


Figura 26 – Distribuição das características seleccionadas pelo método $mRMR$ no tempo [-0.5 à 4 s] e na frequência [4 à 40 Hz].

5 Conclusão

Neste trabalho foram avaliadas quatro abordagens na implementação do método FBCSP multiclasse bem como os impactos da etapa de seleção de características no desempenho de um sistema ICM de quatro classes. Dentre as quatro abordagens FBCSP multiclasse, o maior desempenho médio foi obtido com o método FBCSP PW + Seleção de Características, com *kappa* médio de 0.6489 para os 9 indivíduos e *kappa* médio de 0.8241 para os 4 melhores indivíduos. O segundo melhor desempenho foi obtido pelo método FBCSP OxR + Seleção de Características, com *kappa* médio de 0.6295 para os 9 indivíduos e 0.7825 para os quatro melhores. Para o método FBCSP DC + Seleção de Características, com terceiro melhor desempenho médio dentre as implementações, o valor médio de *kappa* para os 9 indivíduos foi de 0.6161 e 0.7772 para os quatro melhores, já o método FBCSP DS + Seleção de Características, do qual foi obtido o menor desempenho médio, o valor médio de *kappa* para os 9 indivíduos foi de 0.5761 e para os quatro melhores foi de 0.7711.

Desta forma, as quatro abordagens propostas permitem a extração de conjuntos características capazes de discriminar as quatro classes de IM, apresentando taxas de desempenho significativas para os 9 Indivíduos em comparação com as abordagens propostas na literatura. Ainda, aplicação da etapa de seleção de características permite a estimação de modelos de classificação mais eficientes, utilizando um conjunto reduzido de características além de aumentar discretamente o desempenho da discriminação. Dessas etapas, aliadas aos métodos de adequação dos sinais, obtém-se resultados que apontam para a conclusão de que a metodologia proposta é capaz de influenciar, de maneira positiva, na acurácia de sistemas ICM.

Referências

- ALLISON, B. Z.; WOLPAW, E. W.; WOLPAW, J. R. Brain–computer interface systems: progress and prospects. *Expert review of medical devices*, Taylor & Francis, v. 4, n. 4, p. 463–474, 2007. Citado na página 16.
- ANG, K. K. et al. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, Frontiers, v. 6, p. 39, 2012. Citado 4 vezes nas páginas 16, 34, 35 e 55.
- ANG, K. K. et al. Filter bank common spatial pattern (fbcs) in brain-computer interface. In: IEEE. *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. [S.l.], 2008. p. 2390–2397. Citado na página 16.
- BAMDAD, M.; ZARSHENAS, H.; AUAIS, M. A. Application of bci systems in neurorehabilitation: a scoping review. *Disability and Rehabilitation: Assistive Technology*, Taylor & Francis, v. 10, n. 5, p. 355–364, 2015. Citado 3 vezes nas páginas 11, 21 e 22.
- BEAR, M. F.; CONNORS, B. W.; PARADISO, M. A. *Neurociências: desvendando o sistema nervoso*. [S.l.]: Artmed Editora, 2008. Citado 2 vezes nas páginas 18 e 19.
- BERAHA, M. et al. Feature selection via mutual information: New theoretical insights. In: IEEE. *2019 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2019. p. 1–9. Citado na página 38.
- BLANKERTZ, B. et al. Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal processing magazine*, IEEE, v. 25, n. 1, p. 41–56, 2007. Citado na página 16.
- 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, p. 1–6, 2008. Citado 2 vezes nas páginas 7 e 25.
- BUNSE-GERSTNER, A.; BYERS, R.; MEHRMANN, V. Numerical methods for simultaneous diagonalization. *SIAM journal on matrix analysis and applications*, SIAM, v. 14, n. 4, p. 927–949, 1993. Citado na página 38.
- BURGES, C. J. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, Springer, v. 2, n. 2, p. 121–167, 1998. Citado na página 43.
- C-H, I. I. *Computational EEG Analysis: Methods and Applications*. Im C.-H., editor. [S.l.]: Springer, 2018. Citado na página 20.
- CARDOSO, J.-F.; SOULOUMIAC, A. Jacobi angles for simultaneous diagonalization. *SIAM journal on matrix analysis and applications*, SIAM, v. 17, n. 1, p. 161–164, 1996. Citado na página 37.
- CHIN, Z. Y. et al. Multi-class filter bank common spatial pattern for four-class motor imagery bci. In: IEEE. *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. [S.l.], 2009. p. 571–574. Citado na página 34.

- CHOW, T. W.; HUANG, D. Estimating optimal feature subsets using efficient estimation of high-dimensional mutual information. *IEEE Transactions on Neural networks*, IEEE, v. 16, n. 1, p. 213–224, 2005. Citado na página 39.
- DASH, M.; LIU, H. Feature selection for classification. *Intelligent data analysis*, IOS Press, v. 1, n. 3, p. 131–156, 1997. Citado na página 24.
- DECETY, J. Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research*, v. 3, n. 2, p. 87–93, 1996. ISSN 09266410. Citado na página 22.
- DEECKE, L.; SCHEID, P.; KORNHUBER, H. H. Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Experimental Brain Research*, Springer, v. 7, n. 2, p. 158–168, 1969. Citado na página 21.
- DORNHEGE, G. et al. *Toward brain-computer interfacing*. [S.l.]: Citeseer, 2007. v. 63. Citado na página 23.
- GROSSE-WENTRUP, M.; BUSS, M. Multiclass common spatial patterns and information theoretic feature extraction. *IEEE transactions on Biomedical Engineering*, IEEE, v. 55, n. 8, p. 1991–2000, 2008. Citado na página 34.
- GUILLOT, A.; COLLET, C. *The neurophysiological foundations of mental and motor imagery*. [S.l.]: Oxford University Press, 2010. Citado na página 22.
- HAYKIN, S. *Neural networks: a comprehensive foundation*. [S.l.]: Prentice Hall PTR, 1994. Citado 2 vezes nas páginas 42 e 43.
- HE, B. et al. Brain–computer interfaces. In: *Neural Engineering*. [S.l.]: Springer, 2013. p. 87–151. Citado 2 vezes nas páginas 22 e 24.
- JAMES, G. et al. *An introduction to statistical learning*. [S.l.]: Springer, 2013. v. 112. Citado 2 vezes nas páginas 42 e 43.
- JEANNEROD, M. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, v. 17, n. 2, p. 187–202, 1994. ISSN 14691825. Citado na página 22.
- KLEM, G. H. et al. The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, v. 52, n. 3, p. 3–6, 1999. Citado na página 20.
- KSCHISCHANG, F. R. The hilbert transform. *University of Toronto*, Citeseer, v. 83, p. 277, 2006. Citado na página 29.
- LIYANAGE, S. R. et al. Eeg signal separation for multi-class motor imagery using common spatial patterns based on joint approximate diagonalization. In: IEEE. *The 2010 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2010. p. 1–6. Citado 3 vezes nas páginas 16, 37 e 38.
- LUDWIG, K. A. et al. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *Journal of neurophysiology*, American Physiological Society, v. 101, n. 3, p. 1679–1689, 2009. Citado na página 28.

MEISHERI, H.; RAMRAO, N.; MITRA, S. K. Multiclass common spatial pattern with artifacts removal methodology for eeg signals. In: IEEE. *2016 4th International Symposium on Computational and Business Intelligence (ISCBI)*. [S.l.], 2016. p. 90–93. Citado na página 31.

MÜLLER-GERKING, J.; PFURTSCHELLER, G.; FLYVBJERG, H. Designing optimal spatial filters for single-trial eeg classification in a movement task. *Clinical neurophysiology*, Elsevier, v. 110, n. 5, p. 787–798, 1999. Citado na página 32.

PENFIELD, W.; BOLDREY, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, v. 60, n. 4, p. 389–443, 1937. ISSN 00068950. Citado na página 18.

PENG, H.; LONG, F.; DING, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 27, n. 8, p. 1226–1238, 2005. Citado 2 vezes nas páginas 16 e 40.

PFURTSCHELLER, G.; LOPES, F. H. Event-related EEG / MEG synchronization and desynchronization : basic principles. *Clinical Neurophysiology*, v. 110, p. 1842–1857, 1999. ISSN 1388-2457. Citado 2 vezes nas páginas 20 e 21.

PFURTSCHELLER, G. et al. Eeg-based discrimination between imagination of right and left hand movement. *Electroencephalography and clinical Neurophysiology*, Elsevier, v. 103, n. 6, p. 642–651, 1997. Citado na página 24.

RAMPIL, I. J. A primer for eeg signal processing in anesthesia. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, The American Society of Anesthesiologists, v. 89, n. 4, p. 980–1002, 1998. Citado na página 19.

SARMA, P. et al. Pre-processing and feature extraction techniques for eeg-bci applications-a review of recent research. *ADBU Journal of Engineering Technology*, v. 5, n. 1, 2016. Citado na página 23.

SYAM, S. H. F. et al. Comparing common average referencing to laplacian referencing in detecting imagination and intention of movement for brain computer interface. In: *MATEC Web of Conferences*. [S.l.: s.n.], 2017. v. 140. Citado na página 28.

TANG, J.; ALELYANI, S.; LIU, H. Feature selection for classification: A review. *Data classification: Algorithms and applications*, CRC Press, p. 37, 2014. Citado na página 24.

VALLABHANENI, A.; HE, B. Motor imagery task classification for brain computer interface applications using spatiotemporal principle component analysis. *Neurological research*, Taylor & Francis, v. 26, n. 3, p. 282–287, 2004. Citado na página 24.

VARELA, F. et al. The brainweb: phase synchronization and large-scale integration. *Nature reviews neuroscience*, Nature Publishing Group, v. 2, n. 4, p. 229–239, 2001. Citado na página 29.

WANG, Y.; GAO, S.; GAO, X. Common spatial pattern method for channel selection in motor imagery based brain-computer interface. In: IEEE. *2005 IEEE engineering in medicine and biology 27th annual conference*. [S.l.], 2006. p. 5392–5395. Citado na página 31.

WANG, Y. et al. Phase synchrony measurement in motor cortex for classifying single-trial eeg during motor imagery. In: IEEE. *2006 international conference of the IEEE engineering in medicine and biology society*. [S.l.], 2006. p. 75–78. Citado 2 vezes nas páginas 16 e 29.

WELCH, P. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, IEEE, v. 15, n. 2, p. 70–73, 1967. Citado 2 vezes nas páginas 16 e 30.

WOLPAW, J. *Brain-computer interfaces*. [S.l.: s.n.], 2013. v. 110. ISSN 00729752. Citado 2 vezes nas páginas 22 e 23.

WOLPAW, J. R. et al. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, Elsevier, v. 113, n. 6, p. 767–791, 2002. Citado na página 23.

WOLPAW, J. R. et al. An eeg-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, Elsevier, v. 78, n. 3, p. 252–259, 1991. Citado 2 vezes nas páginas 22 e 23.

ZHANG, R. et al. Eeg function network analysis of left and right hand motor imagery. In: IEEE. *Chinese Automation Congress (CAC), 2017*. [S.l.], 2017. p. 346–350. Citado na página 19.

6 Materiais elaborados pelo autor

Resultados das classificações binárias CSP, PLV e Welch

Tabela 10 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método CSP. As características foram extraídas do intervalo de $t=-0.5s$ à $t=3s$ após filtragem de 8-30Hz

Taxas de acerto para a classificação binária Algoritmo CSP (%)						
Indivíduo	Pé x Ling.	MD x Ling.	MD. x Pé	ME x Ling.	ME x Pé	ME x MD
1	82.61	98.55	98.55	92.03	97.10	89.13
2	82.61	76.09	92.75	73.91	76.09	60.14
3	68.66	94.78	90.30	96.27	97.76	91.79
4	65.09	85.85	95.28	81.13	77.36	69.81
5	58.46	70	62.31	63.85	73.08	54.62
6	68.87	70.75	64.15	66.98	62.26	74.53
7	81.06	96.21	97.73	93.94	97.73	81.06
8	81.06	87.88	87.12	93.18	88.64	97.73
9	83.08	83.85	87.69	95.38	94.62	90.77
Média	74.61	84.88	86.21	84.08	84.96	78.84

Tabela 11 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método PLV. As características foram extraídas pelo método de PLV Global dos sinais de EEG segmentados de $t=0s$ à $t=3s$ filtrados de 8-30Hz

Taxas de acerto para a classificação binária Algoritmo PLV (%)						
Indivíduo	Pé x Ling.	MD x Ling.	MD. x Pé	ME x Ling.	ME x Pé	ME x MD
1	64.49	93.48	89.86	92.75	92.03	65.22
2	65.94	55.07	75.36	51.45	69.57	50.72
3	56.72	88.81	79.85	80.60	79.85	84.33
4	48.11	66.04	63.21	59.43	64.15	72.64
5	47.69	50.	56.92	50.77	54.62	51.54
6	65.09	63.21	58.49	58.49	61.32	61.32
7	53.03	68.94	75	69.70	72.73	48.48
8	53.79	75	71.97	80.30	72.73	87.12
9	56.92	56.92	60.77	81.54	90.77	86.92
Média	56.87	68.61	70.16	69.45	73.08	67.59

Tabela 12 – Taxas de acerto para a classificação binária das combinações das quatro classes {IM ME, IM MD, IM PÉ, IM LI} pelo método do Periodograma de Welch. As características foram extraídas do sinal segmentado de t=0s à t=3s em trechos de 1s com sobreposição de 500ms

Taxas de acerto para a classificação binária Algoritmo Welch (%)						
Indivíduo	Pé x Ling.	MD x Ling.	MD. x Pé	ME x Ling.	ME x Pé	ME x MD
1	55.80	90.58	84.78	86.96	78.99	55.80
2	66.67	52.17	73.91	53.62	63.77	42.75
3	75.37	85.82	84.33	79.10	78.36	68.66
4	59.43	66.98	65.09	65.09	66.04	57.55
5	51.54	55.38	56.92	53.85	54.62	44.62
6	66.04	57.55	62.26	51.89	65.09	55.66
7	71.21	72.73	61.36	79.55	68.94	50
8	63.64	77.27	68.18	68.94	56.82	79.55
9	73.85	76.92	63.08	94.62	79.23	72.31
Média	64.84	70.60	68.88	70.40	67.98	58.54

Modelos e Resultados do método FBCSP DC

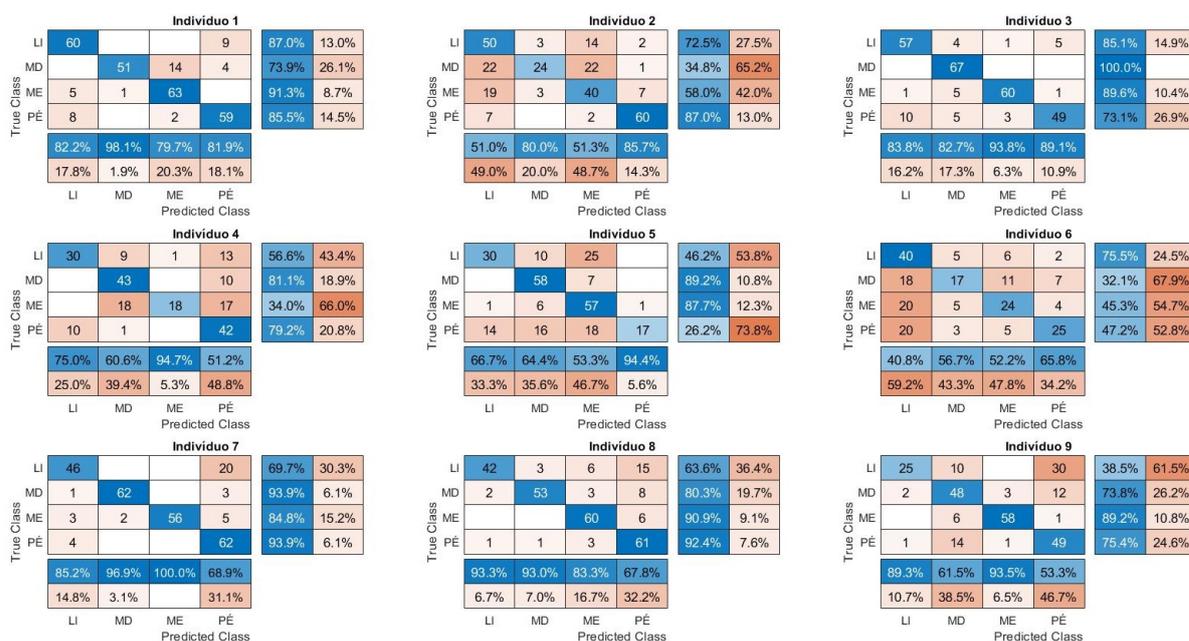


Figura 27 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP DC

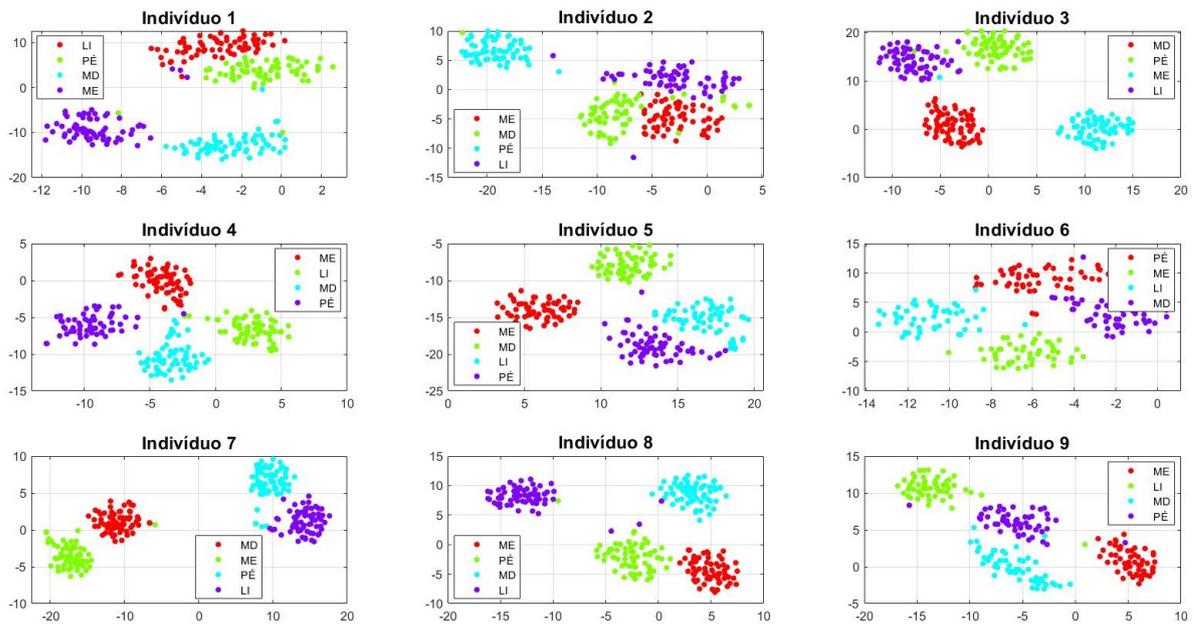


Figura 28 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem para o método FBCSP DC

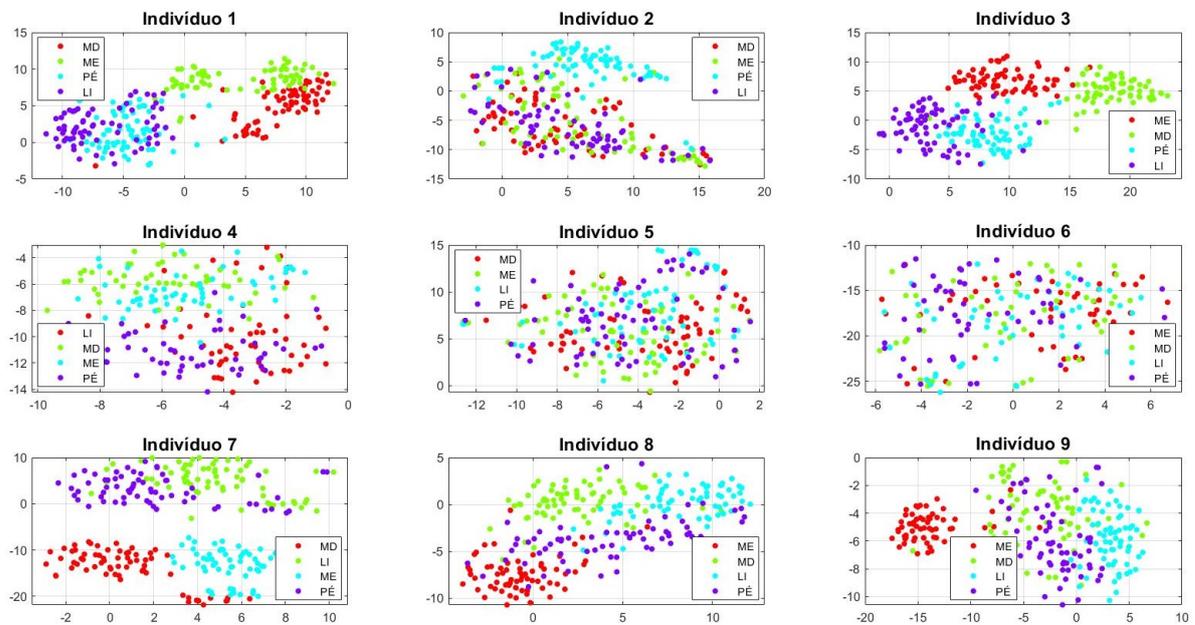


Figura 29 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP DC

Modelos e Resultados do método FBCSP PW

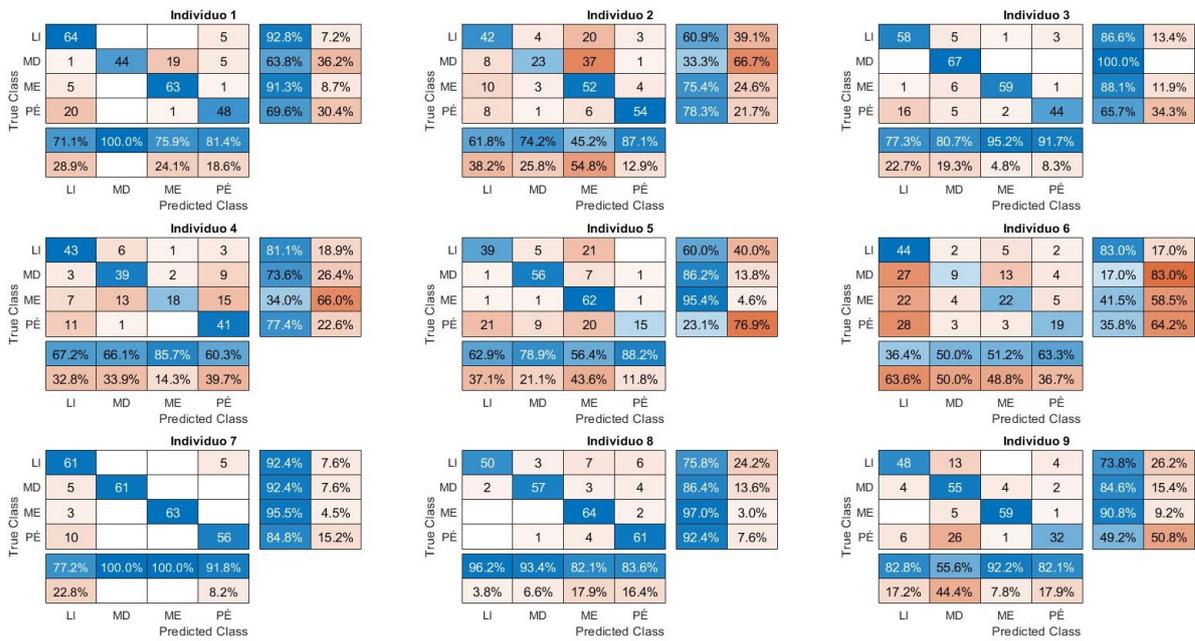


Figura 30 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP PW

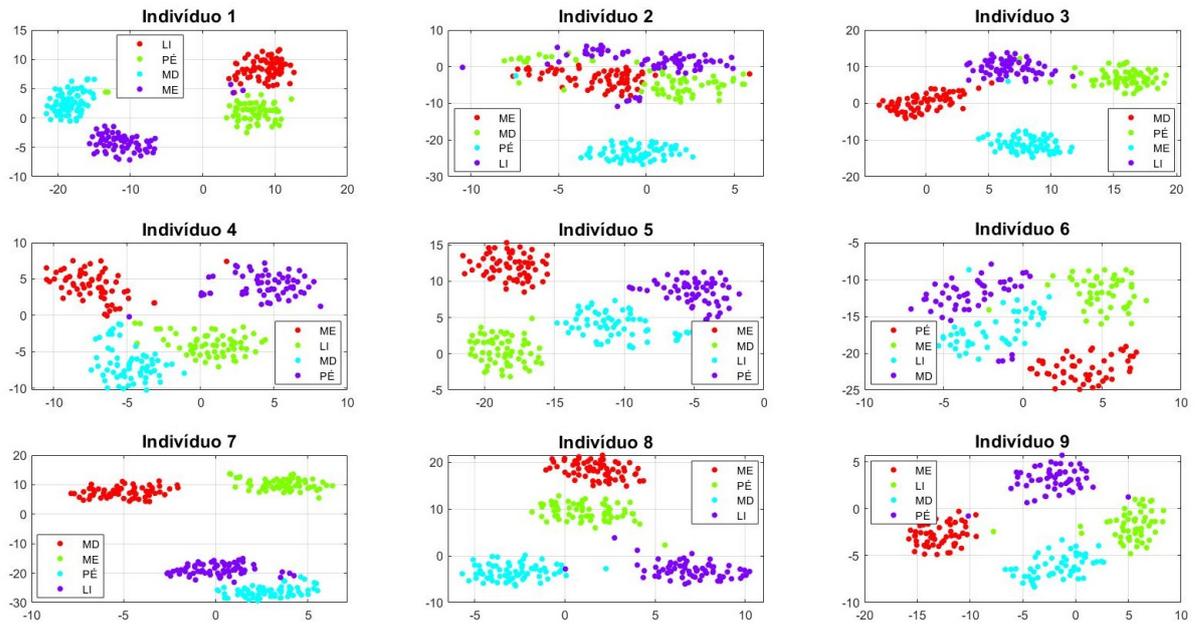


Figura 31 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem do método FBCSP PW

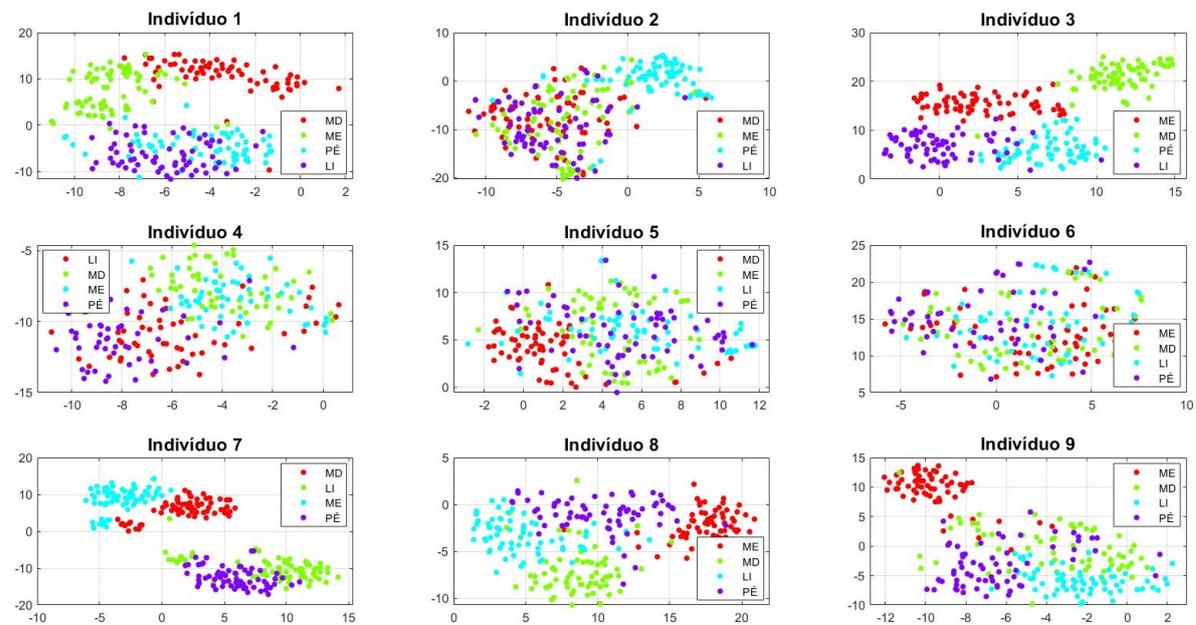


Figura 32 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP PW

Modelos e Resultados do método FBCSP O_xR

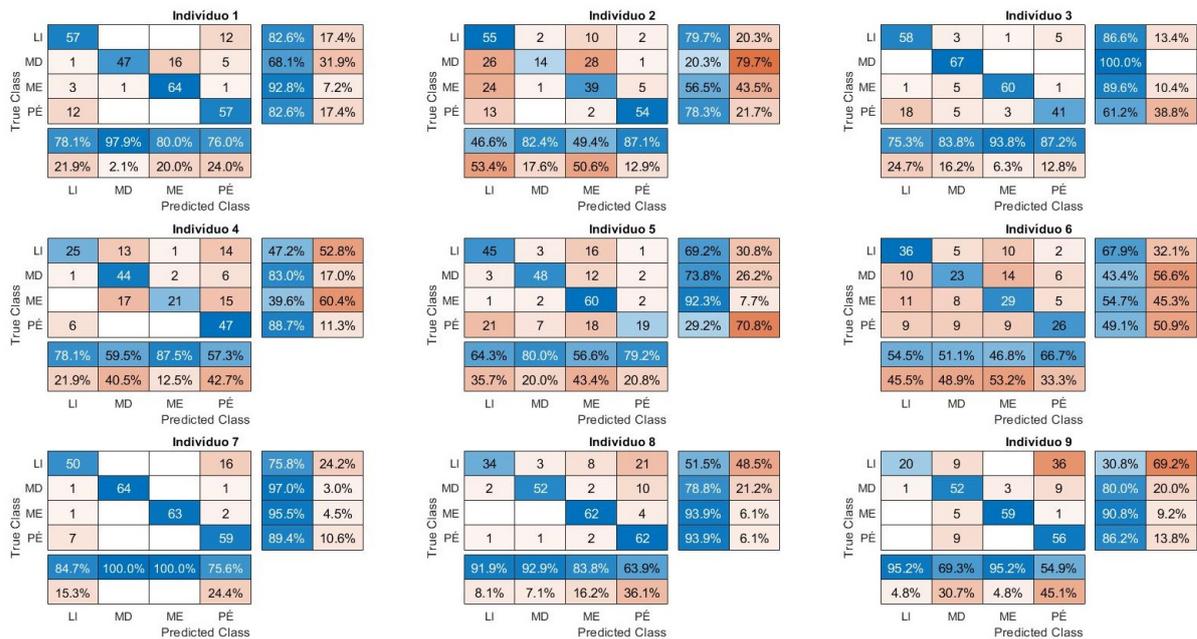


Figura 33 – Matrizes de Confusão do para a classificação das observações cegas de cada indivíduo com modelos estimados pelas características extraídas à partir do método FBCSP O_xR

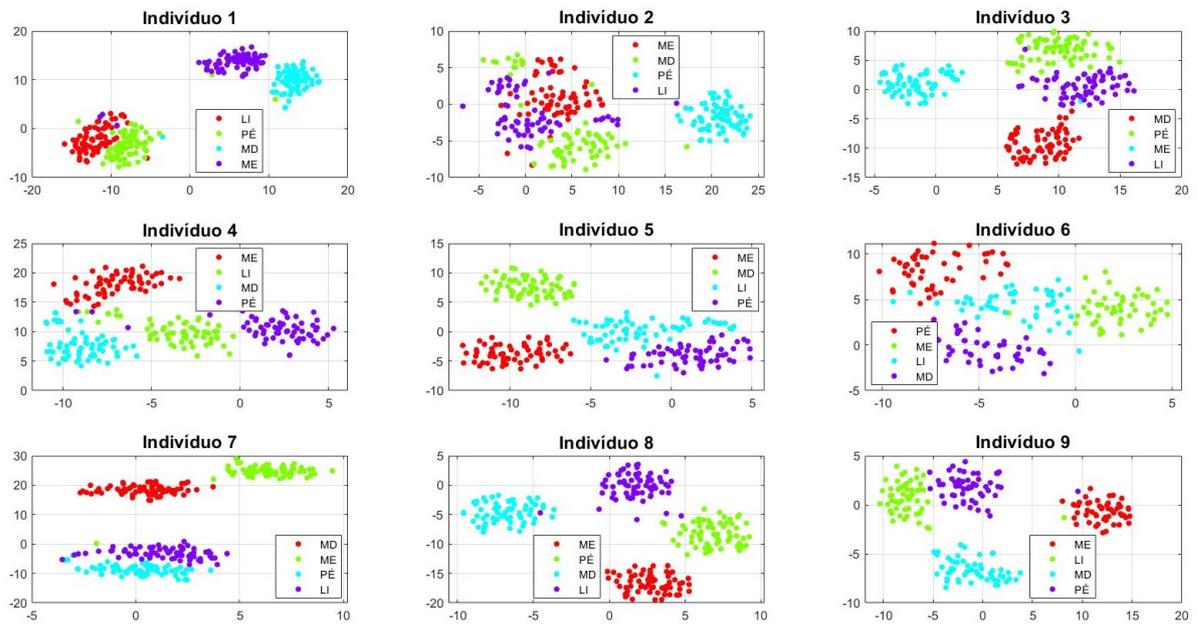


Figura 34 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de modelagem para o método FBCSP OxR

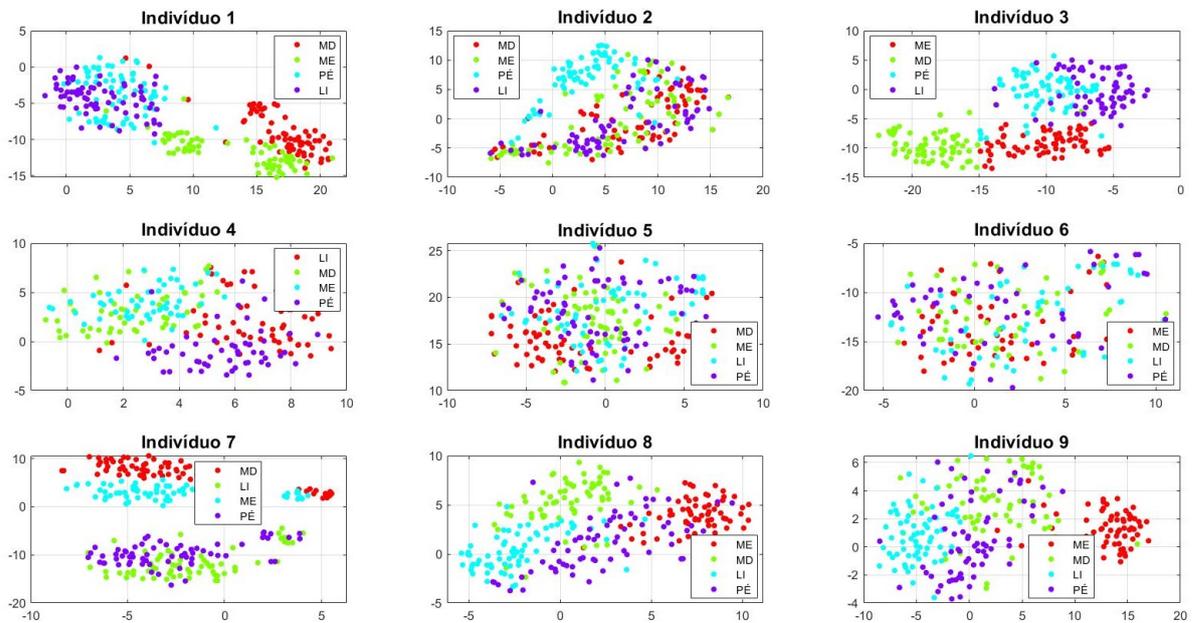


Figura 35 – Visualização em duas dimensões do conjunto multidimensional de características extraídas para os 9 indivíduos à partir das observações de teste pelo método FBCSP OxR

TERMO DE RESPONSABILIDADE

O texto do trabalho de conclusão intitulado "Análise do desempenho de técnicas multiclasse de extração de características em interfaces cérebro máquina ativadas por imaginação de movimento" é de minha inteira responsabilidade. Declaro que não há utilização indevida de texto, material fotográfico ou qualquer outro material pertencente a terceiros sem a devida citação ou consentimento dos referidos autores.

João Monlevade, 10 de Junho de 2022



Jefferson da Silva Costa