



UNIVERSIDAD DE GUANAJUATO

---

---

CAMPUS IRAPUATO - SALAMANCA  
DIVISIÓN DE INGENIERÍAS

*“Clasificación de Señales de Intención Motora en Electroencefalograma (EEG) utilizando Aprendizaje Automático”*

**TESIS**

Que para obtener el Grado de  
**Maestría en Ingeniería Eléctrica**

PRESENTA:

*Ing. Alberto Carlos Molina Porras*

DIRECTORES:

**Dr. Juan Gabriel Aviña Cervantes**

**Dr. Emmanuel Ovalle Magallanes**

Salamanca, Guanajuato,

Septiembre, 2023

Salamanca, Gto., a 19 de octubre del 2023.

**M. en I. HERIBERTO GUTIÉRREZ MARTIN**  
**COORDINADOR DE ASUNTOS ESCOLARES**  
**P R E S E N T E.-**

Por medio de la presente, se otorga autorización para proceder a los trámites de impresión, empastado de tesis y titulación al alumno Alberto Carlos Molina Porras del **Programa de Maestría en Ingeniería Eléctrica (Instrumentación y Sistemas Digitales)** y cuyo número de **NUA** es: 826842 del cual soy director. El título de la tesis es: *Clasificación de Señales de Intención Motora en Electroencefalograma (EEG) utilizando Aprendizaje Automático.*

Hago constar que he revisado dicho trabajo y he tenido comunicación con los sinodales asignados para la revisión de la tesis, por lo que no hay impedimento alguno para fijar la fecha de examen de titulación.

**A T E N T A M E N T E**



DR. JUAN GABRIEL AVIÑA CERVANTES  
**DIRECTOR DE TESIS**  
**SECRETARIO**



DR. EMMANUEL OVALLE MAGALLANES  
**DIRECTOR DE TESIS**



DRA. DORA LUZ ALMANZA OJEDA  
**PRESIDENTE**



DR. JOSÉ RUIZ PINALES  
**VOCAL**

---

## Dedicatoria

---

*A mi esposa Dayana Paola Duarte Zambrano.*

*A mis padres, Sra. Liliana Porras Fuenmayor y Sr. Diogenes Alberto Molina Altuve.*

*A mi hermano Juan Marco Molina Porras.*

*A toda mi familia y amigos cercanos.*

---

## Agradecimientos

---

*Al Dr. Juan Gabriel Aviña Cervantes y al Dr. Emmanuel Ovalle Magallanes por todo su apoyo.*

*Al Ph.D. Candidate Daniel Fernando Zambrano Gutierrez gran amigo de vida y a su esposa M.I. Grecia Carolina Duque Gimenez, por todo su apoyo en todo lo que he realizado.*

*A todas aquellas personas que de una u otra manera hicieron posible la realización de este trabajo de tesis.*

---

## Agradecimientos Institucionales

---

*Expreso mi más sincera gratitud hacia la Universidad de Guanajuato, especialmente a la División de Ingenierías del Campus Irapuato-Salamanca (DICIS) por la formación académica y por el apoyo financiero que he recibido durante mis estudios en esta institución.*



*Este trabajo de tesis fue realizado gracias al apoyo invaluable recibido a través del Consejo Nacional de Humanidades, Ciencia y Tecnología, CONAHCyT de México, bajo el número de becario 826842 y CVU 1156648.*



**CONAHCYT**

CONSEJO NACIONAL DE HUMANIDADES  
CIENCIAS Y TECNOLOGÍAS

---

## Índice General

---

Dedicatoria	ii
Agradecimientos	iii
Agradecimientos institucionales	iv
Índice de Figuras	x
Índice de Tablas	xiii
Resumen	xvii
Abstract	xx
<b>1 Introducción</b>	<b>3</b>
1.1 Introducción . . . . .	3

1.2	Justificación . . . . .	5
1.3	Planteamiento general del problema . . . . .	7
1.4	Objetivos . . . . .	7
1.4.1	Objetivo general . . . . .	7
1.4.2	Objetivos específicos . . . . .	8
1.5	Organización de la tesis . . . . .	8
<b>2</b>	<b>Fundamentos de BCI</b>	<b>9</b>
2.1	Conceptos en Interfaces Cerebro-Computadora . . . . .	10
2.2	Anatomía del Cerebro . . . . .	10
2.2.1	Neuronas . . . . .	10
2.2.2	Regiones del Cerebro . . . . .	12
2.2.2.1	Encéfalo . . . . .	12
2.2.2.2	Tronco Encefálico . . . . .	13
2.2.2.3	Cerebelo . . . . .	14
2.2.3	Lóbulos del Cerebro . . . . .	14
2.3	Señales de Electroencefalograma . . . . .	15
2.3.1	Adquisición de señales EEG . . . . .	16
2.3.2	Caracterización de las señales EEG . . . . .	17
2.3.3	Artefactos de ruido en señales EEG . . . . .	19
2.4	Imaginación Motora . . . . .	21
2.5	Conclusiones del capítulo . . . . .	23

<b>3</b>	<b>Procesamiento de Señales EEG</b>	<b>24</b>
3.1	Trabajos Relacionados . . . . .	25
3.2	Eliminación Artefactos de Ruido . . . . .	27
3.2.1	Filtro Digital Pasa-banda . . . . .	27
3.2.2	Diseño de banco de filtros IIR . . . . .	30
3.2.3	Descomposición de paquetes de onda . . . . .	32
3.3	Filtro Espacial . . . . .	37
3.3.1	Patrones Espaciales Comunes . . . . .	37
3.3.2	Diagonalización Aproximada Conjunta . . . . .	39
3.4	Técnicas Selección de Características . . . . .	40
3.4.1	Selección de filtros espaciales . . . . .	40
3.4.2	Selección por Valor Laplaciano . . . . .	42
3.5	Métodos de Clasificación . . . . .	44
3.5.1	Maquinas de Vectores de Soporte . . . . .	44
3.5.2	K-Vecinos Más Cercanos . . . . .	47
3.5.3	Análisis Discriminante Lineal . . . . .	49
3.6	Conclusiones del capítulo . . . . .	50
<b>4</b>	<b>Metodología</b>	<b>52</b>
4.1	Descripción de Bases de Datos . . . . .	54
4.1.1	BCI Competition IV Dataset 2a . . . . .	54
4.1.2	BCI Competition IV Dataset 2b . . . . .	55

4.2	Etapa de Pre-procesamiento . . . . .	57
4.2.1	Filtrado y Reducción de Artefactos de Ruido . . . . .	57
4.2.1.1	Banco de filtros con Solapamiento: . . . . .	57
4.2.1.2	Descomposición de Paquetes de Ondas: . . . . .	58
4.2.2	Filtrado Espacial . . . . .	60
4.2.2.1	CSP para clasificación binaria: . . . . .	60
4.2.2.2	CSP para clasificación multi-clase: . . . . .	62
4.3	Etapa de extracción y selección de características . . . . .	63
4.3.1	Extracción de Características . . . . .	63
4.3.2	Selección de Mejores Características . . . . .	64
4.4	Clasificación . . . . .	67
4.4.1	Modelo basado en SVM . . . . .	67
4.4.2	Modelo basado en LDA . . . . .	68
4.4.3	Modelo basado en KNN . . . . .	68
4.5	Métricas de desempeño . . . . .	69
4.6	Descripción de Marcos de Trabajo . . . . .	71
4.6.1	Propuesta 1: OFB-CSP-FS . . . . .	71
4.6.2	Propuesta 2: WPD-CSP-FS . . . . .	71
<b>5</b>	<b>Resultados experimentales</b>	<b>73</b>
5.1	Resultados con BCI Competition IV 2a . . . . .	74
5.1.1	Clasificación Binaria . . . . .	74

5.1.1.1	Resultados con Clasificador SVM . . . . .	75
5.1.1.2	Resultados con Clasificador KNN . . . . .	77
5.1.1.3	Resultados con Clasificador LDA . . . . .	79
5.1.1.4	Análisis de Resultados . . . . .	80
5.1.2	Clasificación Multiclases . . . . .	83
5.1.2.1	Resultados con modelo SVM . . . . .	84
5.1.2.2	Resultados con modelo KNN . . . . .	85
5.1.2.3	Resultados con modelo LDA . . . . .	86
5.1.2.4	Análisis de Resultados . . . . .	88
5.2	Resultados con base de datos BCI Competition IV 2b . . . . .	92
5.2.1	Resultados con Clasificador SVM . . . . .	94
5.2.2	Resultados con Clasificador KNN . . . . .	95
5.2.3	Resultados con Clasificador LDA . . . . .	97
5.2.4	Análisis de Resultados . . . . .	99
5.3	Comparativa con la literatura . . . . .	101
<b>6</b>	<b>Conclusiones Generales</b>	<b>104</b>
6.1	Trabajo futuro . . . . .	106
	<b>Referencias</b>	<b>107</b>

---

## Índice de Figuras

---

1.1	Bloques fundamentales en aplicaciones de Interfaz Cerebro - Computadora. Los bloques son: adquisición de las señales, sistema de reconocimiento y aplicación de la intención motora. . . . .	4
1.2	Contribuciones en la literatura sobre Interfaces Cerebro-Computadora en diferentes campos de investigación desde el año 2005 hasta 2023 según la base de datos Scopus. . . . .	6
2.1	Descripción general de la anatomía de una neurona, compuesta por el soma, las dendritas, el cono, el axón y los terminales. . . . .	11
2.2	Regiones del Cerebro. . . . .	12
2.3	Vista lateral de la ubicación de los lóbulos del Cerebro. . . . .	14
2.4	Ubicación de electrodos no invasivos en el cuero cabelludo correspondiente al sistema de localización 10/20 [Klem et al., 1999]. . . . .	17

2.5	Identificación de artefactos oculares obtenidos con lecturas EOG (selección marcada por las bandas amarillas) con escala en $\mu V$ , y su efecto en una señal EEG obtenida de la base de datos BCI Competition IV dataset 2a [Tangermann et al., 2012]. . . . .	20
2.6	Distribución de regiones en la corteza motora primaria y la corteza somato-sensorial primaria del cerebro, y su influencia en los diferentes tipos de intención motora [Marieb and Hoehn, 2010]. . . . .	22
2.7	Efecto de la ERD sobre la densidad espectral de potencia de la señal de EEG. (a) Atenuación de potencia en las bandas $\alpha$ y $\beta$ relacionadas con la señal de pensamiento motor de brazo derecho (R). (b) Atenuación de potencia en las bandas $\alpha$ y $\beta$ relacionadas con la señal de pensamiento motor de brazo izquierdo (L). . . . .	23
3.1	Contribuciones en la literatura relacionadas con técnicas de aprendizaje automático aplicado a la detección de imaginación motora en BCI, consulta realizada en Scopues el 29 de marzo del 2023. . . . .	25
3.2	Parámetros de rendimiento en respuesta de magnitud de un filtro digital. . . .	29
3.3	Diagrama de descomposición en dos niveles con la técnica de descomposición de paquetes de onda (WPD). . . . .	35
3.4	Interpretación geométrica del SVM con margen regularizado ( <i>soft margin</i> ). . .	45
3.5	Interpretación del modelo KNN identificando la clase de una instancia no clasificada $y$ con $k = 3$ . . . . .	47
3.6	Interpretación del modelo LDA para la identificación de la clase de una instancia no clasificada. . . . .	50
4.1	Esquema general de las etapas desarrolladas en la metodología de la investigación.	53
4.2	Detalles de la adquisición de los datos suministrados en la <i>BCI Competition IV dataset 2a</i> por Tangermann et al. [2012]. . . . .	55
4.3	Estructura de cada sesión en el <i>BCI Competition IV dataset 2b</i> . . . . .	56

4.4	Esquema de experimento realizado para la toma de datos en la <i>BCI Competition IV dataset 2b</i> por <i>Tangermann et al. [2012]</i> . . . . .	56
4.5	Aplicación del banco de filtros IIR Butterworth de Orden 2 sobre las señales adquiridas en la base de datos. . . . .	58
4.6	Descomposición de 5 niveles con WPD sobre la señal EEG. . . . .	59
4.7	Gráfica de la aproximación FIR Discreta del Wavelet de Meyer. . . . .	59
4.8	Esquema de las etapas de procesamiento de la señal implementando selección de filtros con ITFE. . . . .	65
4.9	Esquema de las etapas de procesamiento de la señal implementando selección de características con Valor Laplaciano. . . . .	66
5.1	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS. . . . .	82
5.2	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS. . . . .	83
5.3	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS en combinación multi-clases. . . . .	91
5.4	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS en combinación multi-clases. . . . .	92
5.5	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS por cada sujeto de prueba de la base de datos BCI Competition IV 2b. . . . .	99
5.6	Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS por cada sujeto de prueba de la base de datos BCI Competition IV 2b. . . . .	100

---

## Índice de Tablas

---

2.1	Clases de señales de electroencefalograma en función del rango de frecuencia [Sörnmo and Laguna, 2005]. . . . .	18
3.1	Kernel de ponderación para cálculo de pesos en KNN. . . . .	49
5.1	Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	75
5.2	Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	76
5.3	Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	77

5.4 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	78
5.5 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	80
5.6 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases. . . . .	81
5.7 Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo OFB-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación binaria de clases. . . . .	82
5.8 Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo WPD-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación binaria de clases. . . . .	84
5.9 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. . . . .	85
5.10 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. . . . .	86
5.11 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. . . . .	87

5.12 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. 88

5.13 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. 89

5.14 Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases. 90

5.15 Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo OFB-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación multi-clases. . . . . 91

5.16 Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo WPD-CSP-FS con información mutua (ITFE) como selector de característica, el logaritmo de la varianza promedio como característica (logvar) y clasificador SVM con combinación multi-clases. . . . . 93

5.17 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas. . . . . 94

5.18 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas. . . . . 95

5.19 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas. . . . . 96

5.20 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas. . . . . 97

5.21 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas.. . . . 98

5.22 Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas.. . . . 98

5.23 Comparación entre métricas de Exactitud [%] y Kappa promedio publicadas en la literatura, y las obtenidas por los marcos de trabajo propuestos durante clasificación de cuatro clases con los datos de prueba de la BCI Competition VI 2a. . . . . 101

5.24 Comparación entre máximo kappa obtenido por los marcos de trabajo propuestos y los algoritmos en la literatura durante la clasificación de las clases L y R con datos de prueba de la BCI Competition VI 2a y 2b. . . . . 102

---

## List of Algoritmos

---

1	Algoritmo CSP aplicado a dos clases . . . . .	61
2	Algoritmo CSP aplicado a Multi-clases . . . . .	62
3	Algoritmo de Selección de Características . . . . .	66

---

## Resumen

---

Las Interfaces Cerebro-Computadora son métodos que permiten la comunicación directa entre el cerebro humano y una computadora u otro dispositivo digital. Esta interfaz utiliza un conjunto de herramientas tecnológicas para leer las señales eléctricas o químicas del cerebro, y traducirlas en comandos que pueden ser interpretados por la computadora. Implementar algoritmos de inteligencia artificial en la creación de interfaces, que permitan controlar sistemas digitales con la actividad cerebral, representa un campo de investigación en constante crecimiento. Para ello, una alternativa popular consiste en utilizar señales de electroencefalograma adquiridas a través de dispositivos externos no invasivos con la finalidad de detectar patrones en pensamientos de intención motora.

Una de las técnicas de procesamiento más exitosas aplicadas en la detección de patrones en señales de electroencefalograma se conoce como Patrones Espaciales Comunes (*Common Spatial Pattern*, CSP), y busca obtener filtros espaciales que maximicen las diferencias entre las diferentes clases de pensamiento motor en función de su varianza. Desafortunadamente, esta técnica es sensible a la presencia de ruido en las señales y las variaciones en el comportamiento de la señal entre individuos diferentes. De igual manera, las variaciones presentes en las señales de un mismo individuo cuando son adquiridas en diferentes días, también representan un desafío que afecta a los sistemas de reconocimiento basados en CSP.

En este trabajo de investigación, se propusieron dos marcos de trabajo para la clasificación de señales de intención motora basados en CSP con el objetivo de aumentar el

rendimiento de métodos clásicos de aprendizaje automático. El primero, llamado OFB-CSP-FS, está definido por la incorporación de una etapa de pre-procesamiento mediante un banco de filtros IIR pasa-banda con solapamiento en frecuencia sobre la señal original. Por otro lado, el segundo marco de trabajo, llamado WPD-CSP-FS, emplea como pre-procesamiento de las señales la Descomposición de Paquetes de Ondas (*Wavelet Packet Decomposition*, WPD). En cada propuesta se realizan pruebas exhaustivas combinando características basadas en el logaritmo de la varianza de la señal procesada y el logaritmo de su potencia promedio, junto con algoritmos de selección de características basados en la información mutua (Método supervisado) o basados el cálculo del valor laplaciano (Método no supervisado).

Finalmente, para cada una de las combinaciones planteadas, se evaluaron tres clasificadores clásicos: Máquina de Vectores de Soporte, K Vecinos más Cercanos y Análisis Discriminante Lineal, sometiéndolos a prueba con las bases de datos de la BCI Competition IV dataset 2a y dataset 2b. Los resultados demuestran que OFB-CSP-FS es la mejor propuesta y logra mejorar el rendimiento de los sistemas de reconocimiento reportados en la literatura que implementan versiones de CSP con técnicas clásicas de aprendizaje automático para el mismo caso de estudio.

---

## Abstract

---

Brain Computer Interfaces (BCI) are systems that allow direct communication between the human brain and a computer or other digital device. This interface uses a set of technological tools to read electrical or chemical signals from the brain, and translate them into commands that can be interpreted by computers. Implementing artificial intelligence algorithms in the creation of interfaces to control digital systems with brain activity is a growing research field. A popular alternative is to use electroencephalogram (EEG) signals acquired through non-invasive external devices to detect patterns during Motor Imagery performings.

One of the most successful processing techniques applied for pattern recognition in EEG signals is known as Common Spatial Pattern (CSP), which try to obtain spatial filters that maximize the differences between different classes of motor imagery as a function of their variance. Unfortunately, the performance of pattern recognition systems with CSP can be significantly affected due to the presence of noise artifacts in the signals, variations in signal behavior between different subjects, even between different days with the same subject, and the presence of inconsequential information that attenuates the performance of the classifier.

In this work, two frameworks were proposed for the classification of motor imagery signals based on CSP with the aim of increasing its performance. The first one, called OFB-CSP-FS, is defined by a band-pass IIR filter bank with overlapping in frequency applied in the original signal, as a previous stage to the application of CSP. On the other hand, the second framework, called WPD-CSP-FS, is defined by Wavelet Packet Decomposition (WPD)

as a prior stage to the CSP implementation. In each proposal, tests are performed combining features based on the logarithm of the variance of the processed signal and the logarithm of its average power, together with supervised and unsupervised feature selection algorithms.

Finally, different combinations of processing techniques were evaluated in conjunction with three classical classifiers, SVM, KNN and LDA, by testing them on the BCI Competition IV dataset 2a and dataset 2b. The results show that OFB-CSP-FS managed to improve the performance of recognition systems implementing CSP with machine learning algorithms, reported in the literature for the same case study.

---

## Lista de acrónimos

---

ASR	Artifact Subspace Reconstruction
BCI	Brain Computer Interface
CSP	Common Spatial Pattern
ECG	Electrocardiogram
EEG	Electroencephalogram
EMD	Empirical Mode Descomposition
EMG	Electromiogram
EMI	Electromagnetic Interference
EOG	Electrooculogram
ERD	Event-Related Desynchronization
ERS	Event-Related Synchronization
FIR	Finite Impulse Response
ICA	Independent Component Analysis
IIR	Infinite Impulse Response

INEGI	Instituto Nacional de Estadística y Geología
ITFE	Information Theoretic Feature Extraction
JAD	Joint Aproximate Diagonalization
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LS	Laplacian Score
LTI	Linear time-invariant
MI	Motor Imagery
MI-BCI	Motor Imagery - Brain Computer Interface
ML	Machine Learning
OFB-CSP-FS	Overlapped Filter Bank Common Spatial Pattern with Feature Selection
OVR	One vs Rest
RBF	Radial Basis Function
SCM	Sample Covariance Matrix
STFT	Short-Time Fourier Transform, STFT
SVD	Singular Value Descomposition
SVM	Support Vector Machines
WPD	Wavelet Packet Descomposition
WPD-CSP-FS	Wavelet Packets Descomposition Common Spatial Pattern with Feature Selection

# CAPÍTULO 1

---

## Introducción

---

---

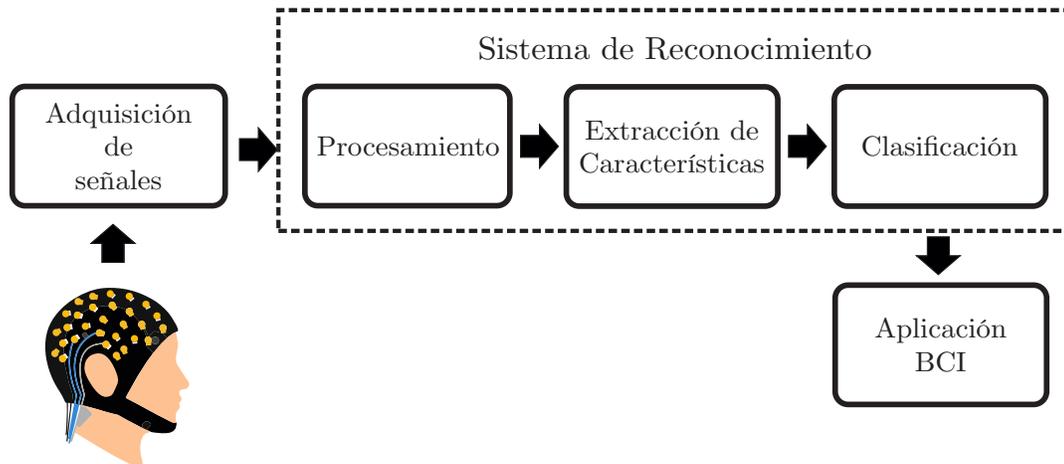
<b>1.1</b>	<b>Introducción</b>	<b>3</b>
<b>1.2</b>	<b>Justificación</b>	<b>5</b>
<b>1.3</b>	<b>Planteamiento general del problema</b>	<b>7</b>
<b>1.4</b>	<b>Objetivos</b>	<b>7</b>
1.4.1	Objetivo general	7
1.4.2	Objetivos específicos	8
<b>1.5</b>	<b>Organización de la tesis</b>	<b>8</b>

---

## 1.1. Introducción

Gracias a los avances de la neurociencia, a partir de 1970, empezaron las primeras investigaciones orientadas a utilizar las señales eléctricas generadas por la actividad neuronal como herramienta para controlar dispositivos electrónicos. A estos sistemas se les denomina Interfaces Cerebro-Computadora (*Brain Computer Interface*, BCI), y pueden ser descritos

como herramientas de hardware y software que permiten a los seres humanos interactuar con su entorno sin la intervención de nervios y músculos periféricos, mediante señales de control generadas por la actividad electroencefalográfica [Nicolas-Alonso and Gomez-Gil, 2012]. Las BCI miden la actividad neuronal en regiones específicas, tal que sea posible extraer características descriptivas que sustituyan, restauren, potencien, complementen o mejoren las funciones humanas [Tandon and Allison, 2018]. Como se observa en la Figura 1.1, las señales del cerebro, tales como: amplitudes de los potenciales, los ritmos de la corteza somatosensorial o la frecuencia de neuronas corticales, se adquieren mediante electrodos colocados en el cuero cabelludo, en la superficie cortical o dentro del cerebro. Estas señales se procesan para extraer atributos o características, las cuales son un conjunto de medidas o propiedades específicas que describen su comportamiento y representa la intención del usuario.



**Figura 1.1.** Bloques fundamentales en aplicaciones de Interfaz Cerebro - Computadora. Los bloques son: adquisición de las señales, sistema de reconocimiento y aplicación de la intención motora.

Extraer características discriminantes entre las diferentes clases dentro del conjunto de datos puede ayudar a capturar patrones y relaciones importantes entre ellos, lo que a su vez permite que los algoritmos de aprendizaje automático los traduzcan en órdenes que hacen funcionar dispositivos físicos [Schalk et al., 2004]. Esto resulta de gran utilidad aplicado a personas con discapacidades motoras para mejorar su calidad de vida y, al mismo tiempo, reducir el costo de los cuidados intensivos.

En este trabajo de investigación se implementarán métodos tecnológicos orientados al procesamiento de señales de electroencefalograma (*Electroencephalogram*, EEG) no invasivos, utilizando electrodos colocados en la superficie del cuero cabelludo, con el objetivo de clasificar

correctamente las señales generadas al momento de pensar en el movimiento de alguna parte del cuerpo, también llamadas imaginación motora. Según los estudios, la imaginación motora de los usuarios estimula el córtex somatosensorial del cerebro, aumentando el metabolismo y el flujo sanguíneo al tiempo que disminuye o bloquea la amplitud de potencia en las bandas de frecuencia  $\alpha$  (8-13 Hz) y  $\beta$  (14-30 Hz) de las señales de EEG [Abdullah et al., 2022].

Aunque el desarrollo de nuevas BCI es un área de interés ampliamente estudiada, todavía presenta algunos retos científicos que deben ser abordados. Entre ellos, haciendo referencia a la detección de intenciones motoras, el principal problema para lograr un sistema de clasificación robusto esta asociado con la calidad de la señal, debido a la alta presencia de artefactos de ruido aditivos originados principalmente por el movimiento, la actividad cardiaca o la actividad muscular. Este fenómeno dificulta una correcta aplicación en situaciones donde el usuario no se encuentra en un ambiente controlado y relajado.

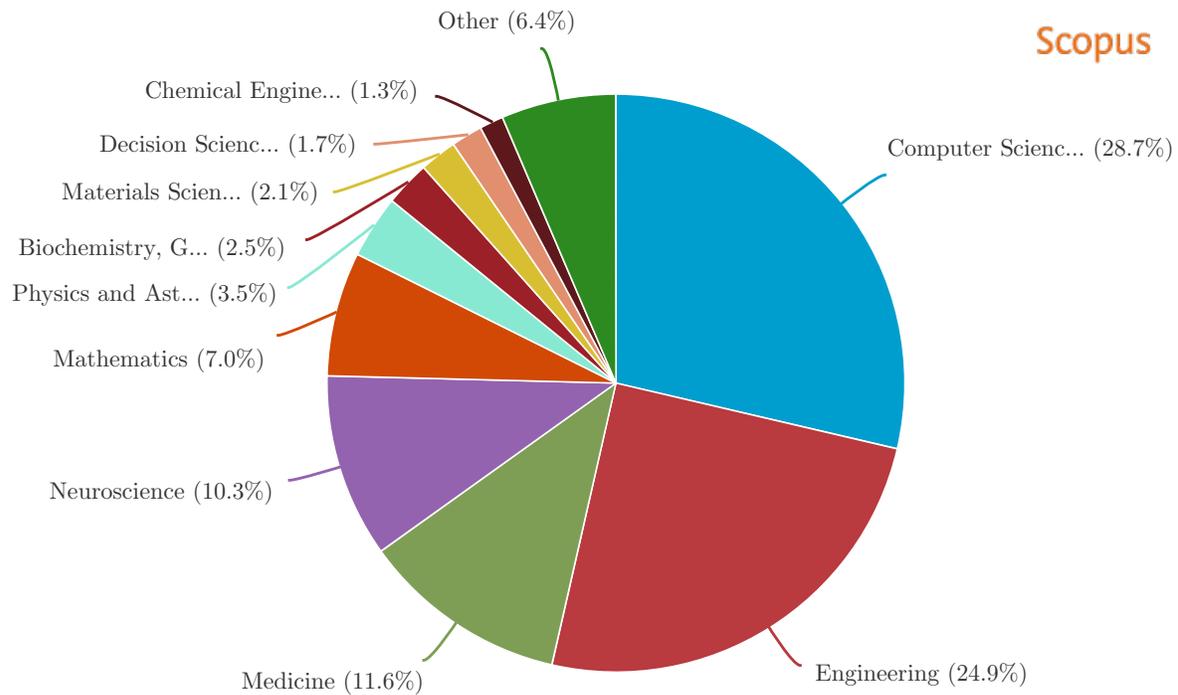
El objetivo principal de esta investigación es establecer diferentes marcos de trabajo fundamentados en técnicas de pre-procesamiento y extracción de características modernas, aplicándolas en bases de datos con señales de EEG orientadas a la detección de intenciones motoras robustas a la presencia de artefactos de ruido. Se someterá a estudio el rendimiento de estas técnicas, y se buscará el mejor marco de trabajo que permita tasas de clasificación adecuadas para dichas señales EEG.

## 1.2. Justificación

En México, según estudios realizados por el Instituto Nacional de Estadística y Geología (INEGI), a través del censo de demografía y sociedad realizado en el año 2020 [INEGI, 2020], identificó un total de 20,477,037 personas, el 16% de la población, como personas con discapacidad o que dijeron tener alguna limitación para realizar actividades de la vida diaria (*e.g.*, caminar, ver, oír, auto-cuidado, hablar o comunicarse). Adicionalmente, según la Organización Mundial de la salud en su informe mundial sobre la discapacidad [Organization and Bank, 2011], aproximadamente una de cada siete personas a nivel mundial cuenta con algún tipo de discapacidad.

Tomando en cuenta las necesidades de este importante sector de la población, y gracias a los avances en el campo de la neurociencia y la tecnología de la información, las herramientas de BCI evolucionan constantemente para ofrecer una solución alternativa orientada a la

interacción de los usuarios con los dispositivos tecnológicos a partir de la detección de intenciones motoras, es decir, pensar en mover alguna parte del cuerpo [Kawala-Sterniuk et al., 2021].



**Figura 1.2.** Contribuciones en la literatura sobre Interfaces Cerebro-Computadora en diferentes campos de investigación desde el año 2005 hasta 2023 según la base de datos Scopus.

Las BCI han contribuido en varios campos de la investigación. Están involucradas en el área de la medicina, neuroergonomía y entorno inteligente, neuromarketing y publicidad, educación, juegos y entretenimiento, así como en campos de seguridad y autenticación [Abdulkader et al., 2015], como se observa en la Figura 1.2. Estudios en el campo de la neurociencia, también reportan avances prometedores en aplicaciones de detección y diagnóstico de enfermedades cerebrales [Abdulkader et al., 2015], sistemas robóticos [Hayta et al., 2022], rehabilitación motora en pacientes con problemas de movilidad [Lim et al., 2022] y en sistemas de control cognitivo para aplicaciones en domótica [Akram et al., 2022].

Por esta razón, la implementación de algoritmos de inteligencia artificial en la creación de interfaces, que interpreten las señales cerebrales de los usuarios como acciones de control, representa un campo de investigación en constante crecimiento. La posibilidad de poder controlar sistemas digitales con la actividad cerebral utilizando señales de EEG adquiridas a través de dispositivos externos no invasivos, sugiere una solución interesante para personas con discapacidades motoras o de comunicación, a las cuales se les dificulta utilizar las interfaces

de control convencionales.

### 1.3. Planteamiento general del problema

A pesar de que las BCI tienen el potencial de revolucionar muchos campos, como la medicina, la educación y la tecnología, aún existen varios desafíos técnicos y científicos que deben abordarse para que puedan ser utilizadas de manera efectiva y segura en la vida cotidiana.

La principal barrera con la que se han topado los investigadores en el desarrollo de sistemas que interpretan las señales de EEG, es que todavía los desempeños alcanzados están lejos de ser ideales debido a la dificultad de lograr un procesamiento de bio-señales preciso y en tiempo real para el reconocimiento de patrones [Mwata-Velu et al., 2021]. La dificultad radica principalmente en la selección de los canales EEG correctos, la relación señal-ruido de estas señales y cómo discernir la información redundante entre ellos [Fumanal-Idocin et al., 2022].

Adicionalmente, otra dificultad que debe ser tomada en cuenta en este tipo de aplicaciones, corresponde a las variaciones existentes en las lecturas de EEG tomadas en días y sujetos diferentes. Al ser señales de baja amplitud, dependientes de factores biológicos y fácilmente alteradas por agentes externos, es necesario implementar técnicas de procesamiento adecuadas para determinar características diferenciadoras y robustas ante los posibles cambios que experimentan las señales. De esta manera, es posible aumentar el rendimiento de los algoritmos de clasificación al discernir entre las diferentes clases de pensamiento.

Por lo tanto, reducir la influencia del ruido, seleccionar las características adecuadas y optimizar los protocolos de entrenamiento en los sistemas de reconocimiento son factores determinantes para lograr resultados adecuados en las BCI basadas en imaginación motora.

### 1.4. Objetivos

#### 1.4.1. Objetivo general

Clasificar señales de intención motora contenidas en lecturas de señales EEG con técnicas robustas de aprendizaje automático.

### 1.4.2. Objetivos específicos

1. Estudiar el estado del arte sobre los métodos de extracción, selección y técnicas de clasificación de características orientadas a la detección de intenciones motoras en señales de EEG.
2. Estudiar e implementar una estrategia para la eliminación de señales de ruido presentes en las señales de EEG.
3. Evaluar los métodos de extracción y selección de características más relevantes, y técnicas de aprendizaje automático adecuadas para la clasificación de intenciones motoras en señales de EEG.
4. Comparar el rendimiento del enfoque propuesto con los resultados que se mencionan en la literatura relacionados con técnicas de aprendizaje automático.

## 1.5. Organización de la tesis

Este trabajo de tesis está organizado en 6 capítulos, los cuales se describen brevemente a continuación:

Capítulo 1: Presenta la introducción y los objetivos de la presente obra.

Capítulo 2: Describe los conceptos teóricos relacionados a la imaginación motora contenidas en señales de EEG, y sus aplicaciones en Interfaces Cerebro-Computadora.

Capítulo 3: Describe las bases científicas y matemáticas de los métodos empleados en las etapas de pre-procesamiento y reconocimiento de intenciones motoras en señales de EEG.

Capítulo 4: Formula la metodología de los marcos de trabajo propuestos para una Interfaz Cerebro-Computadora basada en imaginación motora (MI-BCI) .

Capítulo 5: Presenta los resultados de los marcos de trabajo establecidos en la fase de pruebas, analizando su funcionamiento con datos orientados a detección de MI-BCI para múltiples clases.

Capítulo 6: Establece las conclusiones finales fundamentadas sobre el análisis de los resultados y la formulación de trabajos futuros basados en las experiencias obtenidas.

---

## Fundamentos de Interfaces Cerebro-Computadora

---

---

<b>2.1</b>	<b>Conceptos en Interfaces Cerebro-Computadora</b>	<b>10</b>
<b>2.2</b>	<b>Anatomía del Cerebro</b>	<b>10</b>
2.2.1	Neuronas	10
2.2.2	Regiones del Cerebro	12
2.2.3	Lóbulos del Cerebro	14
<b>2.3</b>	<b>Señales de Electroencefalograma</b>	<b>15</b>
2.3.1	Adquisición de señales EEG	16
2.3.2	Caracterización de las señales EEG	17
2.3.3	Artefactos de ruido en señales EEG	19
<b>2.4</b>	<b>Imaginación Motora</b>	<b>21</b>
<b>2.5</b>	<b>Conclusiones del capítulo</b>	<b>23</b>

---

La identificación de intenciones motoras en señales de EEG es una tarea compleja que requiere lograr una correcta manipulación de los datos, y aplicación de técnicas de aprendizaje automático, para obtener los patrones comunes entre cada tipo de pensamiento.

En este capítulo se realiza una revisión de los aspectos teóricos que fundamentan la aplicación de las BCI orientadas a la detección de intenciones motoras. Se estudiará el comportamiento de las señales EEG, las principales barreras que se presentan al momento de clasificarlas y los aspectos biológicos que intervienen en el procesamiento de las señales de intenciones motoras.

## 2.1. Conceptos en Interfaces Cerebro-Computadora

**Definición 2.1.1 (Neuronas).** Una neurona es el componente principal del sistema nervioso, cuya función principal es recibir, procesar y transmitir información a través de señales químicas y eléctricas gracias a la excitabilidad eléctrica de su membrana plasmática. Están especializadas en la recepción de estímulos y conducción del impulso nervioso entre ellas (en forma de potencial de acción), mediante conexiones llamadas sinapsis, o con otros tipos de células como, por ejemplo, las fibras musculares de la placa motora.

**Definición 2.1.2 (Cerebro).** Es la parte más grande del encéfalo. Se divide en dos hemisferios o mitades que se llaman hemisferios cerebrales. Las áreas internas del cerebro controlan el funcionamiento de los músculos; también controlan el habla, el pensamiento, las emociones, la lectura, la redacción y el aprendizaje.

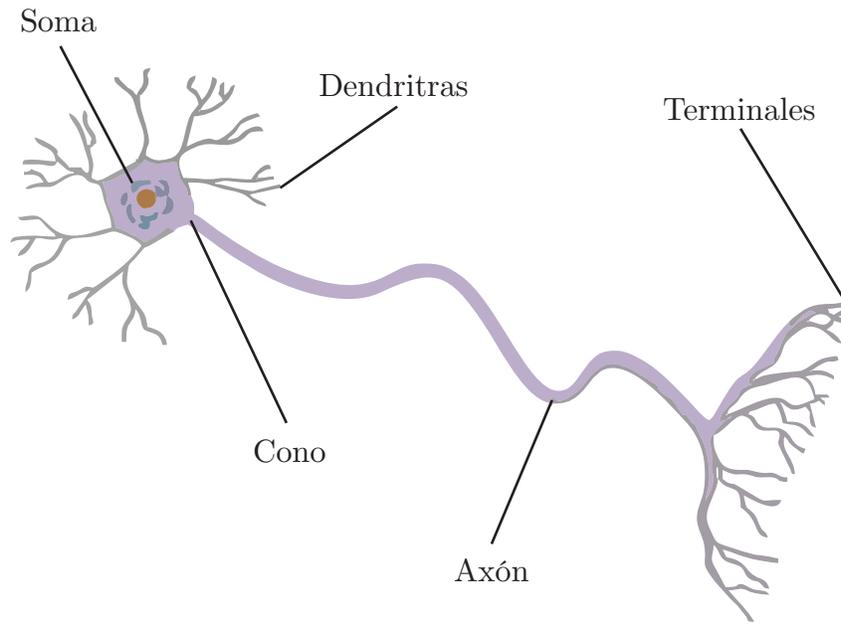
**Definición 2.1.3 (Interfaz Cerebro-Computadora).** Una BCI es un sistema de inteligencia artificial capaz de reconocer un determinado conjunto de patrones en las señales cerebrales. Está compuesto por cinco etapas: adquisición de la señal, preprocesamiento o mejora de la señal, extracción de características, clasificación e interfaz de control.

## 2.2. Anatomía del Cerebro

### 2.2.1. Neuronas

Una neurona es una célula especializada del sistema nervioso que se encarga de transmitir información a través de impulsos eléctricos y químicos. Las neuronas se comunican entre sí a través de conexiones llamadas sinapsis, y son responsables de procesar y transmitir información sensorial, motora y cognitiva. En el cerebro, las neuronas forman redes complejas que permiten la realización de funciones complejas como el pensamiento, la memoria y el

movimiento. Cada neurona consta de un cuerpo celular, dendritas que reciben señales de otras neuronas, un axón que envía señales a otras células, y terminales axónicos que se comunican con las dendritas de las células vecinas.



**Figura 2.1.** Descripción general de la anatomía de una neurona, compuesta por el soma, las dendritas, el cono, el axón y los terminales.

Como se observa en la Figura 2.1, las partes principales de una neurona son las siguientes:

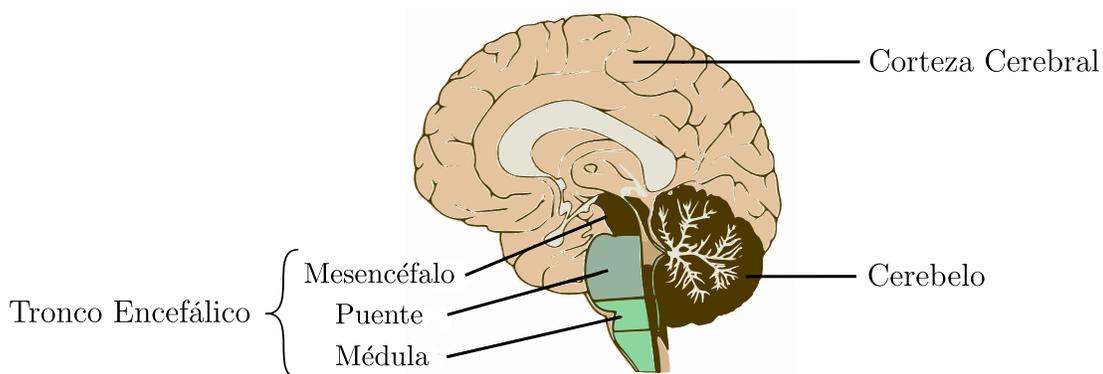
1. **Cuerpo celular o soma:** Es la parte de la neurona que contiene el núcleo y otros orgánulos celulares necesarios para la función celular normal.
2. **Dendritas:** Son las ramificaciones que se extienden desde el soma y reciben señales de otras neuronas o células sensoriales.
3. **Axón:** Es una prolongación alargada de la célula que lleva señales eléctricas y químicas desde el soma hacia otras células.
4. **Cono de inicio:** Es el sitio donde comienza el axón, y donde se generan los potenciales de acción que viajan a lo largo del axón.
5. **Terminales axónicos:** Son las terminaciones del axón que establecen contactos con otras células, como neuronas o músculos, y liberan neurotransmisores para transmitir información.

Además de estas partes, también se encuentran en las neuronas otros elementos como las vainas de mielina, que son capas aislantes de lípidos que recubren algunos axones para mejorar la transmisión de señales eléctricas, y las sinapsis, que son las conexiones especializadas entre neuronas donde se lleva a cabo la comunicación química.

### 2.2.2. Regiones del Cerebro

El cerebro es un órgano que envía y recibe señales químicas y eléctricas por todo el cuerpo. Diferentes señales controlan distintos procesos, y el cerebro interpreta cada una de ellas. Algunos mensajes se mantienen en el cerebro, mientras que otros se transmiten a través de la columna vertebral y la vasta red de nervios del cuerpo hasta las extremidades distantes. Para ello, el sistema nervioso central cuenta con miles de millones de neuronas (células nerviosas).

Como se indica en la Figura 2.2, el cerebro puede dividirse en tres regiones llamadas encéfalo, tronco encefálico y cerebelo.



**Figura 2.2.** Regiones del Cerebro.

#### 2.2.2.1. Encéfalo

El encéfalo (parte frontal del cerebro) está comprendido por la corteza cerebral (materia gris) y la materia blanca contenida en su centro. El encéfalo es la parte más grande del cerebro, inicia y coordina el movimiento, y regula la temperatura. Otras áreas del cerebro permiten el habla, el juicio, el pensamiento y el razonamiento, la resolución de problemas, las emociones y el aprendizaje. Otras funciones están relacionadas con la visión, el oído, el tacto y otros sentidos.

**Definición 2.2.1 (Corteza Cerebral).** Describe la cubierta externa de materia gris del cerebro. La corteza tiene una gran superficie debido a sus pliegues, y comprende aproximadamente la mitad del peso del cerebro. La corteza cerebral está dividida en dos mitades o hemisferios. Está cubierta de crestas (giros) y pliegues (surcos). Las dos mitades se unen en un surco grande y profundo (la fisura interhemisférica, también conocida como fisura longitudinal medial) que va desde la parte frontal de la cabeza hasta la parte posterior. El hemisferio derecho controla el lado izquierdo del cuerpo, y el izquierdo controla el lado derecho. Las dos mitades se comunican entre sí a través de una gran estructura en forma de C de materia blanca y vías nerviosas llamada cuerpo caloso. El cuerpo caloso está en el centro del cerebro.

#### 2.2.2.2. Tronco Encefálico

El tronco encefálico, parte media del encéfalo, conecta el cerebro con la médula espinal. Incluye el mesencéfalo, el puente de varolio y la médula.

**Definición 2.2.2 (Mesencéfalo).** El mesencéfalo es una estructura muy compleja con diversos grupos de neuronas (núcleos y colículos), vías neuronales y otras estructuras. Estas características facilitan diversas funciones, desde la audición y el movimiento hasta el cálculo de respuestas y cambios ambientales. El mesencéfalo también contiene la sustancia negra, una zona afectada por la enfermedad de Parkinson rica en neuronas dopaminérgicas y parte de los ganglios basales, que permiten el movimiento y la coordinación.

**Definición 2.2.3 (Puente de Varolio).** El puente de Varolio es la conexión entre el mesencéfalo y la médula, origen de cuatro de los 12 nervios craneales, que permite una serie de actividades como la producción de lágrimas, la masticación, el parpadeo, el enfoque de la visión, el equilibrio, la audición y la expresión facial.

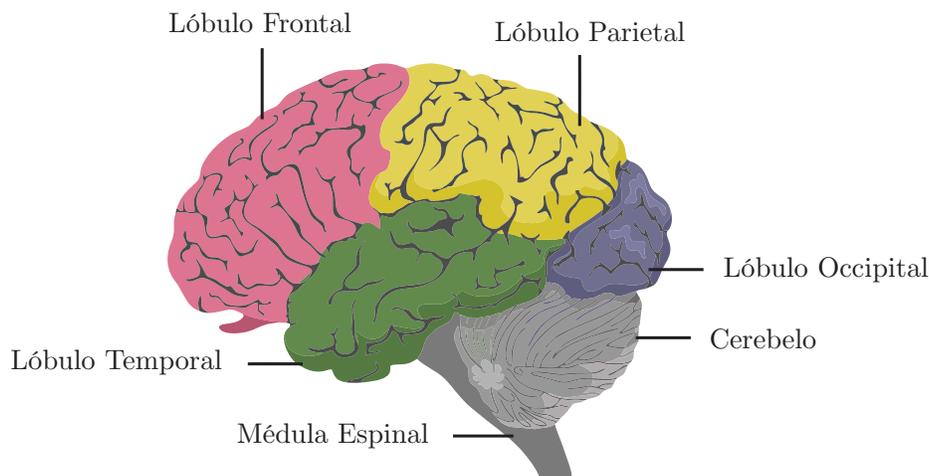
**Definición 2.2.4 (Médula).** En la parte inferior del tronco encefálico, la médula se encuentra con la médula espinal. La médula es esencial para la supervivencia. Las funciones de la médula regulan muchas actividades corporales, como el ritmo cardiaco, la respiración, el flujo sanguíneo, los niveles de oxígeno y dióxido de carbono. La médula produce actividades reflejas como estornudar, vomitar, toser y tragar.

### 2.2.2.3. Cerebelo

El cerebelo es una porción del cerebro del tamaño de un puño situada en la parte posterior de la cabeza, debajo de los lóbulos temporal y occipital, y encima del tronco encefálico. Al igual que la corteza cerebral, tiene dos hemisferios. La porción externa contiene neuronas y la zona interna se comunica con la corteza cerebral. Su función es coordinar los movimientos musculares voluntarios, mantener la postura, el equilibrio y el balance. También se ha explorado el papel del cerebelo en el pensamiento, las emociones y el comportamiento social, así como su posible implicación en la adicción, el autismo y la esquizofrenia.

### 2.2.3. Lóbulos del Cerebro

El cerebro está dividido en dos secciones denominadas hemisferios, a través de una fisura longitudinal. Al mismo tiempo, en cada hemisferio cerebral se han determinado cuatro secciones diferentes llamadas lóbulos: frontal, parietal, temporal y occipital, representadas en la Figura 2.3. A pesar de que cada lóbulo desempeña funciones específicas en la actividad humana, la mayoría de estas funciones implican la cooperación de múltiples áreas en ambos hemisferios. Por lo tanto, aunque el lóbulo frontal desempeña un papel esencial en el movimiento de las partes del cuerpo, regiones de los lóbulos parietal, temporal y occipital de ambos hemisferios también participan en los procesos de coordinación del movimiento.



**Figura 2.3.** Vista lateral de la ubicación de los lóbulos del Cerebro.

**Definición 2.2.5 (Lóbulo Frontal).** El lóbulo frontal, el mayor del cerebro y situado en la parte frontal de la cabeza, interviene en las características de la personalidad, la toma de decisiones y el movimiento. En el reconocimiento del olfato suelen intervenir partes del lóbulo frontal. El lóbulo frontal contiene el *área de Broca*, asociada a la capacidad del habla.

**Definición 2.2.6 (Lóbulo Parietal).** El lóbulo parietal, la parte media del cerebro, ayuda a la persona a identificar objetos y a comprender las relaciones espaciales con las que se compara el propio cuerpo con los objetos que rodean a la persona. El lóbulo parietal también interviene en la interpretación del dolor y el tacto en el cuerpo. El lóbulo parietal alberga el *área de Wernicke*, que ayuda al cerebro a comprender el lenguaje hablado.

**Definición 2.2.7 (Lóbulo Occipital).** El lóbulo occipital se encuentra en la parte posterior del cerebro y es el principal centro de procesamiento de la información visual. Interviene en la interpretación de la visión, la formación de recuerdos visuales y la integración de las percepciones visuales con la información espacial.

**Definición 2.2.8 (Lóbulo Temporal).** A los lados del cerebro, los lóbulos temporales intervienen en la memoria a corto plazo, el habla, el ritmo musical y cierto grado de reconocimiento de olores.

## 2.3. Señales de Electroencefalograma

La Electroencefalografía es un estudio no invasivo que mide la actividad eléctrica en el cerebro mediante pequeños discos, llamados electrodos, colocados sobre el cuero cabelludo. Se utiliza ampliamente en neurociencia y entornos clínicos para estudiar la función cerebral, diagnosticar trastornos neurológicos y controlar la actividad cerebral durante diversas tareas. Las neuronas del cerebro se comunican a través de impulsos eléctricos, los cuales con instrumentación adecuada, pueden ser captados y amplificados para interpretar la actividad de diferentes zonas del cerebro en forma de líneas onduladas almacenadas en un registro electroencefalográfico. Las ondas cerebrales EEG registradas, reflejan la actividad eléctrica en la que se basan las funciones mentales superiores, ya que la función cerebral normal implica una actividad eléctrica continua de las neuronas.

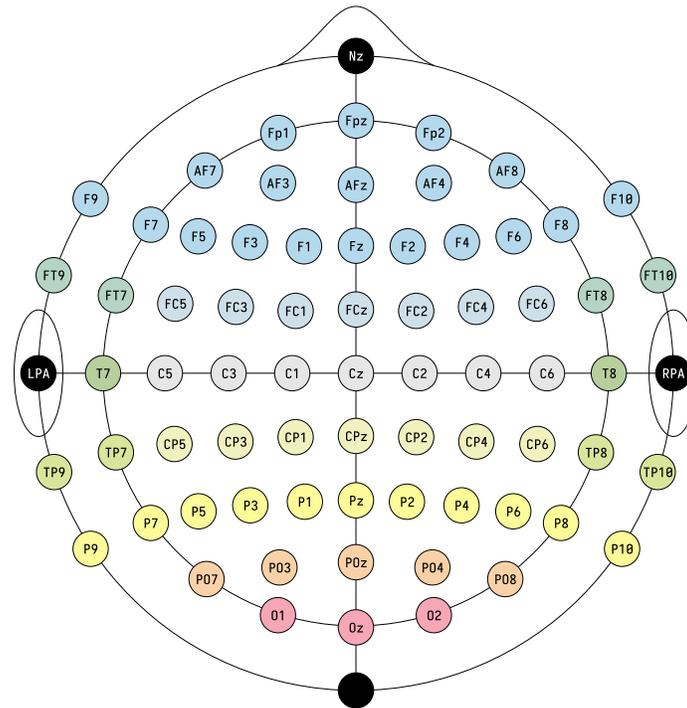
### 2.3.1. Adquisición de señales EEG

Constantemente, la interacción bioquímica entre las neuronas al momento de comunicarse entre ellas, produce una pequeña actividad eléctrica generada por la actividad sináptica en la superficie de la corteza del cerebro, y no por potenciales de acción en la materia blanca interna del mismo. La actividad eléctrica generada por la sinapsis entre solo dos neuronas es muy débil para ser registrada, pero cuando millones de neuronas se sincronizan, el campo eléctrico generado se puede medir externamente con dispositivos de medición especializados. Con la finalidad de registrar estos potenciales generados por la actividad neuronal de forma no invasiva, es necesario utilizar electrodos colocados en la superficie del cuero cabelludo, y conectarlos a un dispositivo de encefalografía con la capacidad de medir las diferencias de voltaje entre varias zonas corticales.

A rasgos generales, existen dos tipos de electrodos no invasivos en función de la forma en que son aplicados: a) Los primeros corresponden a los electrodos húmedos, los cuales, son electrodos generalmente fabricados a base de plata/cloruro de plata (Ag/AgCl) que utilizan un material en gel electrolítico como conductor entre la piel y el electrodo. b) Los segundos corresponden a los electrodos secos que consisten de un solo metal que actúa como conductor entre la piel y el electrodo, generalmente son fabricados de acero inoxidable.

De igual forma, los electrodos también pueden clasificarse de acuerdo a su funcionamiento como electrodos activos o pasivos. Los electrodos activos cuentan con un módulo de pre-amplificación inmediatamente después del material conductor entre la piel y el electrodo. Esto permite amplificar la señal antes de que se sumen ruidos adicionales entre el electrodo y el sistema de adquisición de la señal. Por otro lado, los electrodos pasivos no cuentan con un módulo de pre-amplificación, en su lugar, simplemente se extiende la conexión desde el material conductor hasta el dispositivo de adquisición de la señal.

Como se observa en la Figura 2.4, los electrodos se colocan sobre los lóbulos frontal, temporal, parietal y occipital, utilizando el sistema de colocación estándar 10/20 [Klem et al., 1999]. Aunque en este sistema hay hasta 75 posiciones diferentes, para aplicaciones orientadas a BCI puede ser suficiente utilizar entre 8 a 32 electrodos [Usakli, 2010]. Los números pares se refieren al hemisferio derecho y los números impares se refieren al hemisferio izquierdo, los electrodos ubicados en el eje central del cráneo se representan con la letra  $Z$ .



Fp: Frontal posterior. A: Apófisis mastoides.  
 F: Frontal. P: Parietal.  
 T: Temporal. O: Occipital.  
 C: Central.

**Figura 2.4.** Ubicación de electrodos no invasivos en el cuero cabelludo correspondiente al sistema de localización 10/20 [Klem et al., 1999].

### 2.3.2. Caracterización de las señales EEG

Con respecto a la naturaleza de las señales EEG, cada onda es un tren continuo de picos y valles, y la frecuencia de onda, expresada en Hertz (Hz), es el número de picos en un segundo. La amplitud de las ondas cerebrales refleja la actividad sincrónica de muchas neuronas, y no el grado de actividad eléctrica de neuronas individuales. Las señales EEG son complejas y dependientes de factores biológicos particulares para cada individuo, su amplitud varía desde unos pocos microvoltios hasta aproximadamente  $100 \mu\text{V}$  en un rango de frecuencia entre 0.1 Hz hasta 100 Hz. Por esta razón, cada persona tiene un patrón de ondas cerebrales único, tal y como sucede con las huellas dactilares [Marieb and Hoehn, 2010].

No obstante, como se observa en el Tabla 2.1, las señales electroencefalográficas que se registran en el estado de vigilia (también llamadas ritmos de fondo) se pueden agrupar en

cinco clases correspondientes al rango de frecuencia en la que operan:

1. Delta ( $\delta$ ), menor a 4 Hz,
2. Theta ( $\theta$ ), ente 4-7 Hz,
3. Alfa ( $\alpha$ ), entre 8-13 Hz,
4. Beta ( $\beta$ ), entre 14-30 Hz,
5. Gamma ( $\gamma$ ), mayor a 30 Hz.

**Tabla 2.1.** Clases de señales de electroencefalograma en función del rango de frecuencia [Sörnmo and Laguna, 2005].

Clase	Descripción
Delta (menor a 4 Hz)	Son ondas de gran amplitud, son frecuentes durante el sueño profundo o cuando se suprime el sistema activador reticular, como durante la anestesia. En adultos despiertos indican daño cerebral.
Theta (4-7 Hz)	Son señales aún más irregulares que las Delta. Aunque son comunes en los niños, las ondas Theta son infrecuentes en adultos despiertos pero pueden aparecer al concentrarse.
Alfa (8-13 Hz)	Señales relativamente regulares y rítmicas, de baja amplitud y sincrónicas. En la mayoría de los casos indican que el cerebro está en un estado de estado de vigilia.
Beta (14-30 Hz)	Son señales rítmicas, pero menos regulares que las ondas alfa, y con una frecuencia mayor. Las ondas beta se producen cuando estamos mentalmente alerta, como cuando nos concentramos en algún problema o estímulo visual.
Gamma (mayor a 30 Hz)	Son ondas cerebrales con una mayor frecuencia y menor amplitud. Aparecen en estados de vigilia y su presencia está relacionada con la aparición de la consciencia, con la ampliación del foco de atención y con la gestión de la memoria.

La interpretación de la información contenida en estas bandas es relativa, y depende de factores biológicos de cada individuo como su edad o salud mental. Por esta razón, para

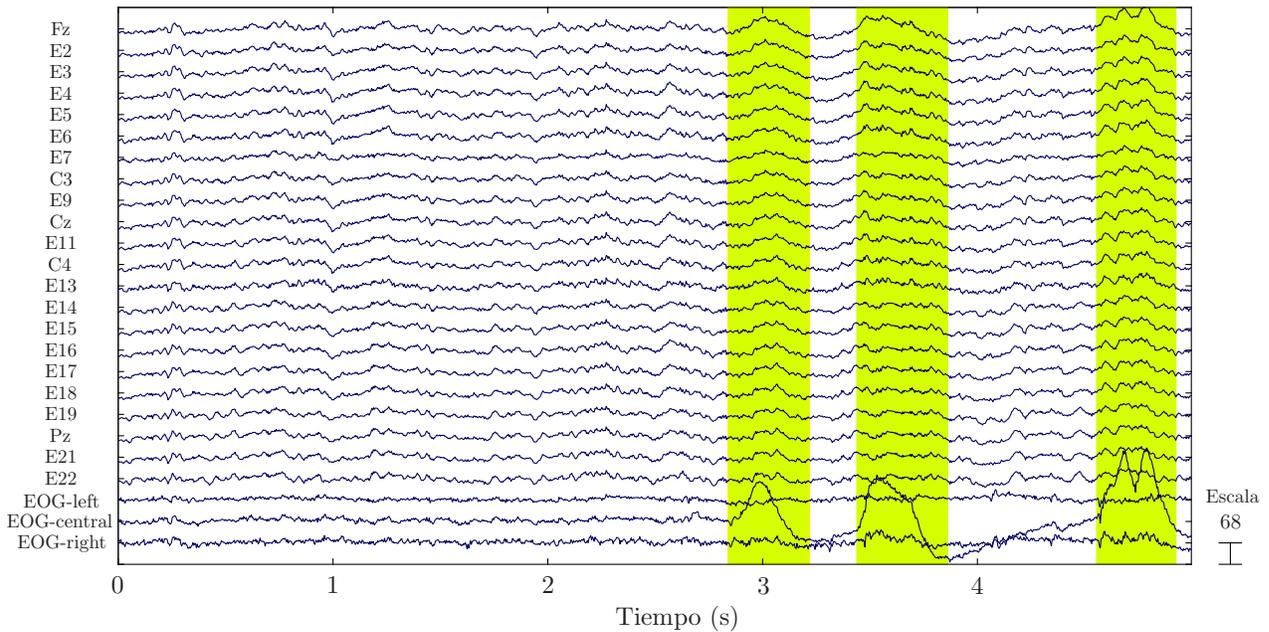
determinar un estado de normalidad o anormalidad en el comportamiento de los componentes frecuenciales en las señales EEG es necesario tomar en cuenta estos factores en conjunto con técnicas adecuadas de procesamiento de los datos. Por ejemplo, tal como indica [Sörnmo and Laguna \[2005\]](#), el EEG de un recién nacido es drásticamente diferente al de un adulto y tiene, en general, un contenido de frecuencia considerablemente mayor.

### 2.3.3. Artefactos de ruido en señales EEG

Las señales EEG se transmiten a través de tejido, hueso y cabello antes de ser registradas, y para entonces su amplitud es más tenue en comparación a la señal que puede adquirirse internamente en la corteza motora del cerebro [[Sörnmo and Laguna, 2005](#)]. Por esta razón, los electrodos utilizados en las mediciones de EEG son elementos muy sensibles y propensos a detectar ciertos artefactos de ruido que se suman a las señales generadas por la actividad neuronal. Estos artefactos representan una perturbación indeseable que puede estar originada por factores biológicos del individuo o factores no biológicos relacionados con las condiciones de la medición. Entre los artefactos de origen biológico se encuentran los movimientos oculares, los movimientos musculares (como los producidos por la actividad del músculo facial) y el pulso cardíaco. Estos artefactos pueden ser eliminados mediante técnicas de filtrado, como el filtrado pasa-banda o el filtrado adaptativo.

Por su parte, los artefactos originados por el movimiento ocular se producen debido a los cambios en la orientación del dipolo retina y córnea, y a las alteraciones del contacto de la córnea con el párpado producidos al momento del parpadeo [[Schlögl et al., 2007](#)]. Como se observa en la Figura 2.5, estos eventos oculares generan señales que pueden registrarse mediante un electro-oculograma (*Electrooculogram*, EOG), poseen amplitudes entre  $100 \mu\text{V}$  y  $3500 \mu\text{V}$  con un rango de frecuencia entre 0.1 Hz y 20.0 Hz, los cuales se propagan por la superficie de la piel y son registrados por los electrodos al mismo tiempo que las señales de la actividad neuronal. Los electrodos ubicados en el lóbulo frontal más cercanos a los ojos (Fz, E2, E3, E4, E5, E6, E7) son los más afectados por la presencia de artefactos de movimiento ocular.

Por otro lado, la contaminación de las señales EEG debido a la actividad muscular es causada por la contracción o estiramiento de cualquier grupo muscular próximo a los sitios de registro de las señales. Estos artefactos pueden ser registrados por un electro-miograma (*Electromiogram*, EMG), y su eliminación representa un problema de alta complejidad debido a que su amplitud y forma de onda (con frecuencias entre 0 Hz y 200 Hz) son afectados significativamente de acuerdo al nivel de la actividad muscular. Por lo tanto, es difícil extraer



**Figura 2.5.** Identificación de artefactos oculares obtenidos con lecturas EOG (selección marcada por las bandas amarillas) con escala en  $\mu V$ , y su efecto en una señal EEG obtenida de la base de datos BCI Competition IV dataset 2a [Tangermann et al., 2012].

la actividad muscular de una medición de EEG en comparación con los artefactos EOG y el seguimiento ocular [Jiang et al., 2019].

Los artefactos cardíacos pueden introducirse debido a la señal eléctrica emitida por el corazón (*Electrocardiogram*, ECG) [Lee et al., 2015] o cuando los electrodos se colocan cerca de un vaso sanguíneo (artefacto de pulso), en el que el movimiento de expansión y contracción se debe al corazón [Goncharova et al., 2003]. Las señales eléctricas ECG emitidas por la actividad cardíaca poseen un patrón regular característico y se pueden registrar aparte de la actividad cerebral, por lo que utilizando una forma de onda de referencia es posible eliminar dichos artefactos [Jiang et al., 2019].

En cuanto a los artefactos de origen no biológico, se incluyen las interferencias electromagnéticas (*Electromagnetic Interference*, EMI) causadas por equipos electrónicos cercanos (*e.g.*, monitores de ordenador, teléfonos móviles), las interferencias de la línea eléctrica y las interferencias causadas por cables mal conectados. Estos artefactos pueden ser minimizados mediante la implementación de equipos de medición adecuados, una correcta colocación de los electrodos y la implementación de blindajes electromagnéticos.

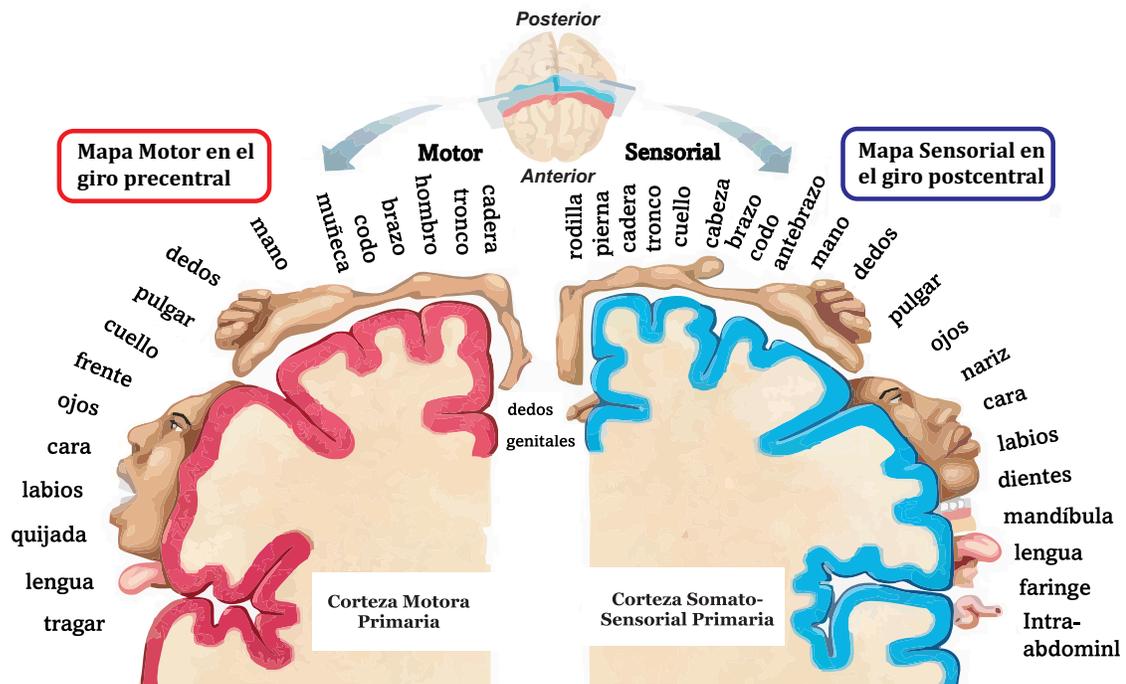
Para tratar de reducir el impacto negativo que estas diferentes fuentes o artefactos de ruido pueden generar sobre el procesamiento de las señales, los principales métodos hasta ahora propuestos en la literatura están basados en algoritmos de descomposición de señales como el Análisis de Componentes Independientes (*Independent Component Analysis*, ICA), la transformada Wavelet o la reconstrucción del subespacio de artefactos (*Artifact Subspace Reconstruction*, ASR), donde se busca rechazar los componentes de la señal relacionados con el ruido y separarlos de los componentes relacionados con la señal de interés. De igual forma, las técnicas de filtrado y análisis en tiempo-frecuencia ofrecen una amplia gama de soluciones para el tratamiento de estas fuentes de ruido. Por esta razón, los filtros pasa-banda, filtros adaptativos y la transformada de Fourier de tiempo corto (*Short-Time Fourier Transform*, STFT) también son algunas propuestas ampliamente utilizadas para la eliminación de ruido en los desarrollos de Interfaces Cerebro-Computadoras.

## 2.4. Imaginación Motora

La implementación de Interfaces Cerebro-Computador puede realizarse con una variedad de dispositivos diferentes y estar orientados a diversas funciones. La aplicación tomada como base para esta investigación se conoce como Interfaz Cerebro Computador de Movimiento Imaginario (*Motor Imagery - Brain Computer Interface*, MI-BCI) . En este tipo de aplicación, el usuario piensa en efectuar un movimiento repetidamente y el sistema debe identificar el tipo de movimiento empleado, por ejemplo, discernir entre el pensamiento motor de manos o pies.

La corteza motora es una parte del cerebro que controla el movimiento voluntario, y los movimientos de cada parte del cuerpo pueden asociarse a la actividad de una región diferente de la corteza motora [Garcia-Ramos and Cruces-Álvarez, 2020], tal como se ejemplifica en la Figura 2.6. Dicha actividad neuronal puede ser medida y detectada a través de un fenómeno conocido como ritmos sensoriomotores.

El principio de funcionamiento de la imaginación motora aplicada a interfaces de control cerebro computadora se basa en dos fenómenos observables en las bandas de frecuencia  $\alpha$  y  $\beta$  de las señales EEG, llamados sincronización relacionada a eventos (*Event-Related Synchronization*, ERS) y desincronización relacionada a eventos (*Event-Related Desynchronization*, ERD) . Como se puede observar en la Figura 2.7, cuando el usuario no está ejecutando acciones o pensamientos motores, las áreas del cerebro relacionadas con el movimiento están en reposo y la potencia de la señal está en su punto máximo, manifestando un ERS. Por el contrario, cuando

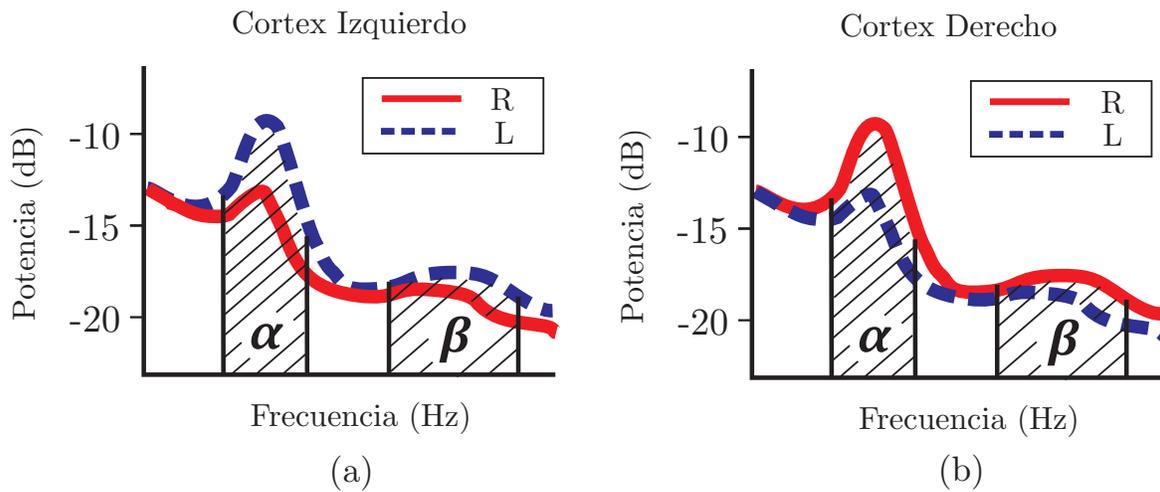


**Figura 2.6.** Distribución de regiones en la corteza motora primaria y la corteza somato-sensorial primaria del cerebro, y su influencia en los diferentes tipos de intención motora [Marieb and Hoehn, 2010].

el usuario realiza un movimiento o piensa en la ejecución del mismo, la zona correspondiente a este movimiento en la corteza motora del cerebro se activa y la potencia de la señal descende, manifestando un ERD.

Gracias a estos cambios de potencia, el usuario puede imaginar un movimiento sin ejecutarlo físicamente y un dispositivo electrónico puede leer dichas señales de EEG, para posteriormente mediante técnicas de aprendizaje automático, discriminar la intención de realizar esta acción respecto a otras opciones y así convertirla en una señal de control para una computadora. Esto abre la posibilidad de utilizarlo en casos en los que el usuario se encuentre en un estado de discapacidad motora.

Este trabajo de investigación analizará diferentes técnicas de aprendizaje automático y sus etapas de pre-procesamiento, extracción, selección de características y clasificación para la creación de un sistema MI-BCI .



**Figura 2.7.** Efecto de la ERD sobre la densidad espectral de potencia de la señal de EEG. (a) Atenuación de potencia en las bandas  $\alpha$  y  $\beta$  relacionadas con la señal de pensamiento motor de brazo derecho (R). (b) Atenuación de potencia en las bandas  $\alpha$  y  $\beta$  relacionadas con la señal de pensamiento motor de brazo izquierdo (L).

## 2.5. Conclusiones del capítulo

En este capítulo se mencionaron aspectos morfológicos relacionados con la anatomía del cerebro, el origen de las señales obtenidas por la lectura de encefalograma, y los aspectos biológicos y físicos que intervienen en la implementación de las Interfaces Cerebro-Computadora, basadas en la interpretación de la imaginación motora.

Adicionalmente, se realizó una breve introducción sobre los aspectos técnicos relacionados a la adquisición de las señales de encefalografía, y las circunstancias que propician la aparición de componentes de ruido, una de los principales problemas presentes en la clasificación de señales relacionadas con la imaginación motora.

En el Capítulo 3, se expondrán los métodos utilizados en la investigación para acondicionar y clasificar las señales EEG, tomando en cuenta sus características y los requerimientos de las aplicaciones de MI-BCI.

# CAPÍTULO 3

---

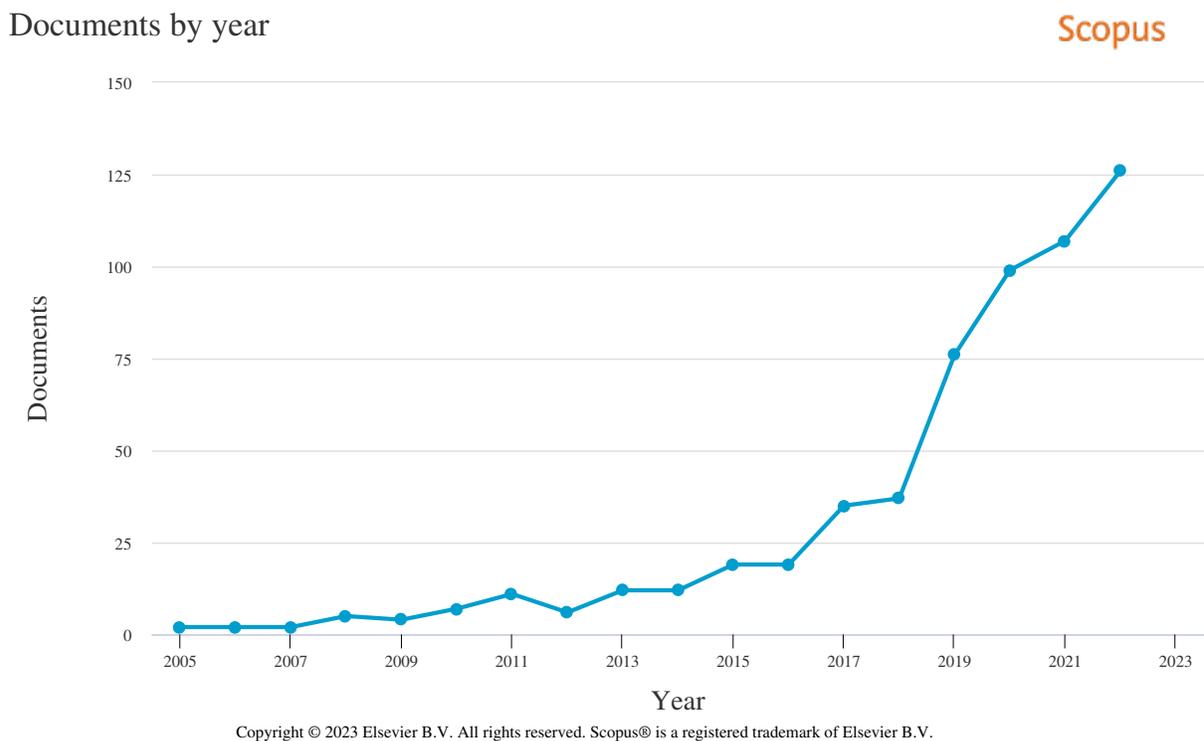
## Procesamiento de Señales EEG

---

<b>3.1</b>	<b>Trabajos Relacionados</b>	<b>25</b>
<b>3.2</b>	<b>Eliminación Artefactos de Ruido</b>	<b>27</b>
3.2.1	Filtro Digital Pasa-banda	27
3.2.2	Diseño de banco de filtros IIR	30
3.2.3	Descomposición de paquetes de onda	32
<b>3.3</b>	<b>Filtro Espacial</b>	<b>37</b>
3.3.1	Patrones Espaciales Comunes	37
3.3.2	Diagonalización Aproximada Conjunta	39
<b>3.4</b>	<b>Técnicas Selección de Características</b>	<b>40</b>
3.4.1	Selección de filtros espaciales	40
3.4.2	Selección por Valor Laplaciano	42
<b>3.5</b>	<b>Métodos de Clasificación</b>	<b>44</b>
3.5.1	Maquinas de Vectores de Soporte	44
3.5.2	K-Vecinos Más Cercanos	47
3.5.3	Análisis Discriminante Lineal	49

### 3.1. Trabajos Relacionados

En cuanto al procesamiento de señales de EEG y sistemas de reconocimiento aplicados a Interfaces Cerebro-Computadora basados en técnicas de aprendizaje automático, se puede observar un creciente interés descrito por los avances publicados en la literatura desde el año 2005, tal como se observa en la Figura 3.1.



**Figura 3.1.** Contribuciones en la literatura relacionadas con técnicas de aprendizaje automático aplicado a la detección de imaginación motora en BCI, consulta realizada en Scopus el 29 de marzo del 2023.

La presencia de diferentes fuentes de ruido y el comportamiento no estacionario de las señales de EEG, han obligado a la comunidad científica a desarrollar técnicas avanzadas de pre-procesamiento y algoritmos de aprendizaje automático, que aporten una correcta detección e interpretación de las señales cerebrales para aplicaciones en BCI [Zhang et al., 2021a].

La importancia que tiene el pre-procesamiento de las señales de EEG en las aplicaciones

para Interface Cerebro-Computadora, se enfatiza en las propuestas planteadas en la literatura. El trabajo desarrollado por [Raza et al. \[2015\]](#), implementan un sistema de clasificación basado en técnicas convencionales de aprendizaje automático (*Machine Learning*, ML), aplicando un pre-procesamiento basado en filtros pasa-banda para separar las señales originales en sub-bandas de interés. Como aporte en su investigación, demuestran la importancia de filtrar las bandas de frecuencia que ofrezcan extraer mejores características para la clasificación. Por otra parte, [Zahid et al. \[2020\]](#) propone una alternativa de procesamiento de las señales EEG basadas en filtrado espacial similar a [Raza et al. \[2015\]](#) pero complementada con algoritmos de inferencia estadística, buscando reducir problemas de óptimos locales. En sus trabajos, ambos autores proponen el cálculo de logaritmo de la varianza normalizada de las señales como atributo para la clasificación.

Los datos de EEG a menudo están contaminados con señales de ruido tales como los EOG, ECG y EMG, creado por movimientos oculares y parpadeos, pulsos cardíacos y actividades musculares, respectivamente [[Mohammadi and Mosavi, 2021](#)]. La presencia de estas fuentes de ruido en las señales de EEG genera una barrera importante en el proceso de clasificación de las actividades cerebrales. Es por esta razón, que técnicas de pre-procesamiento como la Descomposición de Paquetes Wavelet (*Wavelet Packet Decomposition*, WPD) [[Feng et al., 2019](#)], Descomposición de Modo Empírico (*Empirical Mode Decomposition*, EMD) [[Mwata-Velu et al., 2022b](#)] o la diferenciación en tiempo [[Fumanal-Idocin et al., 2022](#)] son aplicadas en las señales de EEG para reducir los efectos de distorsión provocados por las fuentes de ruido.

Por otra parte, poder extraer características robustas durante el procesamiento de las señales es una tarea indispensable para obtener resultados adecuados y superar las dificultades inherentes a la clasificación. [Tian and Liu \[2018\]](#) en su trabajo proponen un clasificador de intenciones motoras en señales EEG combinando información espacial, en tiempo y en frecuencia. Las señales originales son convertidas en espectrogramas que son clasificados por técnicas de aprendizaje profundo. En su propuesta, los autores también comprueban el efecto que tienen las diferentes bandas de frecuencia de la señal sobre la respuesta del clasificador, determinando que cada individuo analizado obtenía su máximo rendimiento en bandas de frecuencia distintas.

La aplicación de técnicas de aprendizaje profundo en la clasificación de intenciones motoras también se ha explorado con el objetivo de obtener mayores rendimientos. [Zhang et al. \[2021a\]](#), proponen una arquitectura de red llamada *EEG-Inception* basada en la red *InceptionTime*, la cual presenta buenas prestaciones en el procesamiento de datos en tiempo.

En esta propuesta, la señal original es utilizada como dato de entrada para la red sin utilizar técnicas de pre-procesamiento previas. Adicionalmente, los autores proponen el aumento de datos de entrenamiento utilizando una mezcla entre las señales y los ruidos extraídos de las mismas. Por su parte, [Mwata-Velu et al. \[2022a\]](#) proponen un análisis comparativo que estudia el efecto de la selección óptima de canales para la clasificación de señales EEG, utilizando la arquitectura de red *EEGNet* como clasificador de intenciones motoras basada en el movimiento de los dedos de la mano.

En resumen, observamos que los métodos del estado del arte constan de cuatro etapas principales: pre-procesamiento o eliminación de artefactos de ruido, extracción de características, selección de características y clasificación. A continuación explicaremos los métodos empleados en cada etapa, los cuales sirvieron de base en esta investigación.

## 3.2. Eliminación Artefactos de Ruido

La adquisición de señales de EEG es una herramienta que permite registrar la actividad eléctrica del cerebro necesaria para discernir entre los diferentes tipos de imaginación motora. Sin embargo, las señales de EEG pueden verse afectadas por varios tipos de artefactos de ruido que producen alteraciones no deseadas, como los originados por movimientos corporales, la actividad muscular, la interferencia electromagnética, entre otros.

Comúnmente, los artefactos de ruido suelen sumarse linealmente con la señal de actividad cerebral EEG, y pueden estar presentes en los registros de todos los electrodos dependiendo de la posición donde se encuentren ubicados. Por lo tanto, un importante paso de procesamiento previo a la interpretación del EEG es la eliminación de todo tipo de artefactos [[Radüntz et al., 2017](#)].

A continuación, se estudiarán dos métodos orientados a la eliminación de estas fuentes de ruido: 1) Banco de filtros digitales pasa-banda IIR Butterworth, 2) Descomposición de paquetes de onda.

### 3.2.1. Filtro Digital Pasa-banda

Un filtro digital es un proceso computacional que toma una secuencia de datos (la señal de entrada) y produce una nueva secuencia de datos diferente. Los datos de entrada

son transformados en otra secuencia numérica llamada señal de salida, de tal manera que el espectro de frecuencia de la salida es diferente al espectro de frecuencia de la señal de entrada. La implementación del filtro digital puede estar basada en una subrutina de software o una secuencia de circuitos integrados correctamente interconectados [Smith, 2007].

Matemáticamente, se puede definir un filtro digital  $\mathcal{F}_n$  como cualquier función que genere un resultado para cada valor entero de entrada  $x(n) \in \mathbb{Z}$ , donde  $\mathbb{Z}$  es el conjunto de los números enteros. Podemos expresar la relación entrada-salida de un filtro digital por la notación,

$$y(n) = \mathcal{F}\{x(n)\}, \quad (3.1)$$

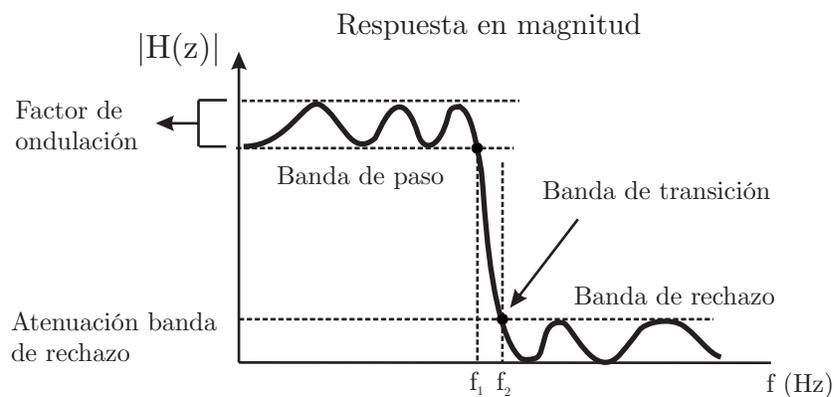
donde  $x(n)$  representa la señal de entrada, y  $y(n)$  es la señal de salida en el tiempo discreto  $n$ . El filtro  $\mathcal{F}$  también se puede definir como un operador en el espacio de la señal  $\mathcal{S} \in \mathbb{Z}$ . El operador  $\mathcal{F}$  transforma cada señal  $x \in \mathcal{S}$  a una nueva señal  $y \in \mathcal{S}$ . Si  $\mathcal{F}$  es lineal, puede ser llamado como un operador lineal en el espacio  $\mathcal{S}$ . Si, adicionalmente, el espacio de la señal  $\mathcal{S}$  está compuesto por señales de longitud finita, para muestras de tamaño  $N$ , *i.e.*,  $\mathcal{S} \subset \mathbb{R}^N$ , entonces cada filtro lineal  $\mathcal{F}$  puede ser denominado como una transformación lineal, la cual es representada por una matriz constante de tamaño  $N \times N$ .

**Definición 3.2.1 (Orden).** En filtros digitales, el orden del filtro se define como el retraso máximo en muestras o número de datos de entrada necesarios por el filtro para crear cada muestra de salida.

**Definición 3.2.2 (Frecuencia de Muestreo).** La frecuencia de muestreo se define como el número de muestras por segundo que se toman de una señal continua para formar una señal discreta o digital. En el caso de las señales en tiempo, las frecuencias se miden en Hertz (Hz) o ciclos por segundo. El teorema de muestreo de Nyquist-Shannon (principio de Nyquist) establece que la reconstrucción perfecta de una señal es posible cuando la frecuencia de muestreo es superior al doble de la frecuencia máxima de la señal muestreada.

Los filtros digitales aplicados en esta investigación se basan en la teoría de sistemas lineales e invariantes en el tiempo (*Linear time-invariant*, LTI), y en la transformada  $Z$ , para determinar una función de transferencia  $H(z)$ , que describe la transformación de la señal de entrada en la señal de salida. En el caso de los filtros pasa-banda, la función de transferencia se diseña para que permita mantener los componentes presentes entre un rango de frecuencias y atenúe los componentes frecuenciales que se encuentren fuera del rango de interés. La función de transferencia se puede expresar en términos de polos y ceros, que determinan las características de frecuencia del filtro, como sus frecuencias de corte y atenuación.

Para medir el rendimiento de un filtro en el dominio de la frecuencia se utilizan tres parámetros fundamentales: factor de caída, factor de ondulación y la atenuación de la banda de rechazo, como se muestra en la Figura 3.2. En casos donde se desean separar componentes frecuenciales muy próximos, el filtro debe tener un factor de caída rápido en su banda de transición, y para que las frecuencias en la banda de paso se procesen a través del filtro sin alteraciones, el filtro debe tener un factor de ondulación bajo. Por último, para bloquear adecuadamente las frecuencias de la banda de rechazo, es necesario tener una buena atenuación de la banda de rechazo.



**Figura 3.2.** Parámetros de rendimiento en respuesta de magnitud de un filtro digital.

Los filtros digitales pueden ser clasificados en dos grupos principales de acuerdo a su respuesta al recibir una señal impulso en su entrada, también llamada respuesta impulso. Por un lado, se encuentran los filtros de respuesta al impulso finita (*Finite Impulse Response*, FIR) cuya respuesta al impulso es de duración finita y están descritos por,

$$y(n) = \sum_{k=0}^{N-1} a_k x(n-k), \quad (3.2)$$

donde  $y(n)$  es la salida del filtro,  $x(n)$  es la señal de entrada al filtro,  $a_i$  son los coeficientes que determinan el comportamiento del filtro, y  $N$  determina el orden del filtro. En este caso, es posible observar que la salida del filtro FIR en cualquier momento depende solo de un número finito de muestras de la entrada.

El segundo grupo de filtros corresponde a los filtros de respuesta infinita al impulso (*Infinite Impulse Response*, IIR), los cuales son filtros recursivos que mezclan la señal de entrada

con realimentación retardada de la señal de salida y están descritos por,

$$y(n) = \sum_{k=0}^{M-1} a_k x(n-k) - \sum_{k=1}^N b_k y(n-k), \quad (3.3)$$

donde  $M$  corresponde al número de ceros,  $N$  corresponde al número de polos,  $a_k$  representa los coeficientes de lazo abierto y  $b_k$  corresponde a los coeficientes de retroalimentación del filtro. El orden del filtro está determinado por el valor máximo entre  $M - 1$  y  $N$ . Esta configuración logra que la salida del filtro IIR en cualquier momento dependa de una combinación de entradas actuales y pasadas.

Es importante señalar que la elección entre utilizar filtros FIR o IIR en el procesamiento de señales depende de los requisitos específicos del análisis y de las compensaciones involucradas. Los filtros FIR suelen ser mejores cuando se desea obtener una respuesta en fase lineal, un control preciso de la respuesta de frecuencia o una alta estabilidad. Mientras que los filtros IIR pueden presentar sensibilidad al ruido y a la acumulación de errores, proporcional al orden del filtro, debido a que incluyen los efectos de todos los datos anteriores [[Austerlitz, 2003](#)].

Sin embargo, los filtros IIR necesitan un orden menor y menos coeficientes para lograr características de respuesta de frecuencia similares, por lo cual, su implementación requiere un menor uso de recursos y requieren menos cálculos para lograr una respuesta de filtro determinada en comparación a los filtros FIR. Esto hace que los filtros IIR sean computacionalmente eficientes, especialmente para el procesamiento en tiempo real de señales EEG en aplicaciones de BCI donde se requiere un procesamiento de baja latencia. Adicionalmente, la propiedad de respuesta al impulso infinito de los filtros IIR puede resultar ventajosa para capturar y enfatizar componentes oscilatorios lentos de larga duración, como ondas  $\delta$  y  $\theta$  en las señales EEG, sin requerir una ventana de filtro excesivamente larga.

### 3.2.2. Diseño de banco de filtros IIR

La metodología más sencilla para construir un filtro IIR digital se basa en el diseño de un filtro analógico en tiempo continuo que contemple las características requeridas por la aplicación, y luego transformarlo al tiempo discreto a través del método de la transformada Z bilineal. Los filtros analógicos clásicos utilizados convencionalmente para el diseño de filtros IIR son el filtro Butterworth, Chebyshev tipo I, Chebyshev tipo II y Elíptico. Cada uno con comportamientos diferentes basados en las expresiones matemáticas utilizadas en el cálculo de

sus coeficientes  $a_k$  y  $b_k$ . De los cuatro tipos de filtros clásicos, el filtro Butterworth tiene una respuesta en magnitud plana tanto en la banda de paso como en la banda de rechazo, el filtro Chebyshev tipo I muestra ondas en la banda de paso pero su respuesta es plana en la banda de rechazo, el filtro Chebyshev tipo II posee una respuesta plana en la banda de paso pero tiene ondas en la banda de rechazo, y el filtro elíptico tiene ondulaciones en ambas bandas [Ambardar, 1999].

Tal como se explicará posteriormente en el Capítulo 4, el objetivo que debe cumplir el banco de filtros IIR es extraer la mayor cantidad de información en las diferentes bandas de frecuencia relacionadas con la imaginación motora en la señal EEG y rechazar los demás componentes que pudieran estar asociados a artefactos de ruido. Por lo tanto, es necesario construir filtros que tengan transiciones suaves y un ancho de banda que permita un solapamiento adecuado entre ellos.

Para cumplir los requerimientos del banco de filtros, se seleccionó el filtro Butterworth descrito por la siguiente función de transferencia en magnitud,

$$|H(j\Omega)|^2 = \frac{1}{1 + \epsilon^2 (\Omega/\Omega_c)^{2N}}, \quad (3.4)$$

donde  $\Omega$  es la frecuencia angular en rad/s,  $\Omega_c$  es la frecuencia de corte del filtro pasa baja en rad/s,  $\epsilon$  es el factor de desviación de la magnitud de la banda de paso unitaria del filtro y  $N$  es el orden del filtro. Los filtros Butterworth presentan una caída en su banda de transición más lenta que los otros filtros clásicos, y su respuesta de magnitud constante en la banda de paso, permite mantener intacto el comportamiento de la señal EEG en las bandas de frecuencia de interés [Kumar et al., 2017b].

La metodología aplicada para la construcción de filtros IIR digitales Butterworth implica los siguientes pasos:

1. Construcción de un filtro pasa baja con frecuencia de corte unitaria  $\Omega_c = 1.0$  rad/s. El diseño de estos filtros analógicos generalmente se basa en las especificaciones de frecuencia como la banda de paso y frecuencias de corte, así como especificaciones de magnitud como la atenuación de banda de paso máxima y atenuación de banda de rechazo mínima, para generar una función de transferencia de filtro con el orden más pequeño posible que cumpla o exceda las especificaciones [Ambardar, 1999]. En este caso, la estrategia de diseño busca convertir las especificaciones de frecuencia originales a aquellas aplicables a un prototipo de filtro pasa baja con una frecuencia de corte de 1 rad/s. La función de

transferencia del filtro Butterworth pasa baja con frecuencia de corte unitaria  $H_{LP}(s)$  está definida como,

$$H_{LP}(s) = \frac{1}{\prod_{k=1}^N (s - s_k)} \text{ con } s_k = e^{\frac{j(2k+N-1)\pi}{2N}}, \quad (3.5)$$

donde  $s_k$  corresponde a los polos de la función de transferencia que describen el comportamiento del filtro.

2. La función de transferencia  $H_{LP}(s)$  es transformada en un filtro pasa-banda con las frecuencias arbitrarias seleccionadas para el diseño  $H_{BP}(s)$ , realizando la siguiente sustitución en (3.5):

$$s \Leftarrow \frac{s^2 + \omega_0^2}{s B}, \quad (3.6)$$

esto es,

$$H_{BP}(s) = H_{LP}(s) \Big|_{s=\frac{s^2+\omega_0^2}{sB}}, \quad (3.7)$$

donde  $\omega_0$  corresponde al centro geométrico entre las frecuencias de corte  $\omega_l$  y  $\omega_h$  que delimitan la banda de paso del filtro pasa-banda que se desea, y  $B$  es el ancho de banda definido por,

$$B = \omega_h - \omega_l, \quad (3.8)$$

3. Por ultimo, se aplica la transformada Z bilineal a la función de transferencia del filtro pasa-banda  $H_{BP}(s)$  con el objetivo de obtener su equivalente en el dominio discreto (versión digital del filtro analógico calculado) realizando la siguiente sustitución,

$$s \Leftarrow \frac{2}{T} \frac{z-1}{z+1}, \quad (3.9)$$

esto es,

$$H_{HP}(z) = H_{BP}(s) \Big|_{s=\frac{2}{T} \frac{z-1}{z+1}}, \quad (3.10)$$

donde  $T$  es el periodo de muestreo del sistema.

### 3.2.3. Descomposición de paquetes de onda

La transformada wavelet es una técnica utilizada en el procesamiento de señales y análisis de datos. Se basa en la aplicación de una serie de operaciones matemáticas que implican la convolución de la señal original con diferentes funciones de onda denominadas *wavelets*,

seguida de la compresión y filtrado de los coeficientes resultantes. El objetivo principal de esta técnica es descomponer señales complejas en sus componentes básicas a diferentes resoluciones de tiempo y frecuencia.

La descomposición de paquetes de onda, o WPD, es una técnica de análisis de señales que extiende la transformada wavelet al permitir descomponer una señal en paquetes de sub-bandas más pequeñas que las descomposiciones wavelet tradicionales. Su implementación permite una mayor flexibilidad en el análisis de señales, ya que logra ajustar la descomposición a niveles de resolución en frecuencia específicas de interés.

La WPD comienza con la descomposición de la señal de entrada en dos sub-bandas de frecuencia mediante dos filtros digitales espejo en cuadratura, es decir un filtro FIR pasa baja  $h(n)$  y un filtro FIR pasa alta  $g(n)$  cada uno con longitud  $T$  expresada por,

$$g(n) = (-1)^{1-n}h(1-n), \quad 0 \leq n \leq T, \quad (3.11)$$

donde un filtro  $h(n)$  está asociado con una función de escalamiento  $\Phi(t)$  y un filtro  $g(n)$  está asociado con una función wavelet madre  $\Psi(t)$ . Ambas funciones son utilizadas para construir una serie de ondas base (wavelets hijas)  $\Psi_{j,k}^{2i}(t)$  y  $\Psi_{j,k}^{2i+1}(t)$  a través de traslaciones y escalas, lo que permite analizar diferentes componentes de frecuencia de la señal a varios niveles de descomposición de la forma:

$$\Psi_{j,k}^{2i}(t) = \frac{1}{\sqrt{2}} \Psi_{j,k}^{2i} \left( \frac{2^j k - t}{2^j} \right) = \sum_{n=0}^{T-1} h(n) \Psi_{j-1, 2k-n}^i(t), \quad (3.12)$$

$$\Psi_{j,k}^{2i+1}(t) = \frac{1}{\sqrt{2}} \Psi_{j,k}^{2i+1} \left( \frac{2^j k - t}{2^j} \right) = \sum_{n=0}^{T-1} g(n) \Psi_{j-1, 2k-n}^i(t), \quad (3.13)$$

donde  $\Psi^0(t) = \Phi(t)$ ,  $\Psi^1(t) = \Psi(t)$ ,  $j$  representa el nivel de descomposición (parámetro de escalamiento),  $k$  es el parámetro de posición en tiempo e  $i$  es el identificador de cada paquete de onda.

**Definición 3.2.3 (Filtro espejo en cuadratura).** Es un filtro utilizado para dividir la señal de entrada en dos bandas (Banda de frecuencias bajas y banda de frecuencias altas, simétricas al rededor de  $\frac{\pi}{2}$ ). Las señales de paso alto y paso bajo resultantes a menudo se reducen en un factor de 2, dando una representación de dos canales muestreada críticamente de la señal original. En este tipo de filtros se cumple la propiedad de potencia complementaria en la cual la suma de potencias de los filtros de paso alto y paso bajo es igual a 1.

**Definición 3.2.4 (Función Wavelet Madre).** Es una función matemática que tiene una duración finita, utilizada para descomponer una señal compleja en diferentes componentes de frecuencia y analizar cada componente de forma separada. La función wavelet madre debe cumplir dos condiciones fundamentales para ser utilizada en el análisis de wavelets, en primer lugar el promedio de la onda debe ser cero y en segundo lugar, la función debe tener energía finita y normalizada. Ejemplos comunes de funciones wavelet madre incluyen la wavelet Haar, la wavelet Daubechies y la wavelet Morlet.

Posteriormente, los coeficientes  $\mathbf{d}_j^p$  que componen cada uno de los paquetes de onda  $p$  de cada nivel  $j$  de la señal original  $F(t)$  se calculan utilizando las siguientes ecuaciones,

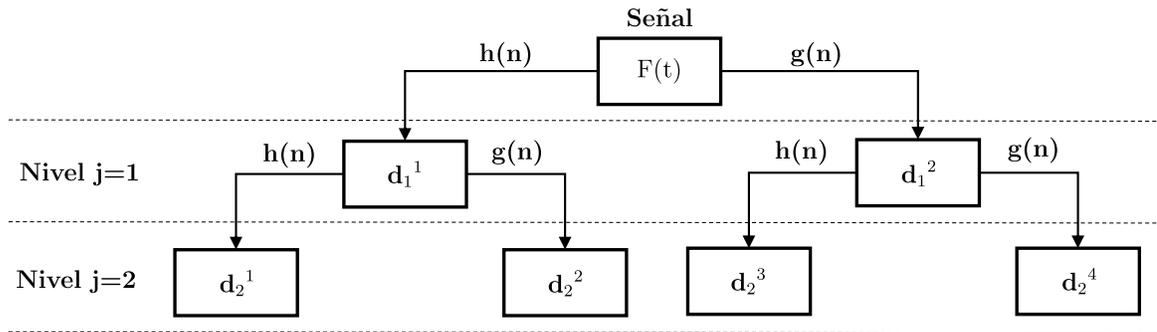
$$d_j^{2i}(k) = \int_{-\infty}^{\infty} F(t) \Psi_{j,k}^{2i}(t) dt = \sum_{n=0}^{T-1} h(n) d_{j-1}^i(2k - n) \quad (3.14)$$

$$d_j^{2i+1}(k) = \int_{-\infty}^{\infty} F(t) \Psi_{j,k}^{2i+1}(t) dt = \sum_{n=0}^{T-1} g(n) d_{j-1}^i(2k - n), \quad (3.15)$$

Cada vector  $\mathbf{d}_j^p$  contiene  $N/2^j$  coeficientes donde  $N$  corresponde al número de muestras de la señal original  $F(t)$ . El valor de estos coeficientes calculados durante la descomposición representan el grado de similaridad entre la función wavelet, con una escala de  $2^j$  en frecuencia y desplazada en cada instante  $k$ , y la señal  $F(t)$  en cada sub-banda de frecuencia  $p$ . Por esta razón, la función wavelet  $\Psi(t)$  utilizada en cada nivel de la descomposición debe ser elegida de tal manera que pueda capturar adecuadamente tanto la información de tiempo como de frecuencia de la señal.

Este proceso se aplica recursivamente en cada sub-banda resultante, dividiéndola a su vez en dos sub-bandas adicionales, y así sucesivamente por cada nivel de descomposición. Como resultado de este proceso, se genera una estructura en forma de árbol, como se muestra en la Figura 3.3, donde cada nodo  $\mathbf{d}_j^p$  está organizado desde  $p = 1, 2, \dots, 2^j$  y representa una sub-banda en un nivel particular de la descomposición.

Uno de los beneficios de la descomposición con WPD, es que permite eliminar componentes no deseados de una señal. Al hacerlo se pretende reducir o eliminar los coeficientes de la descomposición que tengan poca correlación con la señal de interés, y que por lo tanto, tengan el potencial de estar vinculados con fuentes indeseables como el ruido, preservando al mismo tiempo las características importantes de la señal. Para lograrlo es necesario identificar



**Figura 3.3.** Diagrama de descomposición en dos niveles con la técnica de descomposición de paquetes de onda (WPD).

los coeficientes en cada nodo relacionados con los componentes de ruido, y reconstruir la señal sin estos coeficientes. Este proceso se llama enfoque de eliminación de ruido, y supone que el ruido se puede encontrar dentro de ciertas escalas de tiempo-frecuencia descartando los coeficientes que excedan un cierto valor límite, o en su defecto, que cada coeficiente contiene ruido y se distribuye en todas las escalas. Para determinar cuales son los coeficientes que deben ser eliminados, [Donoho \[1995\]](#) plantea el siguiente proceso de umbralización suave,

$$d_j^p(k) = \begin{cases} \text{sign}(d_j^p(k)) (|d_j^p(k)| - th_{j,p}) & \text{si } |d_j^p(k)| > th_{j,p} \\ 0 & \text{si } |d_j^p(k)| \leq th_{j,p}, \end{cases} \quad (3.16)$$

donde  $d_j^p(k)$  es el coeficiente del nivel  $j$  en el nodo  $p$  del instante  $k$ , y  $th_{j,p}$  es el umbral aplicado. En este proceso de umbralización, los coeficientes de WPD cuyos valores absolutos son inferiores al umbral se establecen primero en cero, y luego los coeficientes restantes distintos de cero se reducen.

Para la selección de  $th_{j,p}$  se han establecido cálculos matemáticos que pueden proporcionar un umbral de ruido representativo. Los principales métodos de umbrales son:

- **Umbral Universal (*Sqtwolog*):**

$$th_{j,p} = \sqrt{2 \log_e(N_{j,p})}, \quad (3.17)$$

donde  $th_{j,p}$  es el umbral universal y  $N_{j,p}$  es la longitud del vector de coeficientes  $\mathbf{d}_j^p$ .

- **Umbral por Estimación de Riesgo No Sesgado de Stein (*Stein's Unbiased Risk***

*Estimate, SURE*): En este método, se toma el valor absoluto de cada elemento de  $\mathbf{d}_j^p$  y luego se ordenan de menor a mayor valor. Posteriormente, cada elemento se eleva al cuadrado para obtener una nueva secuencia:

$$Q_{j,p}(k) = (\text{sort}(|\mathbf{d}_j^p|))^2, \quad k = 0, 1, 2, \dots, N_{j,p} - 1 \quad (3.18)$$

Luego se calcula el riesgo  $R_{j,p}$  generado por cada elemento de  $\mathbf{Q}_{j,p}$ ,

$$R_{j,p}(k) = \left[ N_{j,p} - 2k + \frac{\sum_{i=1}^k Q_{j,p}(i) + (N_{j,p} - k) Q_{j,p}(N - k)}{N_{j,p}} \right] \quad (3.19)$$

El valor  $k$  donde se encuentra el elemento con menor riesgo en  $\mathbf{R}_{j,p}$  se denomina  $k_{min}$ . De esta manera, el umbral SURE se define como:

$$th_{j,p} = \sqrt{Q_{j,p}(k_{min})} \quad (3.20)$$

- **Umbral Heurístico con SURE (*Heursure*)**: Es una combinación de los métodos Sqrtwolog y SURE. Esto debido a que la estimación del método SURE es pobre cuando la relación señal a ruido en la señal es pequeña, mientras que en estos casos el método Sqrtwolog proporciona un mejor umbral, y viceversa. Por esta razón, se establece un criterio de selección entre estas dos técnicas determinado por,

$$c_{j,p} = \sqrt{\frac{1}{N_{j,p}} \left( \frac{\ln N_{j,p}}{\ln 2} \right)}, \quad (3.21)$$

$$e_{j,p} = \frac{\sum_{i=1}^{N_{j,p}} |d_j^p(i)|^2 - N_{j,p}}{N_{j,p}}, \quad (3.22)$$

donde el criterio de selección del umbral se establece de la siguiente manera,

$$th_{j,p} = \begin{cases} \sqrt{2 \log_e(N_{j,p})} & \text{si } e_{j,p} < c_{j,p} \\ \min(\sqrt{2 \log_e(N_{j,p})}, \sqrt{Q_{j,p}(k_{min})}) & \text{si } e_{j,p} \geq c_{j,p}, \end{cases} \quad (3.23)$$

es decir, si  $e_{j,p} < c_{j,p}$  se toma como umbral el resultado del método sqrtwolog. De lo contrario, se selecciona el menor umbral obtenido entre el método sqrtwolog y el método SURE.

Finalmente, es posible obtener una señal reconstruida  $Z_{j,p}$  de longitud  $N$  para cada nodo  $\mathbf{d}_j^p$ , asignándole un valor de cero a todos coeficientes de los demás nodos del mismo nivel

$j$  y realizando el proceso inverso de reconstrucción sobre sus coeficientes, ya modificados por la umbralización. El ancho de banda en frecuencia correspondiente a cada una de las señales  $Z_{j,p}$  reconstruidas es igual a  $f_j = fs/2^{j+1}$ , donde  $fs$  corresponde a la frecuencia de muestreo de la señal original. De esta manera, la frecuencia mínima de la sub-banda  $Z_{j,p}$  está definida por  $f_{pmin} = p * f_j$  y su frecuencia máxima por  $f_{pmax} = (p + 1) * f_j$ .

### 3.3. Filtro Espacial

#### 3.3.1. Patrones Espaciales Comunes

Patrones Espaciales Comunes (*Common Spatial Pattern*, CSP) es una técnica de procesamiento de señales que se utiliza en el análisis de electroencefalografía para extraer características discriminantes entre dos clases de datos. Este método extrae patrones espaciales de la señal basado en la covarianza entre las variaciones de potencial de los electrodos en una señal EEG multicanal [Koles et al., 1990].

La técnica de CSP se basa en la idea de que la variabilidad en la señal registrada por diferentes canales EEG está relacionada con la actividad cerebral o muscular de interés. El CSP busca encontrar una transformación lineal que maximice la varianza de la señal en el conjunto de datos perteneciente a una clase, mientras minimiza la varianza en otro conjunto de datos perteneciente a una segunda clase. Esta transformación lineal se puede utilizar para extraer características de la señal que son relevantes para etiquetar los diferentes tipos de imaginación motora.

Cada muestra  $i$  de la señal EEG está representada por  $\mathbf{X}_i \in \mathbb{R}^{N \times c}$ , donde  $c$  representa la cantidad de canales y  $N$  la cantidad de datos en tiempo. La estimación de la matriz de covarianza entre canales de la señal  $\mathbf{X}_i$  puede ser calculada como,

$$\Sigma_i = \frac{1}{N}(\mathbf{X}_i - \mu_i)(\mathbf{X}_i - \mu_i)^T, \quad (3.24)$$

donde  $\Sigma_i \in \mathbb{R}^{c \times c}$  es la matriz de covarianza de  $\mathbf{X}_i$  y  $\mu_i$  es el vector de medias de cada canal de la señal. Obteniendo una matriz  $\Sigma_i$  para cada muestra, se procede a determinar una matriz de covarianza general por cada clase  $\Sigma_p$  a través de un promedio Euclidiano, de la forma:

$$\boldsymbol{\Sigma}_p = \frac{1}{N_p} \sum_{i=1}^{N_p} \boldsymbol{\Sigma}_i, \quad (3.25)$$

donde  $N_p$  es el número de muestras de la clase  $p$ .

La técnica de CSP propuesta por Fukunaga [1990] y Koles et al. [1990] consiste en calcular una matriz  $\mathbf{W} \in \mathbb{R}^{c \times c}$  tal que,

$$\mathbf{W} \boldsymbol{\Sigma}_{C1} \mathbf{W}^T = \mathbf{D}, \quad (3.26)$$

$$\mathbf{W} \boldsymbol{\Sigma}_{C2} \mathbf{W}^T = \mathbf{I} - \mathbf{D}, \quad (3.27)$$

donde  $\boldsymbol{\Sigma}_{C1}$  y  $\boldsymbol{\Sigma}_{C2}$  son las matrices de covarianza promedio de la clase 1 y clase 2 respectivamente,  $\mathbf{D}$  es una matriz diagonal e  $\mathbf{I}$  es una matriz identidad.

Para ello, es necesario descomponer la covarianza compuesta  $\boldsymbol{\Sigma}_T$  (suma de las covarianzas de ambas clases) en sus eigenvalores  $\lambda_T$  y eigenvectores  $\mathbf{B}_T$  utilizando el método de eigenanálisis. Es decir, tenemos que:

$$\boldsymbol{\Sigma}_T = \boldsymbol{\Sigma}_{C1} + \boldsymbol{\Sigma}_{C2} = \mathbf{B}_T \lambda_T \mathbf{B}_T^T, \quad (3.28)$$

donde  $\mathbf{B}_T$  está sujeto a  $\mathbf{B}_T \mathbf{B}_T^T = \mathbf{I}$ .

Posteriormente, se calcula la matriz de transformación de blanqueo  $\mathbf{P}$  (Whitening) mediante:

$$\mathbf{P} = \lambda_T^{-1/2} \mathbf{B}_T^T, \quad (3.29)$$

y se utiliza para transformar las covarianzas promedio de cada clase de la siguiente manera:

$$\begin{aligned} \mathbf{S}_{C1} &= \mathbf{P} \boldsymbol{\Sigma}_{C1} \mathbf{P}^T, \\ \mathbf{S}_{C2} &= \mathbf{P} \boldsymbol{\Sigma}_{C2} \mathbf{P}^T. \end{aligned} \quad (3.30)$$

Por las restricciones del proceso de diagonalización de las matrices de covarianza,  $\mathbf{S}_{C1}$  y  $\mathbf{S}_{C2}$  comparten los mismos eigenvectores  $\mathbf{U}$ , pero en cuanto a sus eigenvalores se cumple que  $\lambda_{C1} + \lambda_{C2} = \mathbf{I}$ . Las columnas de  $\mathbf{U}$  con los mayores valores en  $\lambda_{C1}$  corresponden a los menores valores en  $\lambda_{C2}$ , y viceversa. Por esta razón, los eigenvectores correspondientes a los mayores eigenvalores de  $\mathbf{S}_{C1}$  y  $\mathbf{S}_{C2}$  son los más óptimos para separar la varianza entre las clases [Wang

et al., 2005] de la forma:

$$\begin{aligned}\mathbf{S}_{C_1} &= \mathbf{U} \lambda_{C_1} \mathbf{U}^T \\ \mathbf{S}_{C_2} &= \mathbf{U} \lambda_{C_2} \mathbf{U}^T\end{aligned}\tag{3.31}$$

Finalmente, la matriz de transformación  $\mathbf{W}$  está definida por la siguiente ecuación,

$$\mathbf{W} = \mathbf{U}^T \mathbf{P},\tag{3.32}$$

con la cual es posible convertir los datos originales y proyectarlos hacia un nuevo dominio definido por,

$$\mathbf{Z} = \mathbf{W} \mathbf{X},\tag{3.33}$$

donde se maximice la varianza de la clase 1 y se minimice la varianza de la clase 2. En esta última expresión,  $\mathbf{Z} \in \mathbb{R}^{N \times c}$  corresponde a la señal de EEG proyectada con  $c$  canales y  $N$  datos en tiempo.

De igual forma, según Blankertz et al. [2008], el proceso de diagonalización propuesto por Koles et al. [1990] es equivalente a la solución de la siguiente ecuación,

$$\Sigma_{C_1} \mathbf{W} = \lambda (\Sigma_{C_1} + \Sigma_{C_2}) \mathbf{W},\tag{3.34}$$

donde  $\mathbf{W}$  es la matriz de transformación del CSP y  $\lambda$  es una matriz diagonal de eigenvalores, los cuales, se interpretan como el índice de relación de varianza entre los canales de la señal de ambas clases. Por lo tanto, las columnas de la matriz  $\mathbf{W}$  corresponden a los eigenvectores asociados a los eigenvalores de mayor magnitud. Es decir, son los filtros espaciales más relevantes para separar las clases en función de la varianza de la señal.

### 3.3.2. Diagonalización Aproximada Conjunta

El método CSP descrito en la sección anterior es una técnica determinista de diagonalización que sólo permite procesar las matrices de covarianza de dos clases. Por otro lado, la Diagonalización Aproximada Conjunta (*Joint Approximate Diagonalization*, JAD) es una técnica no determinista que permite encontrar una matriz  $\mathbf{W}$  que diagonalice lo mejor posible un conjunto de matrices simétricas definidas positivas. Su aplicación surge como una solución para lograr calcular filtros espaciales con el mismo principio del CSP pero en los casos donde se procesan señales de más de dos clases.

Considerando un conjunto de  $M$  matrices de covarianza  $\Sigma \triangleq (\Sigma_1, \Sigma_2, \dots, \Sigma_M)$ , cada una de tamaño  $c \times c$ . Su diagonalización aproximada conjunta consiste en encontrar una matriz  $\mathbf{W} \in \mathbb{R}^{c \times c}$  tal que el conjunto de matrices  $\mathbf{W} \Sigma \mathbf{W}^T \triangleq (\mathbf{W} \Sigma_1 \mathbf{W}^T, \mathbf{W} \Sigma_2 \mathbf{W}^T, \dots, \mathbf{W} \Sigma_M \mathbf{W}^T)$  genere un conjunto de matrices tan diagonales como sea posible, siguiendo un criterio de diagonalización [Ablin et al., 2019].

El criterio utilizado para la diagonalización aproximada de las covarianzas en esta investigación fue establecido por Pham [2001], el cual está definido por:

$$\mathcal{L}(\mathbf{W}) = \frac{1}{2M} \sum_{i=1}^M [\log \det \text{diag}(\mathbf{W} \Sigma_i \mathbf{W}^T) - \log \det(\mathbf{W} \Sigma_i \mathbf{W}^T)], \quad (3.35)$$

donde propone un algoritmo con enfoque de descenso de coordenadas por bloque para la minimización de  $\mathcal{L}(\mathbf{W})$ . Para cada iteración del algoritmo de optimización, la función objetivo  $\mathcal{L}(\mathbf{W})$  es minimizada secuencialmente modificando los pesos (coordenadas) correspondientes a un par de columnas de la matriz  $\mathbf{W}$ , mientras que las demás columnas de  $\mathbf{W}$  se mantienen fijas. Este proceso se repite para cada combinación de columnas posible hasta que el algoritmo logra converger cuadráticamente encontrando una matriz  $\mathbf{W}$  que diagonaliza, con la mayor aproximación posible, todas las matrices de covarianza.

### 3.4. Técnicas Selección de Características

Con la finalidad de seleccionar la información más discriminante para la posterior clasificación de las señales EEG, se plantearon dos metodologías diferentes. La primera, basada en la selección de los filtros espaciales (columnas de las matrices  $\mathbf{W}$ ) que ofreciera la mejor separación entre las clases, y la segunda, basada en la selección de las mejores características luego de aplicar las técnicas de pre-procesamiento.

A continuación se describe cada una de las técnicas de selección utilizadas en ambas metodologías.

#### 3.4.1. Selección de filtros espaciales

Como se describió en secciones anteriores, la implementación de la técnica de filtrado espacial CSP está asociada con la diagonalización de las matrices de covarianza de cada clase

a través de una matriz de transformación  $\mathbf{W}$ . Cada columna de la matriz  $\mathbf{W}$  corresponde con un filtro espacial, y poder determinar cuales son los filtros que otorgan la mayor separación posible entre las varianzas de las clases, es un factor crucial para su posterior clasificación.

En los casos donde se intenta discriminar entre sólo dos clases, el cálculo de la matriz  $\mathbf{W}$  es determinista, y existe una metodología canónica para determinar los filtros más relevantes verificando el valor de los eigenvalores de la matriz diagonal resultante (3.34). En este caso, los filtros en  $\mathbf{W}$  con mayor relevancia son los eigenvectores asociados con los eigenvalores en  $\lambda$  de mayor magnitud, de esta manera, ordenando de mayor a menor valor los eigenvalores es posible reordenar la matriz  $\mathbf{W}$  desde los filtros más relevantes hasta los menos relevantes.

En el caso de trabajar con señales de múltiples clases, la construcción de  $W$  se basa en la implementación de algoritmos no deterministas, y por lo tanto, no es posible establecer un procedimiento canónico que permita determinar los filtros con mayor relevancia [Dornhege et al., 2004]. Como solución, Grosse-Wentrup and Buss [2008] propone una técnica de aproximación de información mutua utilizada como un método de extracción de características mediante la teoría de la información (*Information Theoretic Feature Extraction*, ITFE).

La información mutua es una medida de la influencia que una variable aleatoria tiene sobre otra variable aleatoria. En el contexto de la selección de características, es un método supervisado que se utiliza para medir la cantidad de información que una característica proporciona sobre la variable objetivo (una mayor información mutua representa una mayor relevancia de la característica sobre la clasificación). En su investigación, Grosse-Wentrup and Buss [2008] lograron demostrar que la Diagonalización Aproximada Conjunta de las matrices de covarianza de las señales EEG condicionadas a las etiquetas de clase es una implementación de ICA, que logra separar la señal limpia del ruido. Por lo tanto, la información mutua encuentra los filtros en  $W$  que separan correctamente el ruido de la información relevante, proporcionando mayor información sobre la intención del usuario de la BCI. Posteriormente, en el entrenamiento de los modelos planteados, esta selección permitirá ignorar aquellos datos de menor relevancia que puedan estar asociados con el ruido aditivo en la señal original.

El criterio para determinar la información mutua aplicada sobre los filtros  $\mathbf{w}_j$  correspondientes a las columnas de la matriz  $\mathbf{W}$  esta descrito por,

$$I(c, \hat{\mathbf{w}}_j^T \mathbf{x}) = - \sum_{i=1}^M P(c_i) \log \sqrt{\hat{\mathbf{w}}_j^T \Sigma_{c_i} \hat{\mathbf{w}}_j} - \frac{3}{16} \left( \sum_{i=1}^M P(c_i) ((\hat{\mathbf{w}}_j^T \Sigma_{c_i} \hat{\mathbf{w}}_j)^2 - 1) \right)^2, \quad (3.36)$$

donde  $I(c, \hat{\mathbf{w}}_j^T \mathbf{x})$  representa la información mutua correspondiente al filtro normalizado  $\hat{\mathbf{w}}_j$ ,  $c$

representa la clase,  $P(c_i)$  representa la probabilidad a priori de cada clase calculada en base a la cantidad de datos que cada una aporta al total de muestras de entrenamiento del modelo (en clases balanceadas la probabilidad es la misma para todas las clases) y  $M$  es la cantidad de clases presentes en los datos. Para poder implementar correctamente este criterio, los filtros  $\mathbf{w}_j$  deben estar normalizados utilizando la siguiente expresión,

$$\hat{\mathbf{w}}_j = \mathbf{w}_j / \sqrt{\mathbf{w}_j^T \boldsymbol{\Sigma}_{pr} \mathbf{w}_j}, \quad (3.37)$$

donde  $\boldsymbol{\Sigma}_{pr}$  es el promedio de todas las matrices de covarianza, tal que  $\hat{\mathbf{w}}_j^T \boldsymbol{\Sigma}_{pr} \hat{\mathbf{w}}_j = 1$ .

### 3.4.2. Selección por Valor Laplaciano

El Valor Laplaciano (*Laplacian Score*, LS) es un método no supervisado que puede ser utilizado para determinar la medida de importancia de cada característica en un conjunto de datos. La idea es aplicar el operador laplaciano a cada una de las características del conjunto de datos y calcular su valor de relevancia. Por lo tanto, a diferencia de la técnica ITFE, que es aplicada como un proceso previo al cálculo de las características, esta técnica debe ser aplicada posterior al cálculo de las mismas.

El método propuesto por [He et al. \[2005\]](#) se basa en la proximidad que suelen tener los datos relacionados con la misma clase. La importancia de una característica se evalúa por la capacidad que tenga de preservar su localidad. Para determinar las localidades presentes en los datos, se modela una estructura geométrica local, representada con un grafo de vecinos más próximos. El LS asigna un valor laplaciano menor a aquellas características que respeten esta estructura gráfica, y por lo tanto, sean más representativas para la clasificación.

Dada la matriz de características  $\mathbf{f} \in \mathbb{R}^{m \times r}$  con  $r$  características por cada una de las  $m$  muestras, el cálculo del valor laplaciano comienza determinando los vecinos más cercanos a cada dato utilizando como métrica de la distancia la Semejanza de Coseno representada por,

$$d(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}\mathbf{y}^T}{\|\mathbf{x}\|\|\mathbf{y}\|}, \quad (3.38)$$

donde  $\mathbf{x}$  y  $\mathbf{y}$  corresponden a los vectores de la característica  $r$  que se están comparando para determinar su cercanía, y  $\|\cdot\|$ , representa la norma 2. Este proceso se realiza para cada dato comparándose con los restantes.

Los pares de datos que estén cercanos (posean menos de  $k$  vecinos cercanos entre ellos) se establecen como datos conectados, y el restante como datos desconectados. De esta manera, se construye una matriz de conexiones  $\mathbf{G} \in \mathbb{R}^{m \times m}$  denominada grafo de vecinos cercanos, donde los pares conectados se establecen con el valor de 1 y los pares desconectados con el valor de 0.

En ciertas aplicaciones, es deseable aplicar un kernel de transformación a la matriz  $\mathbf{G}$  para determinar valores intermedios de similitud entre los pares de datos. Para ello, se calcula cada elemento  $S_{ij}$  de la matriz de similaridad mediante:

$$S_{ij} = \begin{cases} \exp \left\{ \frac{\|x_i - x_j\|^2}{\sigma^2} \right\} & \text{si } G_{ij} = 1 \\ 0 & \text{si } G_{ij} = 0, \end{cases} \quad (3.39)$$

donde  $\|x_i - x_j\|$  es la distancia euclidiana entre los datos vecinos y  $\sigma$  es un factor de escala. Para la selección de características planteada en esta investigación, se utilizó únicamente la matriz  $\mathbf{G}$  con valores binarios, pero de ser necesario puede ser sustituida por la matriz  $S_{ij}$  siguiendo la misma metodología.

Posteriormente, se define el vector de características  $\mathbf{f}_r = [f_{r_1}, f_{r_2}, \dots, f_{r_m}]^T$ , el vector unitario  $\mathbf{I} = [1, 1, \dots, 1]^T \in \mathbb{R}^m$  y la matriz diagonal  $\mathbf{D}$  calculada de la siguiente manera,

$$D(i, i) = \sum_{j=1}^m S_{i,j}, \quad (3.40)$$

y se procede a centrar los datos de  $\mathbf{f}_r$  removiendo su media utilizando la siguiente ecuación,

$$\hat{\mathbf{f}}_r = \mathbf{f}_r - \frac{\mathbf{f}_r^T \mathbf{D} \mathbf{I}}{\mathbf{I}^T \mathbf{D} \mathbf{I}} \mathbf{I}. \quad (3.41)$$

Finalmente el valor laplaciano de la característica  $r$  se calcula como,

$$L_r = 1 - \frac{\hat{\mathbf{f}}_r^T \mathbf{G} \hat{\mathbf{f}}_r}{\hat{\mathbf{f}}_r^T \mathbf{D} \hat{\mathbf{f}}_r}. \quad (3.42)$$

De acuerdo con He et al. [2005], la ecuación (3.42) surge como resultado de la minimización de la siguiente función,

$$L_r = \text{mín} \left( \frac{\sum_{ij} (f_{ri} - f_{rj}) G_{ij}}{\text{var}(\mathbf{f}_r)} \right), \quad (3.43)$$

donde  $\text{var}(\mathbf{f}_r)$  es la varianza del vector de características  $\mathbf{f}_r$ . Lo que se busca con esta función

objetivo es conseguir las características que posean menor varianza, y una mejor disposición a respetar las estructuras locales definidas en el grafo de vecinos cercanos (menores valores de  $(f_{ri} - f_{rj})$  y menor cantidad de conexiones en  $\mathbf{G}$ ).

Para determinar la relevancia del conjunto de características, se calcula  $L_r$  para cada una de ellas y se ordena desde el menor hasta el mayor valor, correspondiendo el menor valor de  $L_r$  a la característica con mayor relevancia, y viceversa.

### 3.5. Métodos de Clasificación

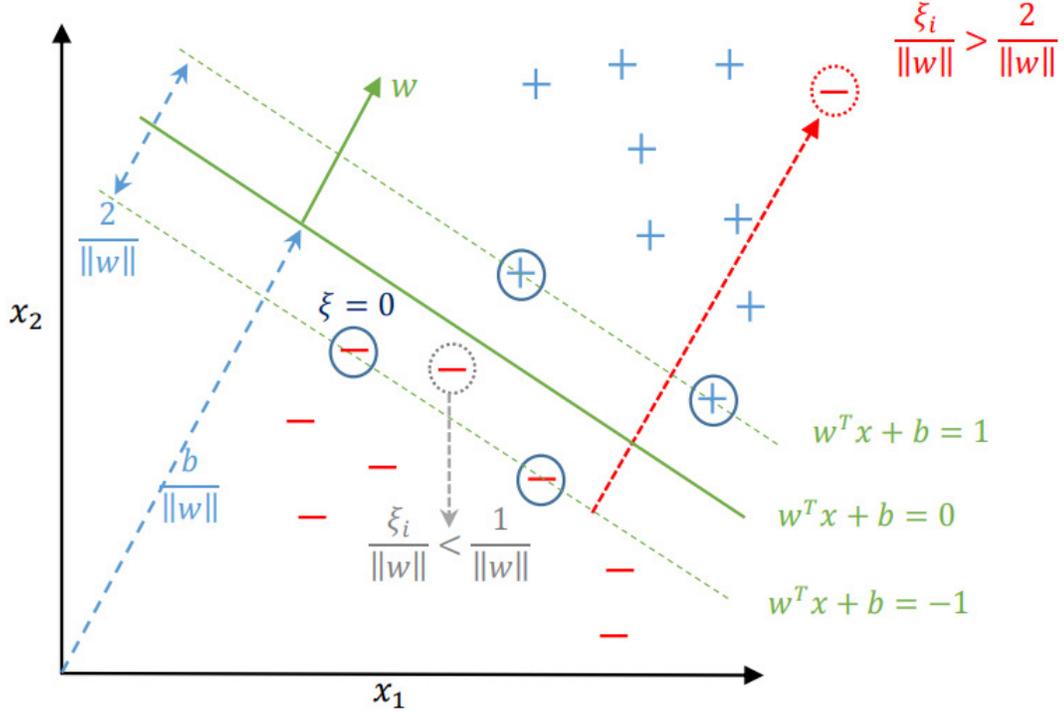
Luego de procesar las señales EEG para obtener características diferenciadoras que describan los patrones más relevantes de la señal, es necesario implementar las herramientas adecuadas para crear las fronteras de decisión que logren utilizar estas características con el fin de identificar de la mejor forma posible cada una de las señales de entrada, y su correspondiente tipo de imaginación motora.

A continuación se explicarán los métodos de clasificación basados en aprendizaje automático, considerados como clásicos, que se utilizaron para separar y clasificar las señales EEG en sus diferentes categorías de imaginación motora.

#### 3.5.1. Maquinas de Vectores de Soporte

Las Máquinas de Vectores Soporte (Support Vector Machines, SVM) es un algoritmo de aprendizaje supervisado compuesto por un conjunto de métodos aplicables tanto a problemas de clasificación como de regresión [Shmilovici, 2005]. Como se observa en la Figura 3.4, el clasificador SVM busca el hiperplano de separación óptimo ( $\mathbf{w}^T \mathbf{x} + \mathbf{b} = 0$ ) entre dos clases, maximizando el margen ( $2/\|\mathbf{w}\|$ ) entre los puntos más cercanos de las clases en un espacio de entrada transformado. Los puntos situados en los límites, denominados vectores de soporte, son los que determinan la posición del hiperplano que genera la separación óptima en relación con las dos clases.

El proceso de entrenamiento de una función de decisión SVM consiste en identificar los parámetros del hiperplano derivados de un problema de optimización cuadrática, que minimicen el valor de  $\|\mathbf{w}\|$ , y a su vez, maximicen la distancia entre los vectores de soporte de



**Figura 3.4.** Interpretación geométrica del SVM con margen regularizado (*soft margin*).

ambas etiquetas de clase  $\Delta$  definido como,

$$\Delta = \frac{2}{\|\mathbf{w}\|}. \quad (3.44)$$

Sea  $m$  el número de instancias de entrenamiento de un conjunto de datos  $\mathbf{X}$ , donde cada entrada  $\mathbf{x}_i \in \mathbf{X}$  contiene  $r$  características y dos clases definidas por  $y_i = \{-1, 1\}$ ; el conjunto de datos se puede representar por,

$$\{\mathbf{x}_i, y_i\} \quad \forall i = 1, \dots, m, y_i \in \{-1, 1\}, \mathbf{x} \in \mathbb{R}^r. \quad (3.45)$$

De acuerdo con [Shawe-Taylor and Sun \[2014\]](#), cuando los datos en  $\mathbf{X}$  no son linealmente separables, entonces el procedimiento para calcular el hiperplano que separa de mejor manera los datos de cada clase se basa en la solución del siguiente problema de minimización cuadrática,

$$\min_{\mathbf{w}, b, \xi} \left( \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \right), \quad (3.46)$$

con la siguiente restricción,

$$y_i (\mathbf{w}^T K(\mathbf{x}_i) + \mathbf{b}) \geq 1 - \xi_i, i = 1, \dots, m, \quad (3.47)$$

donde  $\mathbf{w}$  y  $\mathbf{b}$  son los parámetros del hiperplano,  $\xi$  es el parámetro de holgura de los datos que no pueden cumplir con el margen de separación,  $C$  es el parámetro de compensación o penalización entre  $\xi$  y el tamaño del margen, y  $K(\cdot)$  es el kernel de transformación aplicado a los datos no lineales para poder llevarlos a un plano donde se puede establecer una frontera lineal entre ellos.

**Definición 3.5.1 (Kernel).** Es una función matemática que permite transformar los datos de entrada  $\mathbf{x}_1$  y  $\mathbf{x}_2$  desde un espacio de características de menor dimensión hacia un espacio de mayor dimensión, y está definida como,

$$\langle \mathbf{x}_1 \cdot \mathbf{x}_2 \rangle \leftarrow K(\mathbf{x}_1, \mathbf{x}_2) = \langle \Phi(\mathbf{x}_1) \cdot \Phi(\mathbf{x}_2) \rangle \quad (3.48)$$

donde  $\Phi(\cdot)$  es una función de transformación que convierte los datos de entrada hacia un espacio de mayor dimensión definida como,

$$\Phi : \mathbf{X} \in \mathbb{R}^r \rightarrow \mathbf{Y} \in \mathbb{R}^q, \text{ con } q > r. \quad (3.49)$$

La idea detrás del uso de un kernel es que, aunque los datos originales puedan no ser linealmente separables, es posible que exista una transformación no lineal de los mismos a un espacio de mayor dimensión donde sí lo sean. La operación descrita en (3.48) puede ser sustituida por cualquiera de las funciones kernel definidas a continuación:

- **Kernel Lineal:**

$$K(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T \mathbf{x}_2, \quad (3.50)$$

- **Kernel Polinomial:**

$$K(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1^T \mathbf{x}_2 + 1)^p, \quad (3.51)$$

donde  $p$  es el orden de la función polinomial.

- **Kernel Función de Base Radial (RBF) :**

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|}{2\sigma^2}\right), \quad (3.52)$$

donde  $\sigma$  es el factor de anchura del kernel.

- **Kernel Sigmoide:**

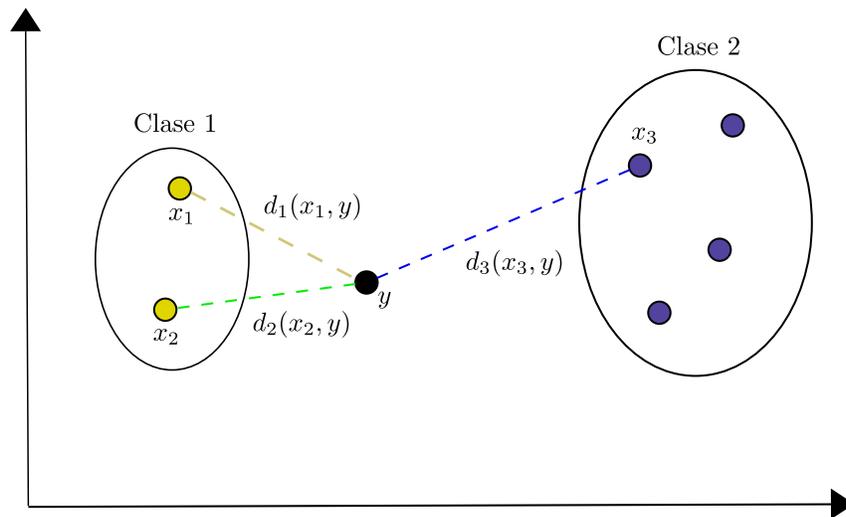
$$K(\mathbf{x}_1, \mathbf{x}_2) = \tanh(\gamma(\mathbf{x}_1^T \mathbf{x}_2) + r), \quad (3.53)$$

donde  $\gamma$  es el factor de escala y  $r$  es el factor de desplazamiento.

### 3.5.2. K-Vecinos Más Cercanos

La técnica de K vecinos más cercanos (K-Nearest Neighbor, KNN) es una técnica de aprendizaje automático supervisado que se utiliza para clasificar objetos basados en las características de los vecinos más cercanos en un conjunto de datos. Se basa en el principio de que las instancias de un conjunto de datos suelen encontrarse próxima a otros casos con propiedades similares [Cover and Hart, 1967].

Como se demuestra en la Figura 3.5, si las instancias de un conjunto de datos están etiquetadas, entonces el valor de la etiqueta de una instancia no clasificada puede determinarse observando la clase de sus vecinos más cercanos. El KNN localiza las  $k$  instancias más cercanas a la instancia de consulta y determina su clase identificando la etiqueta más frecuente [Kotsiantis, 2007].



**Figura 3.5.** Interpretación del modelo KNN identificando la clase de una instancia no clasificada  $y$  con  $k = 3$ .

En general, las instancias  $\mathbf{x}_i$  pueden considerarse muestras o puntos dentro de un espacio de datos  $\mathbf{X} \in \mathbb{R}^{r \times m}$ , donde  $m$  corresponde a la cantidad de instancias del conjunto y  $r$  corresponde al número de características que se utilizan para describir cada instancia. La posición absoluta de estas muestras dentro del espacio no es tan significativa como la distancia

relativa entre ellas. Esta distancia relativa se determina utilizando una métrica de distancia. Idealmente, la métrica utilizada debe minimizar la distancia entre dos muestras clasificadas de forma similar, mientras que maximiza la distancia entre muestras de clases diferentes. Las principales métricas de distancia utilizadas en clasificadores KNN son:

- **Minkowsky:**

$$d(\mathbf{x}, \mathbf{y}) = \left( \sum_{i=1}^m |x_i - y_i|^r \right)^{1/r}, \quad (3.54)$$

donde  $m$  representa el número de elementos de los vectores  $\mathbf{x}$  y  $\mathbf{y}$ , y  $r$  representa el orden de la métrica.

- **Manhattan:**

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m |x_i - y_i|, \quad (3.55)$$

- **Euclideana:**

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^m |x_i - y_i|^2}, \quad (3.56)$$

- **Chebychev:**

$$d(\mathbf{x}, \mathbf{y}) = \max(|\mathbf{x} - \mathbf{y}|), \quad (3.57)$$

- **Semejanza de Coseno:**

$$d(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}\mathbf{y}^T}{\|\mathbf{x}\|\|\mathbf{y}\|}, \quad (3.58)$$

Para obtener resultados más precisos, es posible utilizar esquemas de ponderación que alteran las medidas de distancia de los vecinos calculados con el objetivo de privilegiar a los vecinos más cercanos y disminuir la influencia de los vecinos lejanos. Por lo tanto, el peso de los vecinos debe variar inversamente con la distancia, de tal manera que los puntos más cercanos tengan mayor peso. Esto se logra aplicando un kernel de transformación definido por,

$$w = K[d(\mathbf{x}, \mathbf{y})], \quad (3.59)$$

que determina el peso  $w$  de cada vecino  $\mathbf{x}$  basado en la distancia a la instancia de referencia  $\mathbf{y}$ . Los principales kernel de ponderación utilizados están descritos en el Tabla 3.1.

**Tabla 3.1.** Kernel de ponderación para cálculo de pesos en KNN.

Kernel	Función
	$1 / d^2$
$K[d(\mathbf{x}, \mathbf{y})]$	$1 / d$
	$\exp(-d)$

### 3.5.3. Análisis Discriminante Lineal

El Análisis Discriminante Lineal (*Linear Discriminant Analysis*, LDA) es una técnica estadística de aprendizaje supervisado que se utiliza para encontrar una combinación lineal de características que permita maximizar la separación entre dos o más clases. Es decir, LDA busca encontrar una transformación lineal que maximice la relación de la varianza entre clases y la varianza dentro de cada clase.

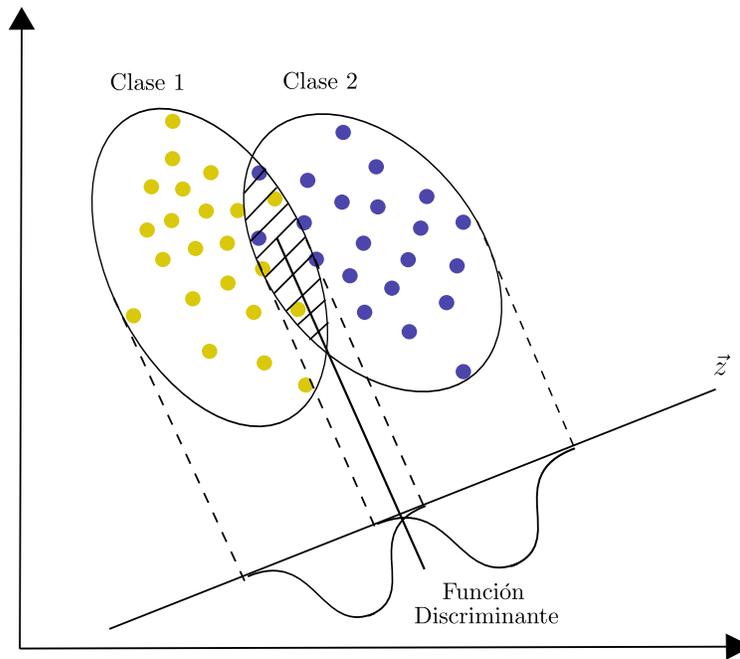
Como se describe en la Figura 3.6, el objetivo del LDA es determinar una frontera lineal que permita predecir la clase a la que pertenece una observación, dada una serie de características o variables predictoras. Para ello, LDA utiliza la información de las clases conocidas y las características de las observaciones para construir una función discriminante lineal que permita clasificar nuevas observaciones en las clases correspondientes.

El clasificador LDA supone que las observaciones en la  $k$ -ésima clase se extraen de una distribución de probabilidad gaussiana multivariable,

$$P(\mathbf{x}; \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mu_k)^T \Sigma_k^{-1} (\mathbf{x} - \mu_k)\right), \quad (3.60)$$

donde  $d$  corresponde al número de características,  $\Sigma_k$  corresponde a la matriz de covarianza de cada una de las clases  $k$ ,  $\mu_k$  es el vector de medias de cada una de las clases y  $\mathbf{x}$  corresponde al vector de características que se desea clasificar. A través de esta expresión, el modelo LDA, busca determinar la distribución condicional o grado de pertenencia del dato de entrada que se desea clasificar en función de cada clase.

Según Duda et al. [2006], con la finalidad de reducir la complejidad computacional en el cálculo de la distribución de probabilidad  $P(\mathbf{x}; \mu_k, \Sigma_k)$ , la expresión (3.60) puede ser reorganizada y linealizada aplicando el operador logaritmo. Adicionalmente, se asume que  $\Sigma_k = \Sigma$ , es decir, que la covarianza de todas las clases es igual. De esta manera, es posible



**Figura 3.6.** Interpretación del modelo LDA para la identificación de la clase de una instancia no clasificada.

determinar una superficie de decisión lineal,

$$\log P(\mathbf{x}; \mu_k, \Sigma) = \mathbf{x}^T \Sigma^{-1} \mu_k + (-1/2) \mu_k^T \Sigma^{-1} - \mu_k + \log P(k) \quad (3.61)$$

donde  $P(k)$  corresponde a la probabilidad a priori de cada clase.

Finalmente, el resultado de la clasificación asignará a la instancia de entrada  $\mathbf{x}$  la clase  $k$  que tenga el mayor  $\log P(\mathbf{x}; \mu_k, \Sigma)$ .

### 3.6. Conclusiones del capítulo

En este capítulo se presentaron las bases teóricas de las técnicas implementadas en la metodología de esta investigación. Adicionalmente, se describieron los fundamentos matemáticos que respaldan el funcionamiento de los algoritmos de pre-procesamiento de las señales EEG, extracción de atributos y posterior clasificación.

En el Capítulo 4, se expondrán los detalles de los marcos de trabajo propuestos y la metodología de implementación de los métodos descritos, orientados a la clasificación de

imaginaciones motoras para una Interfaz Cerebro-Computadora (MI-BCI).

# CAPÍTULO 4

---

## Metodología

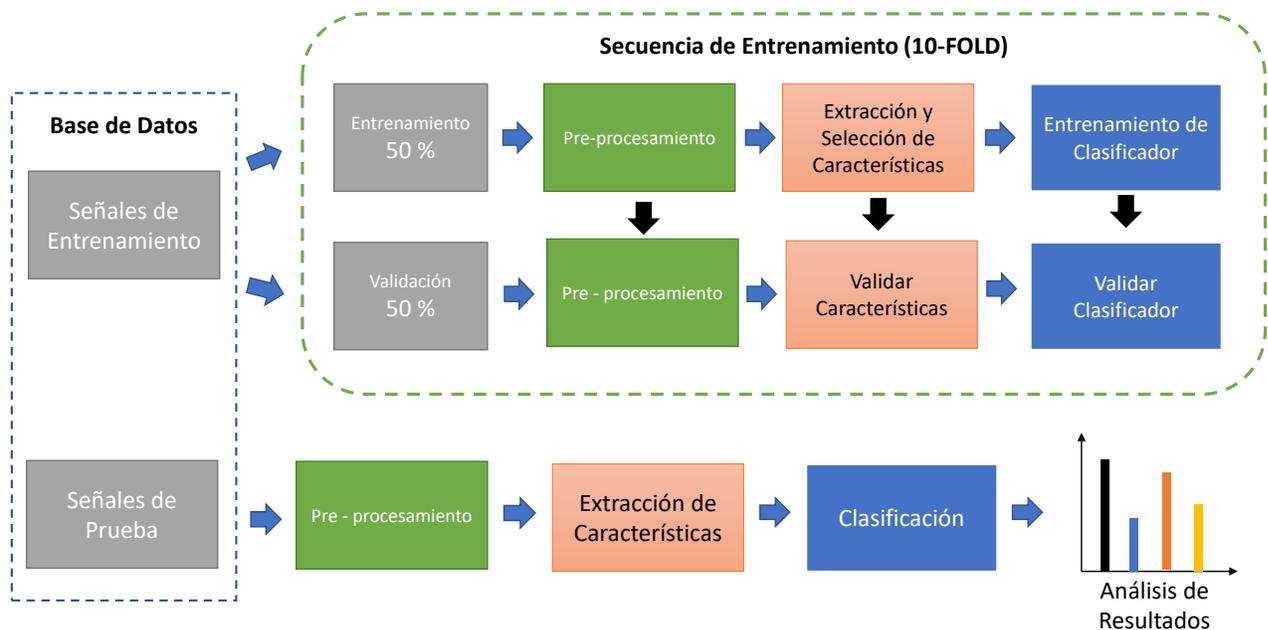
---

---

<b>4.1</b>	<b>Descripción de Bases de Datos</b>	<b>54</b>
4.1.1	BCI Competition IV Dataset 2a	54
4.1.2	BCI Competition IV Dataset 2b	55
<b>4.2</b>	<b>Etapa de Pre-procesamiento</b>	<b>57</b>
4.2.1	Filtrado y Reducción de Artefactos de Ruido	57
4.2.2	Filtrado Espacial	60
<b>4.3</b>	<b>Etapa de extracción y selección de características</b>	<b>63</b>
4.3.1	Extracción de Características	63
4.3.2	Selección de Mejores Características	64
<b>4.4</b>	<b>Clasificación</b>	<b>67</b>
4.4.1	Modelo basado en SVM	67
4.4.2	Modelo basado en LDA	68
4.4.3	Modelo basado en KNN	68
<b>4.5</b>	<b>Métricas de desempeño</b>	<b>69</b>
<b>4.6</b>	<b>Descripción de Marcos de Trabajo</b>	<b>71</b>

4.6.1 Propuesta 1: OFB-CSP-FS . . . . . 71  
 4.6.2 Propuesta 2: WPD-CSP-FS . . . . . 71

La metodología general aplicada en la investigación está compuesta por una secuencia de etapas definidas para el entrenamiento, validación y prueba de los marcos de trabajo propuestos. En la figura 4.1 se puede observar que inicialmente se contempla la adquisición de los datos de entrenamiento y pruebas (definidos por las bases de datos utilizadas), posteriormente la partición de datos de entrenamiento es sub-dividida de forma estratificada con un 50% de datos para el entrenamiento del modelo y un 50% para la validación del rendimiento del clasificador con datos desconocidos. Este esquema de partición de los datos se utilizó bajo la técnica de validación cruzada, repitiendo el experimento 10 veces, y en cada una de ellas, haciendo una selección de datos aleatoria. El objetivo principal de este proceso es evaluar diferentes técnicas de selección de características y escoger aquella con el mejor rendimiento promedio obtenido en validación, y de esta manera, obtener una mayor generalización del clasificador ante datos desconocidos.



**Figura 4.1.** Esquema general de las etapas desarrolladas en la metodología de la investigación.

Durante las pruebas de cada modelo, se empleó el esquema de entrenamiento por sujeto (within subject training), en el cual por la naturaleza única que poseen las señales del EEG,

se entrena un modelo diferente por cada sujeto de prueba utilizando los datos de cada uno de ellos de forma independiente.

Para cada marco de trabajo, se desarrolló un algoritmo exploratorio apoyado con técnicas de selección de características, con el propósito de definir las características que ayudaran al clasificador a generalizar lo mejor posible. Para ello, se propuso utilizar una selección de características optimizadas para cada sujeto de prueba tomando en cuenta el mejor rendimiento individual promedio de la métrica kappa en la clasificación de los datos de validación durante el proceso de validación cruzada.

Finalmente, una vez que los hiper-parámetros del modelo y las mejores características han sido definidas, la partición de datos de prueba es utilizada para comprobar el funcionamiento de cada uno de modelos. A continuación se describe con detalles cada uno de los marcos de trabajo propuestos en la investigación, las especificaciones de los datos utilizados durante los experimentos, las etapas desarrolladas y los algoritmos empleados.

## 4.1. Descripción de Bases de Datos

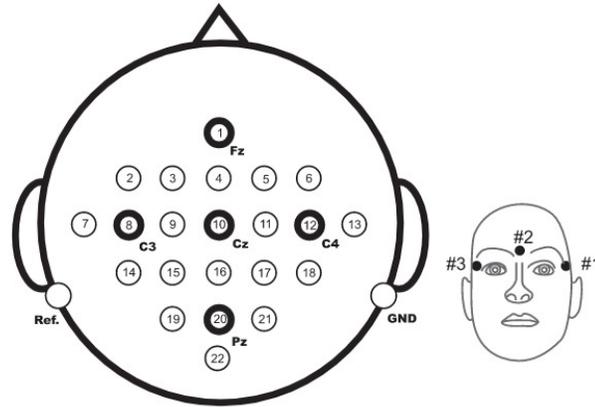
### 4.1.1. BCI Competition IV Dataset 2a

Con el objetivo de poder poner a prueba el sistema propuesto, es necesario contemplar una base de datos robusta, cuyos datos representen fielmente las dificultades reales presentes en la clasificación de intenciones motoras, obtenidas con señales de EEG, y estén respaldados por la comunidad científica.

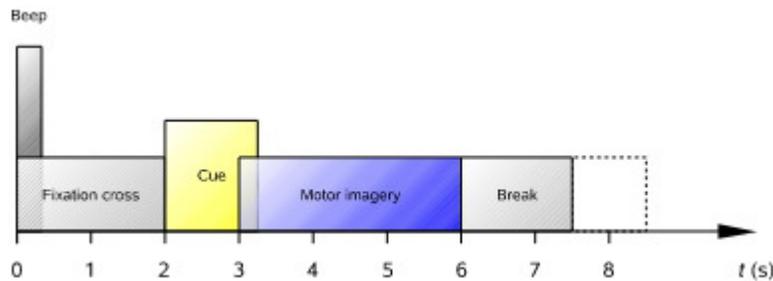
Por esta razón, se utilizó la base de datos BCI Competition IV 2a, publicada por el Laboratory of Brain-Computer Interfaces del Graz University of Technology [[Tangemann et al., 2012](#)]. Esta base de datos consta de cuatro clases basadas en intenciones motoras (pensar en el movimiento de lengua, pie, mano izquierda y mano derecha), realizadas por 9 voluntarios y adquiridas por 22 canales de EEG a una frecuencia de muestreo de 250 Hz con una duración de tres segundos por muestra.

Los datos se registraron en dos sesiones adquiridas en días diferentes. Cada sesión se compone de 6 pruebas separadas por recesos. Cada prueba consta de 48 intentos (12 para cada uno de los cuatro posibles clases), dando un total de 288 intentos por sesión. La estructura de

cada intento y la posición de los electrodos en los sujetos de prueba durante la adquisición de las señales puede verse con detalle en la Figura 4.2.



(a) Ubicación de los 22 canales adquiridos



(b) Secuencia de acciones en cada intento por sujeto

**Figura 4.2.** Detalles de la adquisición de los datos suministrados en la *BCI Competition IV dataset 2a* por [Tangemann et al. \[2012\]](#).

#### 4.1.2. BCI Competition IV Dataset 2b

Esta base de datos está compuesta por señales de EEG de 9 sujetos distintos publicados por el Laboratory of Brain-Computer Interfaces del Graz University of Technology. Los sujetos de prueba eran diestros y se les pagó por participar en los experimentos. Todos los voluntarios estaban sentados en un sillón, mirando un monitor de pantalla plana colocado a aproximadamente 1 m de distancia, a la altura de los ojos.

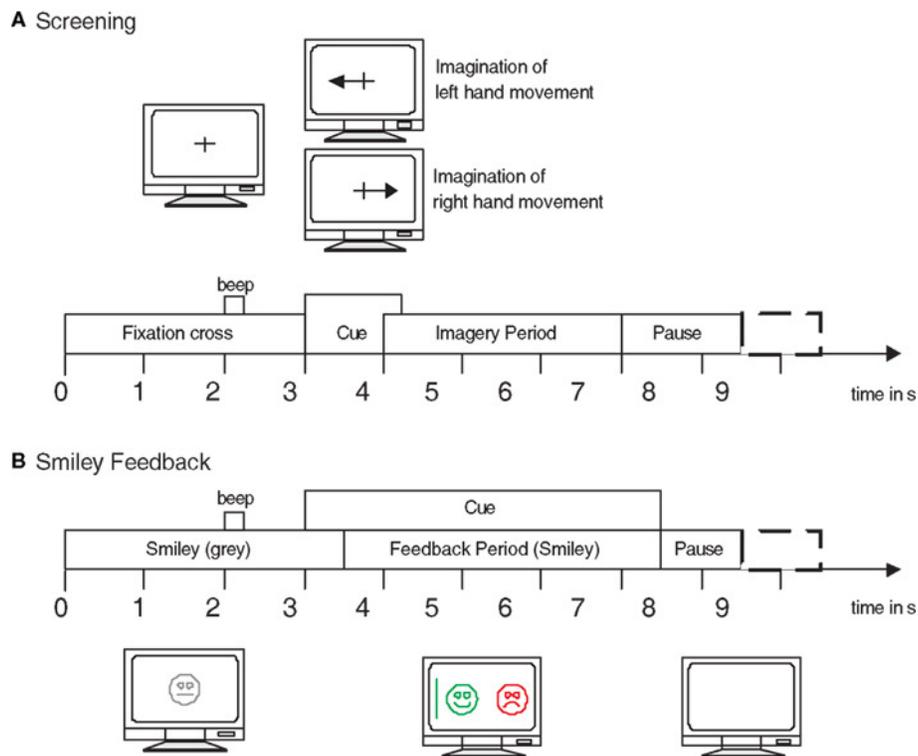
El *BCI Competition IV dataset 2b* se centra en la clasificación de señales de EEG afectadas por artefactos de movimiento ocular, tomando en cuenta la transferencia entre sesiones de los algoritmos, ya que todos los conjuntos de datos de entrenamiento y prueba

se graban en cinco días distintos [Tangermann et al., 2012].



**Figura 4.3.** Estructura de cada sesión en el *BCI Competition IV dataset 2b*.

Como se observa en la Figura 4.3, al principio de cada sesión se realizó una grabación de aproximadamente 5 minutos para estimar la influencia de los artefactos de ruido EOG. La grabación se dividió en 3 partes: (1) dos minutos con los ojos abiertos (mirando una cruz de fijación en la pantalla), (2) un minuto con los ojos cerrados y (3) un minuto con movimientos oculares. El bloque de artefactos se dividió en cuatro secciones (15 segundos de artefactos y 5 segundos de descanso entre ellos).



**Figura 4.4.** Esquema de experimento realizado para la toma de datos en la *BCI Competition IV dataset 2b* por Tangermann et al. [2012].

Cada sujeto participó en dos sesiones sin retroalimentación y tres sesiones con retroalimentación, grabadas en días diferentes en un plazo de dos semanas. Cada sesión consistió en seis ejecuciones de pensamiento motor con diez muestras de cada una. El resultado fue 20 muestras por ejecución representando 120 muestras por sesión.

Cada muestra comenzaba con una cruz de fijación y un tono acústico, como se muestra en la Figura 4.4. Unos segundos después, una señal visual (una flecha derecha o izquierda, según la clase solicitada) durante 1,25 segundos. A continuación, los sujetos debían imaginar el movimiento de la mano correspondiente durante 4 segundos. Cada muestra finalizaba con una breve pausa de al menos 1,5 segundos.

Las señales de EEG contenidas en esta base de datos se adquirieron a través de tres canales bipolares ( $C_3$ ,  $C_z$  y  $C_4$ ), para representar dos clases de pensamiento motor (mano izquierda y mano derecha). Los datos se procesaron con un filtro pasa-banda entre 0,5 y 100 Hz, y un filtro notch a 50 Hz para eliminar el ruido de la red eléctrica. La frecuencia de muestreo es de 250 Hz. La señal EOG se registró con tres electrodos monopolares de la misma manera que en la base de datos BCI Competition IV 2a.

## 4.2. Etapa de Pre-procesamiento

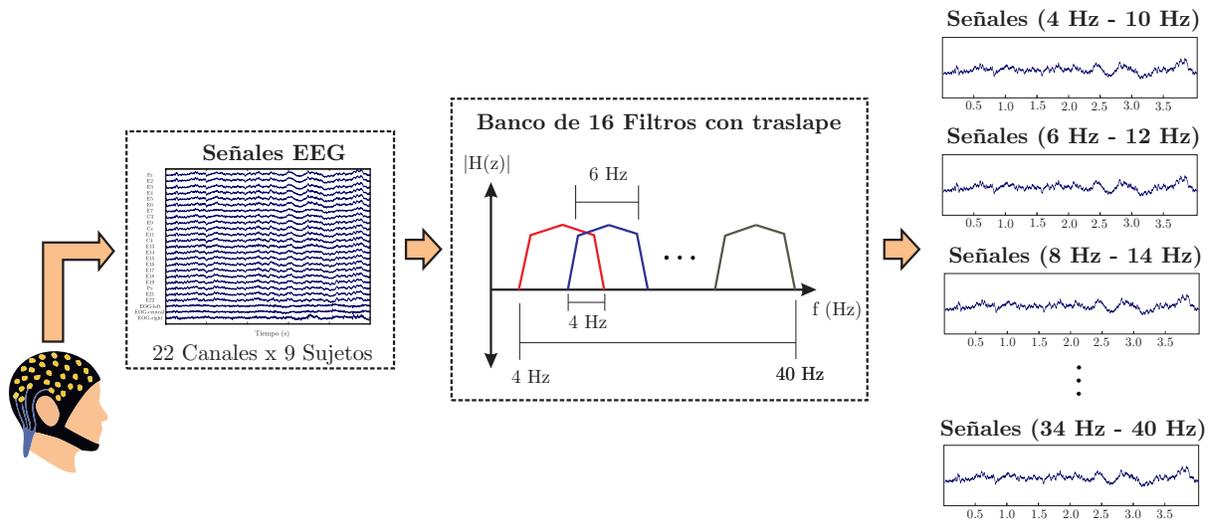
### 4.2.1. Filtrado y Reducción de Artefactos de Ruido

Obtener una correcta manipulación de la información contenida en las señales de EEG requiere filtrar adecuadamente las fuentes de ruido externo, también denominadas artefactos de ruido. Para lograr este objetivo se incorporó la primera etapa de pre-procesamiento utilizada en todos los marcos de trabajo propuestos, como se ejemplifica en la Figura 4.5, en donde se emplean herramientas de filtrado orientadas a eliminar o reducir el efecto de los artefactos de ruido contenida en las señales.

#### 4.2.1.1. Banco de filtros con Solapamiento:

Como primera propuesta de pre-procesamiento se aplicó un banco de 16 filtros no causales pasa-banda IIR Butterworth de Orden 2, debido a su respuesta plana en la banda de paso que limita la introducción de nuevos máximos locales a la señal [Smith, 1997]. El

ancho de banda de cada filtro es de 6 Hz con un traslape de 4 Hz entre cada uno. Los filtros fueron aplicados dos veces (hacia adelante en el tiempo y de regreso) con el objetivo de evitar que el filtro generara desfase en la señal. Esto asegura que los máximos locales de la señal permanezcan en sus posiciones originales independientemente de su frecuencia [Michel and Brunet, 2019].



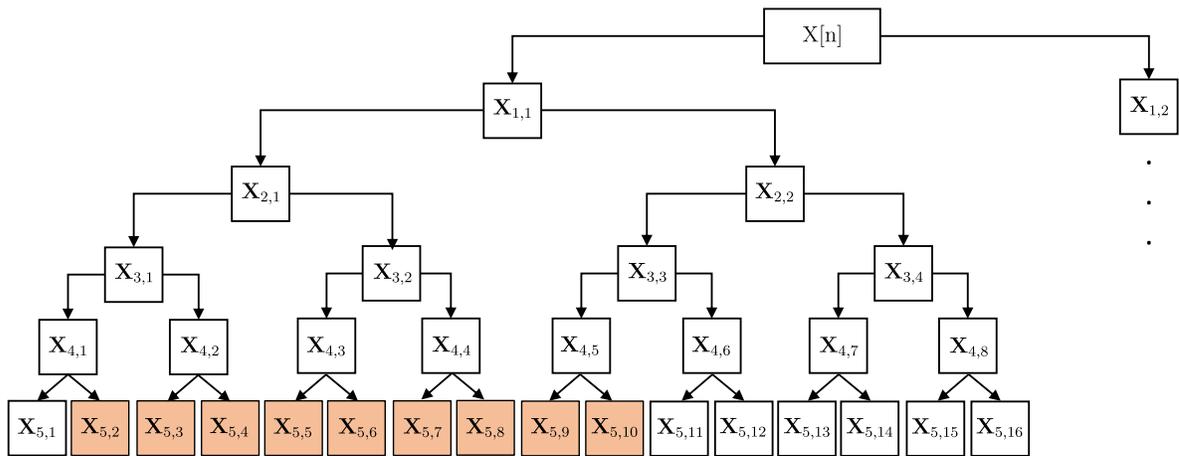
**Figura 4.5.** Aplicación del banco de filtros IIR Butterworth de Orden 2 sobre las señales adquiridas en la base de datos.

Con esta etapa de filtrado se pretende analizar las bandas de frecuencia más relevantes al pensamiento motor contenidas entre 4.0 y 40.0 Hz en cada una de las bases de datos evaluadas, obteniendo un total de 16 señales diferentes a partir de cada canal original. Posteriormente, cada una de estas sub-señales filtradas serán procesadas para obtener las mejores características posibles.

#### 4.2.1.2. Descomposición de Paquetes de Ondas:

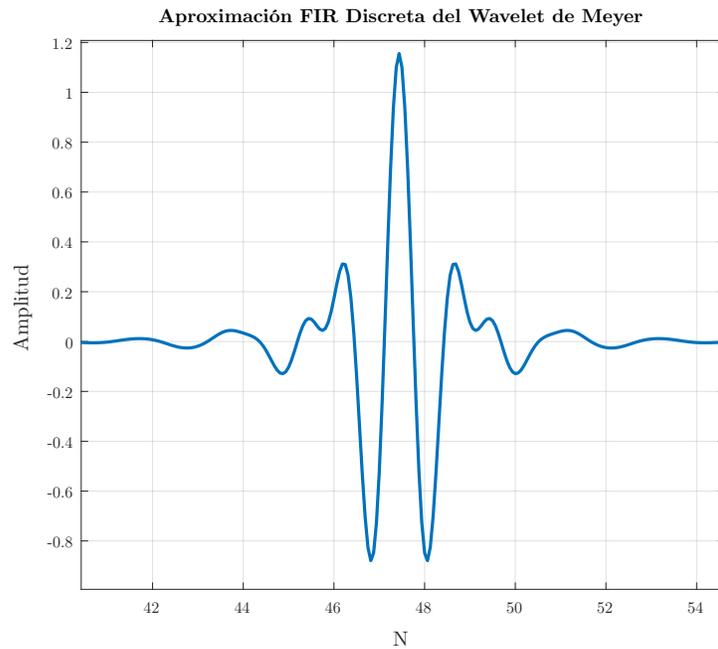
La Descomposición de Paquete de Ondas, WPD, es una técnica de filtrado que se utiliza para analizar señales o datos en diferentes escalas de resolución y frecuencia. La WPD utiliza una base de ondas o *wavelets*, y un árbol binario para descomponer la señal en sub-bandas. En cada nivel del árbol binario, la señal se divide en dos sub-bandas de frecuencia, y cada una de estas sub-bandas se divide a su vez en dos sub-bandas más pequeñas.

La WPD es útil para analizar señales que contienen información de diferentes frecuencias y escalas de tiempo como las señales EEG. En la propuesta de esta investigación, se realizó



**Figura 4.6.** Descomposición de 5 niveles con WPD sobre la señal EEG.

una descomposición de 5 niveles a cada una de las señales en las bases de datos evaluadas, con el objetivo de dividir la señal en paquetes de 3.9 Hz cada una (ver Figura 4.6). Se utilizó la función de onda de aproximación FIR Discreta de Meyer, que se muestra en la Figura 4.7, la cual es una función ortogonal que permite obtener una subdivisión equitativa de la energía de la señal entre las sub-bandas.



**Figura 4.7.** Gráfica de la aproximación FIR Discreta del Wavelet de Meyer.

Como se observa en la Figura 4.6, sólo los coeficientes de las nueve (9) sub-bandas del quinto nivel de descomposición comprendidas entre 4.0 y 40.0 Hz ( $\mathbf{X}_{5,2}, \mathbf{X}_{5,3}, \dots, \mathbf{X}_{5,10}$ ) fueron seleccionados y procesados, debido a que corresponden a las sub-bandas de frecuencia vinculadas con la imaginación motora. Posteriormente, a cada nodo del nivel cinco de la descomposición se le aplicó un proceso de selección por umbral suave (*soft thresholding*), utilizando el umbral heursure, esto con el objetivo de hacer cero los coeficientes que estuvieran vinculados con los componentes de ruido en la señal. Finalmente cada nodo fue reconstruido, utilizando los coeficientes modificados por el proceso de umbralización, para obtener una señal por cada sub-banda de frecuencia con menor presencia de artefactos de ruido, conservando la información relevante.

### 4.2.2. Filtrado Espacial

La técnica de filtrado espacial CSP es aplicada a cada una de las señales generadas por el banco de filtros o las señales descompuestas por la transformada Wavelet de forma independiente. La idea principal de CSP es calcular una matriz  $\mathbf{W} \in \mathbb{R}^{c \times c}$ , donde  $c$  es el número de canales adquiridos, que sea capaz de proyectar los datos filtrados  $\mathbf{E}$  hacia un nuevo dominio  $\mathbf{S}$  que obtenga la mayor separación entre las clases maximizando la diferencia entre sus varianzas. Esta técnica se aplica a cada una de las señales resultantes de la etapa de filtrado, por lo tanto, permite obtener una matriz de transformación  $\mathbf{W}$  diferente por cada sub-banda de frecuencia para cada sujeto de prueba.

#### 4.2.2.1. CSP para clasificación binaria:

En el caso donde los datos contemplan sólo dos clases, esta técnica busca obtener la diagonalización de las matrices de covarianza promedio de cada una de las clases. La implementación del algoritmo para calcular las matrices de transformación sobre este caso binario se basó en la propuesta presentada por Koles et al. [1990] y Blankertz et al. [2008]. Los aspectos matemáticos que fundamentan esta implementación están descritos detalladamente en el Capítulo 3, en la sección dedicada a la técnica de CSP.

El algoritmo desarrollado para la aplicación de la técnica CSP se muestra en el Algoritmo 1, el cual recibe como variables de entrada los datos de cada clase  $\mathbf{X}_{C1} \in \mathbb{R}^{m \times c \times N}$  y  $\mathbf{X}_{C2} \in \mathbb{R}^{m \times c \times N}$ , separados y previamente filtrados, donde  $m$  representa el número de muestras de cada clase,  $c$  la cantidad de canales y  $N$  la cantidad de datos por cada segmento de señal.

Cada matriz de datos es inicialmente centrada, restando su media, para proceder a calcular la Matriz de Covarianza Muestreada (*Sample Covariance Matrix*, SCM) por cada muestra de las clases. El cálculo de los patrones espaciales se aplica después de centrar los datos con respecto a su media para que la optimización de la señal se enfoque en la relación de varianza, de lo contrario, CSP optimiza la relación del momento de segundo orden.

El resultado de esta operación genera dos matrices  $\Sigma_{C1,n1} \in \mathbb{R}^{m \times c \times c}$  y  $\Sigma_{C2,n2} \in \mathbb{R}^{m \times c \times c}$ , las cuales posteriormente son promediadas con métrica euclidiana, para obtener  $\Sigma_{C1} \in \mathbb{R}^{c \times c}$  y  $\Sigma_{C2} \in \mathbb{R}^{c \times c}$ . Estas dos matrices de covarianza promedio, una por cada clase, luego son utilizadas para resolver la siguiente ecuación,

$$\Sigma_{C1} \cdot \mathbf{V} = \lambda(\Sigma_{C1} + \Sigma_{C2}) \cdot \mathbf{V}, \quad (4.1)$$

mediante una descomposición de eigenvalores.

Finalmente, se obtiene la matriz de eigenvectores  $\mathbf{V} \in \mathbb{R}^{c \times c}$  que funcionará como filtro espacial. Los filtros de la matriz  $\mathbf{V}$  (columnas) son ordenados de mayor a menor relevancia utilizando como referencia la magnitud de cada uno de los eigenvalores asociados a ellos en  $\lambda$ , los cuales se interpretan como los índices de relación de varianza entre los canales de la señal de ambas clases. Por lo tanto, los eigenvalores de mayor magnitud en  $\lambda$  poseen mayor relevancia para ser utilizados como filtros al momento de maximizar la diferencia de varianza entre clases.

---

**Algoritmo 1** Algoritmo CSP aplicado a dos clases

---

**Input** :  $\mathbf{X}_{C1}, \mathbf{X}_{C2}$

**Output:** Matriz de proyección  $\mathbf{W}$

$\Sigma_{C1,n1} \leftarrow (\mathbf{X}_{C1} - \mu_1)(\mathbf{X}_{C1} - \mu_1)^T / N$        $\triangleright$  Calc. Covarianza Clase 1 y Clase 2 por muestra

$\Sigma_{C2,n2} \leftarrow (\mathbf{X}_{C2} - \mu_2)(\mathbf{X}_{C2} - \mu_2)^T / N$

$\Sigma_{C1} \leftarrow \text{avg}(\Sigma_{C1,n1})$        $\triangleright$  Promediar todas las matrices de covarianza de la clase 1

$\Sigma_{C2} \leftarrow \text{avg}(\Sigma_{C2,n2})$        $\triangleright$  Promediar todas las matrices de covarianza de la clase 2

$[\mathbf{D}, \mathbf{V}] \leftarrow \text{eigh}(\Sigma_{C1}, \Sigma_{C1} + \Sigma_{C2})$        $\triangleright$  Obtener solución de (4.1)

$\mathbf{W} \leftarrow \mathbf{V}^T$        $\triangleright$  Construir la matriz de filtros espaciales

---

**4.2.2.2. CSP para clasificación multi-clase:**

El segundo caso corresponde a la aplicación del CSP cuando los datos contemplan más de dos clases. Su aplicación puede verse con mayor detalle en el Algoritmo 2.

Para aplicar el filtro CSP en un caso de multi-clases (Un máximo de cuatro clases fueron utilizadas para las pruebas de esta investigación), se escogió utilizar el algoritmo desarrollado por [Grosse-Wentrup and Buss \[2008\]](#), en el cual demostraron que al aplicar JAD sobre el conjunto de matrices de covarianza promedio por clase, es posible obtener una matriz de diagonalización que permite maximizar la separabilidad entre las clases, al mismo tiempo que divide la señal EEG del subespacio de ruido. Por lo tanto, proporciona un conjunto adecuado de filtros espaciales potenciales  $\mathbf{W}$  que no pueden ser calculados de forma exacta pero sí a través de algoritmos de diagonalización como el *Algoritmo de Pham* [[Pham, 2001](#)].

---

**Algoritmo 2** Algoritmo CSP aplicado a Multi-clases
 

---

**Input** :  $\mathbf{X}_{C_1}, \mathbf{X}_{C_2}, \dots, \mathbf{X}_{C_n}$

**Output:** Matriz de proyección  $\mathbf{W}$

**for**  $n$  clases **do**

$\Sigma_{C_n, n_n} \leftarrow (\mathbf{X}_{C_n} - \mu_n)(\mathbf{X}_{C_n} - \mu_n)^T / N$	▷ Calc. Covarianza de cada epoca por clase
$\Sigma_{C_n} \leftarrow \text{avg}(\Sigma_{C_n, n_n})$	▷ Promediar todas las matrices de covarianza de cada clase

**end**

$[\mathbf{D}, \mathbf{V}] \leftarrow \text{ajd\_pham}(\Sigma_{C_1}, \Sigma_{C_2}, \dots, \Sigma_{C_n})$  ▷ Diagonalizar matrices con [[Pham, 2001](#)]

$\Sigma_T \leftarrow \text{avg}(\Sigma_{C_1}, \Sigma_{C_2}, \dots, \Sigma_{C_n})$  ▷ Promedio total de todas las matrices

**for**  $i$  filtros **do**

$tmp \leftarrow \mathbf{V}(:, i)^T \cdot \Sigma_T \cdot \mathbf{V}(:, i)$	▷ Proyectar Matriz promedio total
$\mathbf{V}(:, i) \leftarrow \mathbf{V}(:, i) / \sqrt{tmp}$	▷ Normalizar filtro

**end**

$\mathbf{W} \leftarrow \mathbf{V}^T$  ▷ Construir la matriz de filtros espaciales

---

### 4.3. Etapa de extracción y selección de características

#### 4.3.1. Extracción de Características

Luego de aplicar la técnica de filtrado espacial, se obtienen las matrices  $\mathbf{W}_p$  definidas por,

$$\mathbf{W}_p = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_c] \in \mathbb{R}^{c \times c}, \quad (4.2)$$

donde  $p$  representa el número de sub-bandas resultantes de la aplicación del banco de filtros o la descomposición wavelet, y  $c$  es el número de canales de la señal.

De esta manera, se define a la matriz  $\mathbf{W}_p^* \subset \mathbf{W}_p$  con  $\mathbf{W}_p^* \in \mathbb{R}^{c \times F}$ , donde  $F \leq c$  es el número de filtros espaciales utilizados en los posteriores cálculos, definido por la técnica de selección de características empleada. Estas matrices  $\mathbf{W}_p^*$  actuarán como filtros espaciales proyectando los datos hacia un nuevo dominio  $\mathbf{S}_p \in \mathbb{R}^{N \times F}$  que maximice la diferencia de varianza inter-clase utilizando la siguiente expresión,

$$\mathbf{S}_p = \mathbf{W}_p^* \mathbf{E}_p, \quad (4.3)$$

donde  $\mathbf{E}_p \in \mathbb{R}^{N \times c}$  representa la señal filtrada en la banda de frecuencia  $p$ , y la varianza de la señal viene siendo la principal característica que se utiliza para discernir entre las clases.

Por lo tanto, el vector de características de cada sub-banda de frecuencia  $\mathbf{f}_b$  se construye calculando el logaritmo de la varianza de cada canal sobre cada una de las señales filtradas de la siguiente manera,

$$\mathbf{f}_b = \log(\text{var}(\mathbf{S}_p)) = [f_1, f_2, \dots, f_F] \in \mathbb{R}^F, \quad (4.4)$$

para posteriormente concatenarlas en un solo vector de características  $\mathbf{F}_{total}$  de la forma,

$$\mathbf{F}_{total} = [\mathbf{f}_{b1} \mathbf{f}_{b2} \dots \mathbf{f}_{bp}] \in \mathbb{R}^d, \quad (4.5)$$

donde  $d = F * p$  corresponde a la cantidad de características totales.

Como alternativa a estas características, y aprovechando la diferencia de potencia promedio que se genera en las bandas de frecuencia relacionadas con el pensamiento motriz ( $\alpha$ ,  $\theta$ ,  $\beta$  y  $\gamma$ ), también se propuso utilizar la potencia promedio en escala logarítmica como

característica expresada como,

$$\mathbf{f}_b = \log \left( \frac{1}{N} \sum_{k=1}^N |\mathbf{S}_p|^2 \right) = [f_1, f_2, \dots, f_F] \in \mathbb{R}^F. \quad (4.6)$$

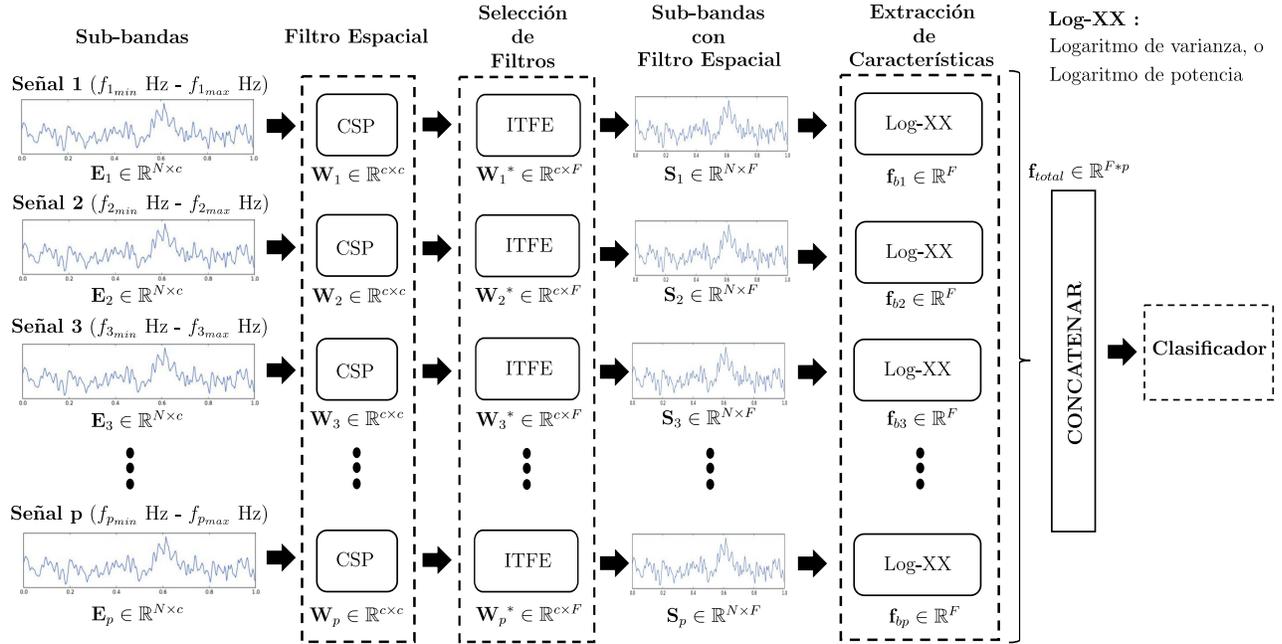
La implementación de este tipo de características se propone con la intención de someter a estudio el desempeño del clasificador basado en los resultados obtenidos por [Aljalal et al. \[2018\]](#), donde se logró demostrar que el logaritmo de la potencia promedio de la señal brinda características más diferenciadoras que la varianza. Al igual que con (4.4), las características de potencia promedio de cada banda son concatenadas construyendo un vector de características total  $\mathbf{F}_{total} \in \mathbb{R}^d$  igual a (4.5).

### 4.3.2. Selección de Mejores Características

Para poder obtener el mayor provecho de las características calculadas se planteó incorporar una etapa de selección previa a la clasificación. Esta selección se realizó probando dos métodos de selección basados en el cálculo de índices de relevancia a priori.

El primer método de selección utilizado fue ITFE, propuesto por [Grosse-Wentrup and Buss \[2008\]](#). Es un método de selección supervisado, basado en el cálculo de aproximación de información mutua, que se utiliza sobre el dominio de matrices de covarianza (Matrices simétricas definidas positivas, SPD) después de obtener las matrices de proyección  $\mathbf{W}_p$  (4.2). Su finalidad es calcular un valor de relevancia para cada una de las columnas  $\mathbf{w}$  de las matrices  $\mathbf{W}_p$ , y con ello, re-estructurar el orden de las mismas desde el mayor hasta el menor grado de relevancia. Este proceso se realiza con el objetivo de probar cada filtro en orden durante el proceso de exploración de características, y determinar cual es la cantidad de filtros ( $1 \leq F \leq c$ ) en  $\mathbf{W}_p$  que permite obtener un mayor índice Kappa durante la clasificación de los datos de validación. La implementación de esta técnica de selección en el procesamiento de la señal puede observarse en la Figura 4.8.

En la Figura 4.9 se describe la implementación del segundo método de selección, llamado Valor Laplaciano [[He et al., 2005](#)], el cual es una técnica no supervisada que se aplica luego de obtener el vector de características  $\mathbf{F}_{total} \in \mathbb{R}^d$ , donde  $d = c * p$  utilizando todos los filtros ( $F = c$ ), y concatenando todas las características basadas en el logaritmo de potencia o el logaritmo de varianza de cada sub-banda de frecuencia. Gracias a esta técnica, el vector de características es re-organizado desde el mayor al menor grado de relevancia, y también es utilizado durante el proceso de exploración de características para determinar la cantidad de



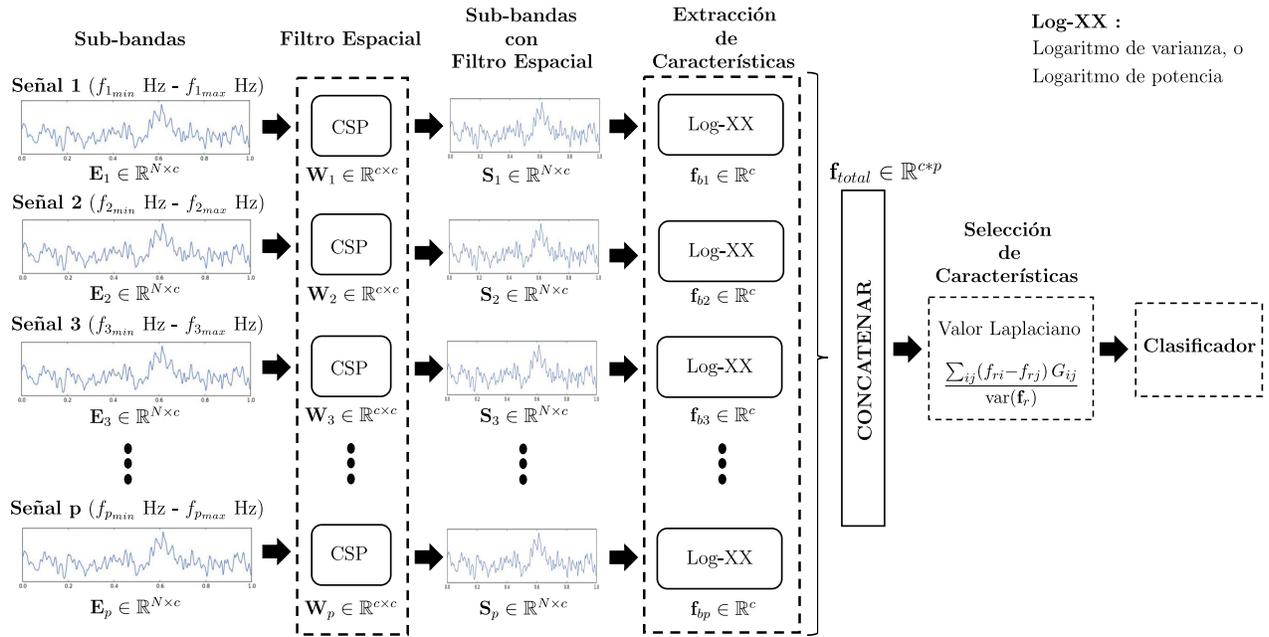
**Figura 4.8.** Esquema de las etapas de procesamiento de la señal implementando selección de filtros con ITFE.

atributos  $1 \leq n_f \leq d$  que permiten obtener un mayor índice Kappa durante la clasificación de los datos de validación.

El Algoritmo 3 describe con detalle los pasos a seguir para la exploración exhaustiva de las características, en donde se agrega cada una de ellas al vector de entrada del clasificador, según se haya establecido la relevancia de las mismas con los algoritmos de selección. Cada vez que se agrega una característica nueva, el algoritmo entrena el clasificador y calcula el resultado de validación para almacenarlo en un historial. De esta manera, el proceso se repite para cada característica agregada hasta que se han probado todas las características extraídas.

Este mismo proceso de evaluación también fue implementado para determinar la cantidad de filtros  $F$  que maximizara el rendimiento del clasificador con los datos de validación utilizando la técnica ITFE.

La división de los datos de entrenamiento y validación para esta etapa de selección es aleatoria y estratificada, y puede influir en los resultados finales si los datos contienen muestras atípicas. Por lo tanto, se utilizó el principio de validación cruzada en donde se repitió el experimento 10 veces seleccionando aleatoriamente la partición de entrenamiento y validación con la finalidad de tener un resultado estadístico más confiable, y poder determinar las mejores



**Figura 4.9.** Esquema de las etapas de procesamiento de la señal implementando selección de características con Valor Laplaciano.

---

### Algoritmo 3 Algoritmo de Selección de Características

---

**Input :**  $\mathbf{X}_{train}$

**Output:** Mejores Características  $n_f$

**for**  $i, \mathbf{X}_t, \mathbf{X}_v \leftarrow split\_10fold(\mathbf{X}_{train})$  **do**

$features \leftarrow cantidad\_columnas(\mathbf{X}_t)$

**for**  $k \leftarrow 1, features$  **do**

$clf \leftarrow modelo()$

▷ crear modelo del clasificador

$clf \leftarrow entrenar(\mathbf{X}_t[0 : k])$

▷ entrenar modelo con  $k$  características

$hist[i, k] \leftarrow predecir(\mathbf{X}_v[0 : k])$

▷ clasificar datos de validación con  $k$  características

**end**

**end**

$\mathbf{p} \leftarrow prom(hist)$

▷ Promediar cada columna

$n_f \leftarrow \arg \max(\mathbf{p}) + 1$

▷ Obtener la columna de mayor rendimiento

**return**  $n_f$

---

características o filtros, en función del mejor resultado de métrica Kappa promedio obtenido al finalizar la validación cruzada.

## 4.4. Clasificación

Para la etapa de reconocimiento se emplearon tres tipos de clasificadores basados en aprendizaje automático (SVM, KNN y LDA). Los fundamentos teóricos de cada uno de los clasificadores propuestos se describe con detalle en el Capítulo 3.

### 4.4.1. Modelo basado en SVM

Para el caso del clasificador desarrollado con SVM, se utilizó la función RBF, como kernel de transformación de los datos a clasificar, la cual se aplica mediante la siguiente ecuación,

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2\right), \quad (4.7)$$

donde  $\gamma$  es el parámetro que define hasta dónde llega la influencia de las muestras de entrenamiento seleccionadas por el modelo como vectores de soporte. Un valor bajo de  $\gamma$  indica que la influencia de los vectores de soporte es más amplia y viceversa.

En conjunto con el parámetro  $\gamma$ , también se seleccionó el factor de regularización  $C$  (ver Ecuación (3.46)), el cual agrega una penalización por cada punto de datos mal clasificado. El valor de  $C$  afecta directamente el margen calculado para la frontera de decisión del clasificador. Para las pruebas llevadas a cabo en esta investigación se definió un  $\gamma$  igual a 0.01 y un  $C$  igual a 1.0.

Debido a que el SVM por naturaleza es un clasificador binario, en los casos donde fue necesario clasificar más de dos clases se empleó el modelo de clasificación Uno contra el Resto (*One vs Rest*, OVR). Esta metodología requiere entrenar un clasificador binario SVM para cada una de las clases en el conjunto de datos. Para cada clasificador SVM, la clase de interés se considera como la clase positiva, y todas las demás clases se agrupan en una sola clase negativa. Estos modelos se utilizan posteriormente al realizar las predicciones de las muestras no clasificadas, generando una probabilidad de pertenencia con respecto a cada clase de interés. La clase que obtiene la mayor probabilidad de pertenencia entre todos los clasificadores se considera la etiqueta resultante de la predicción para la muestra.

### 4.4.2. Modelo basado en LDA

La propuesta de incorporar un modelo de reconocimiento con LDA se basó en la capacidad que tiene este clasificador de encontrar una transformación lineal que maximice la separación entre las medias de las diferentes clases, y minimice la variación dentro de cada clase. Esta transformación fue utilizada para proyectar los datos utilizados en los experimentos de la investigación hacia un espacio de menor dimensión donde la separación entre las diferentes clases es más fácil de visualizar y pueda separarse a través de una frontera lineal. El procedimiento matemático de esta técnica está descrito con mayor detalle en el Capítulo 3.

LDA requiere calcular las matrices de covarianza de las clases para poder desarrollar la descomposición de eigenvalores. Sin embargo, cuando la cantidad de características es mayor al número de muestras por clase, las estimaciones de la covarianza no logran tener un rango completo, y no puede ser invertida.

Para resolver esta limitante, la implementación del modelo LDA se realizó aplicando la técnica de Descomposición de Valores Singulares (*Singular Value Decomposition*, SVD). Con esta técnica no es necesario calcular las matrices de covarianza ya que por definición es equivalente a,

$$\Sigma = \frac{1}{N-1} \mathbf{X}^T \mathbf{X} = \frac{1}{N-1} \mathbf{V} \mathbf{S}^T \mathbf{S} \mathbf{V}^T, \quad (4.8)$$

donde  $\mathbf{X}^T \mathbf{X}$  es el segundo momento (cuando los datos  $\mathbf{X}$  están centrados con respecto a su media es igual a la matriz de covarianza),  $\mathbf{V}$  es igual a los eigenvectores de  $\mathbf{X}^T \mathbf{X}$  y  $\mathbf{S}$  es la matriz diagonal con los valores singulares de la descomposición (equivalentes a la raíz cuadrada de los eigenvalores de  $\mathbf{X}^T \mathbf{X}$ ). El beneficio de este modelo de clasificación basado en SVD es que no requiere el ajuste de hiper-parámetros adicionales y puede ser aplicado en clasificación multi-clases.

### 4.4.3. Modelo basado en KNN

El tercer modelo de reconocimiento evaluado es el clasificador KNN, el cual es un algoritmo de aprendizaje automático de clasificación supervisada que permite encontrar los  $k$  datos de entrenamiento más cercanos a un dato de entrada en un conjunto de datos etiquetados. En la clasificación KNN, se utiliza la mayoría de los votos de los  $k$  datos más cercanos para determinar la etiqueta del dato de entrada.

La métrica de distancia utilizada para determinar la cercanía espacial de los datos fue la Semejanza de Coseno (3.38), donde  $x$  representa el dato vecino de entrenamiento y  $y$  representa el dato de entrada que se desea clasificar.

El número de  $K$  vecinos a tomar en cuenta se definió como,

$$K = \sqrt{N}, \quad (4.9)$$

donde  $N$  representa el número de datos de entrenamiento.

El resultado final se realizó tomando en cuenta los votos ponderados de los  $K$  vecinos más cercanos donde el peso de los vecinos se pondera por el inverso de su distancia. Esto permite asignar más importancia a los puntos más cercanos al dato de entrada que a los que se encuentren más alejados.

Este modelo basado en KNN se incorporó a la investigación ya que es un algoritmo simple, no necesita determinar la distribución estadística del conjunto de datos y puede funcionar correctamente con pocos datos de entrenamiento. Esta última cualidad cobra importancia tomando en cuenta que las bases de datos utilizadas en la investigación ofrecen pocos datos para entrenar. La principal desventaja de este modelo es el mayor costo computacional requerido en comparación a los dos modelos anteriores, pero sigue siendo una referencia robusta en cuanto a clasificadores clásicos que debe ser tomada en cuenta al momento de determinar los mejores rendimientos en los experimentos.

## 4.5. Métricas de desempeño

Las métricas de desempeño utilizadas para evaluar el rendimiento de los modelos propuestos, en función de su capacidad para predecir las clases correctas de imaginación motoras contenidas en las bases de datos de estudio, son:

- **Coefficiente Kappa de Cohen:** El coeficiente Kappa es una medida cuantitativa de la concordancia entre las predicciones del modelo y las clases reales, comparando la tasa de acuerdo observado con la tasa de acuerdo esperado por azar.

$$k = \frac{P_o - P_e}{1 - P_e}, \quad (4.10)$$

donde  $P_o$  es el acuerdo observado relativo entre los observadores y  $P_e$  es la tasa de acuerdo esperado por azar. La tasa de acuerdo por azar  $P_e$  se calcula de la siguiente manera:

$$P_e = \frac{1}{N^2} \sum_{k=1}^K n_{k1} n_{k2}, \quad (4.11)$$

donde  $N$  es el número total de predicciones,  $K$  es el número de clases,  $n_{k1}$  es el número de veces en el que la etiqueta  $k$  aparece en las predicciones y  $n_{k2}$  es el número de veces en el que la etiqueta  $k$  es verdadera. Y la tasa de acuerdo relativo se describe como,

$$P_o = \frac{\sum_{k=1}^K n_{k2}}{N}. \quad (4.12)$$

- **Exactitud:** La exactitud mide la proporción de predicciones correctas realizadas por el modelo en comparación con el total de predicciones,

$$Exactitud = \frac{\sum_{k=1}^K n_{k2}}{N}. \quad (4.13)$$

donde  $K$  corresponde al número de clases,  $n_{k2}$  es el número de veces en el que la etiqueta  $k$  es verdadera y  $N$  es el número total de predicciones.

- **Precisión:** Se refiere a la proporción de predicciones positivas verdaderas con respecto a todas las predicciones positivas realizadas por el modelo,

$$Precision = \frac{TP}{TP + FP}. \quad (4.14)$$

donde  $TP$  es el número de verdaderos positivos y  $FP$  es el número de falsos positivos.

- **Exhaustividad:** La exhaustividad mide la proporción de verdaderos positivos con respecto a todos los casos positivos reales en el conjunto de datos,

$$Exhaustividad = \frac{TP}{TP + FN}. \quad (4.15)$$

donde  $FN$  es el número de falsos negativos.

- **Valor F1:** Es la media armónica de precisión y exhaustividad. Proporciona un equilibrio entre ambas métricas y es útil cuando se necesita una medida única que combine ambas,

$$ValorF1 = \frac{2 * precision * exhaustividad}{precision + exhaustividad}. \quad (4.16)$$

- **Área bajo la curva ROC:** La curva Característica Operativa del Receptor (*Receiver Operating Characteristic*, ROC) representa la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos en diferentes umbrales de clasificación. El AUC-ROC mide la capacidad discriminativa global del modelo.

## 4.6. Descripción de Marcos de Trabajo

### 4.6.1. Propuesta 1: OFB-CSP-FS

En el primer marco de trabajo propuesto, llamado OFB-CSP-FS (*Overlapped Filter Bank Common Spatial Pattern with Feature Selection*), se utilizó el banco de filtros pasa-banda IIR Butterwoth con el objetivo de reducir la influencia de los artefactos de ruido presentes en las señales y extraer la información contenida en las bandas de frecuencia de mayor interés relacionadas con el pensamiento motor. Posteriormente, durante la etapa de pre-procesamiento se utilizó la técnica de filtrado espacial CSP a cada una de las sub-bandas generadas por el banco de filtros, para acondicionar los datos y aumentar la diferencia entre las clases. Se probaron separadamente, el logaritmo de la varianza y el logaritmo de la potencia promedio de la señal como las principales características del sistema de reconocimiento, y adicionalmente, se comprobó el efecto de cada una de las técnicas de selección de características propuestas. Finalmente, se verificó el rendimiento de los tres clasificadores clásicos descritos.

### 4.6.2. Propuesta 2: WPD-CSP-FS

En la segunda propuesta, llamada WPD-CSP-FS (*Wavelet Packets Descomposition Common Spatial Pattern with Feature Selection*), se definió como técnica de pre-procesamiento el método de Descomposición WPD en sustitución al banco de filtros IIR aplicado en la propuesta 1. El objetivo de incorporar esta etapa era verificar el efecto que ejerce la Descomposición Wavelet al momento de obtener la información relevante al pensamiento motor en las diferentes bandas de frecuencia.

Las siguientes etapas de pre-procesamiento, filtrado espacial, extracción de características y clasificación, se conservaron de la misma manera que las utilizadas en la propuesta OFB-CSP-FS. Esto con la finalidad de determinar el impacto de la Descomposición

Wavelet con respecto al Banco de Filtros IIR como herramientas de pre-procesamiento para la detección del pensamiento motor aplicado a interfaces cerebro-computadora.

Los resultados obtenidos de cada uno de los experimentos y las diferentes combinaciones de técnicas propuestas pueden verse con mas detalles en el Capítulo 5, referente a los resultados de la investigación.

# CAPÍTULO 5

---

## Resultados experimentales

---

---

<b>5.1</b>	<b>Resultados con BCI Competition IV 2a</b>	<b>74</b>
5.1.1	Clasificación Binaria	74
5.1.2	Clasificación Multiclases	83
<b>5.2</b>	<b>Resultados con base de datos BCI Competition IV 2b</b>	<b>92</b>
5.2.1	Resultados con Clasificador SVM	94
5.2.2	Resultados con Clasificador KNN	95
5.2.3	Resultados con Clasificador LDA	97
5.2.4	Análisis de Resultados	99
<b>5.3</b>	<b>Comparativa con la literatura</b>	<b>101</b>

---

Con el objetivo de comprobar el desempeño de los marcos de trabajo propuestos en esta investigación, tanto el marco de trabajo OFB-CSP-FS como el WPD-CSP-FS, se utilizaron para la clasificación de las señales suministradas por las bases de datos BCI Competition IV dataset 2a y BCI Competition IV dataset 2b. En cada marco de trabajo se sometió a prueba el efecto que tendría utilizar el logaritmo de la potencia promedio de las señales (logpower) o el logaritmo de la varianza de las señales (logvar) como método de extracción de características

después de aplicar la etapa de pre-procesamiento. Igualmente se contempló la combinación de dos técnicas de selección de características, una basada en información mutua (ITFE) antes del cálculo del vector de características, y la otra basada en el valor laplaciano (LS) después de calcularse el vector de características.

En esta sección se presentarán todos los resultados obtenidos durante las pruebas realizadas con ambas bases de datos, aplicando las variantes propuestas en los modelos OFB-CSP-FS y WPD-CSP-FS, con el objetivo de verificar el desempeño de los mismos aplicados a la detección de MI-BCI.

## 5.1. Resultados con BCI Competition IV 2a

En cuanto a esta primera base de datos, [Tangermann et al. \[2012\]](#) proponen dos grandes retos. El primero es la presencia de artefactos de ruido EOG durante la adquisición de los datos, que puede afectar directamente el desempeño de los modelos de clasificación, y el segundo está asociado a la dificultad de obtener patrones comunes entre señales de sesiones diferentes. Por esta razón, para someter a prueba los modelos, el entrenamiento se realizó con datos tomados en días diferentes a los datos de prueba.

Durante las pruebas con esta base de datos, se utilizó una partición de 72 muestras por clase para el entrenamiento de los modelos, y otra de 72 muestras por clase para la prueba de los mismos. Ambas particiones de datos fueron adquiridas en sesiones realizadas en días diferentes con nueve sujetos de prueba distintos, y están previamente definidas por la BCI Competition IV. Adicionalmente, la base de datos contempla un total de cuatro clases, determinadas por pensamientos de movilidad en brazo izquierdo (L), brazo derecho (R), pies (F) y lengua (T). Con la finalidad de observar el efecto de las diferentes clases durante la clasificación, se propuso una etapa de pruebas donde cada marco de trabajo se utilizó con las seis combinaciones binarias posibles entre las cuatro clases, y luego una etapa de clasificación multiclase donde se realizaron cuatro combinaciones posibles.

### 5.1.1. Clasificación Binaria

A continuación se presentan los resultados obtenidos utilizando los clasificadores SVM, KNN y LDA, observando su influencia en cada una de las diferentes combinaciones binarias

de clases junto con las técnicas de extracción y selección de características propuestas.

Las métricas utilizadas para valorar el desempeño de los modelos fueron el índice kappa, la exactitud, el área bajo la curva (AUC), la precisión y la exhaustividad (Recall). Los resultados expuestos corresponden a los valores promedio obtenidos después de realizar cinco ejecuciones de cada modelo clasificando la porción de datos de prueba, donde la selección y validación de las mejores características se realizó a través de una validación cruzada de 10 ejecuciones con datos escogidos aleatoriamente de la porción de datos de entrenamiento, así como se describe en el Capítulo 4. Los modelos con las mejores métricas para cada caso de estudio han sido resaltados con color azul.

#### 5.1.1.1. Resultados con Clasificador SVM

**Tabla 5.1.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom $\pm$ Std
ITFE con logpower	Kappa	0.613	<b>0.759</b>	0.722	0.782	0.74	0.597	0.702 $\pm$ 0.030
	Exactitud [%]	80.68	<b>87.95</b>	86.08	89.07	87.02	79.85	85.11 $\pm$ 1.56
	AUC [%]	89.95	<b>96.17</b>	94.17	97.09	94.78	91.88	94.00 $\pm$ 0.42
	Precisión [%]	78.98	<b>86.08</b>	86.61	90.12	92.95	82.66	86.23 $\pm$ 2.03
	Recall [%]	87.66	<b>92.62</b>	86.45	88.73	79.78	80.09	85.89 $\pm$ 2.43
ITFE con logvar	Kappa	0.61	0.737	0.724	0.794	0.724	0.592	0.697 $\pm$ 0.032
	Exactitud [%]	80.51	86.85	86.22	89.69	86.17	79.61	84.84 $\pm$ 1.58
	AUC [%]	89.74	95.99	94.22	97.15	94.52	92.13	93.95 $\pm$ 0.46
	Precisión [%]	78.66	84.85	87.26	90.79	90.86	82.85	85.87 $\pm$ 1.94
	Recall [%]	87.22	92.53	85.96	89.07	81.11	79.57	85.91 $\pm$ 2.62
LS con logpower	Kappa	0.615	0.743	0.727	0.795	<b>0.762</b>	<b>0.621</b>	<b>0.711 <math>\pm</math> 0.038</b>
	Exactitud [%]	80.77	87.13	86.37	89.74	<b>88.07</b>	<b>81.05</b>	<b>85.52 <math>\pm</math> 1.88</b>
	AUC [%]	90.07	96.77	94.38	97.04	<b>95.07</b>	<b>92.01</b>	<b>94.22 <math>\pm</math> 0.86</b>
	Precisión [%]	78.82	88.22	87.8	92.03	<b>92.67</b>	<b>82.65</b>	<b>87.03 <math>\pm</math> 2.44</b>
	Recall [%]	83.27	88.02	84.85	87.47	<b>82.13</b>	<b>81.42</b>	<b>84.52 <math>\pm</math> 2.85</b>
LS con logvar	Kappa	<b>0.614</b>	0.719	<b>0.729</b>	<b>0.801</b>	0.758	0.614	0.706 $\pm$ 0.039
	Exactitud [%]	<b>80.71</b>	85.96	<b>86.47</b>	<b>90.05</b>	87.92	80.68	85.29 $\pm$ 1.96
	AUC [%]	<b>88.71</b>	96.57	<b>94.79</b>	<b>97.09</b>	95.17	91.92	94.04 $\pm$ 0.94
	Precisión [%]	<b>79.67</b>	87.63	<b>87.92</b>	<b>92.25</b>	92.88	82.86	87.20 $\pm$ 02.44
	Recall [%]	<b>84.6</b>	87.19	<b>84.78</b>	<b>87.71</b>	81.51	81.14	84.49 $\pm$ 2.76

Como puede observarse en la Tabla 5.1, el marco de trabajo OFB-CSP-FS con clasificador SVM utilizando LS con logpower, obtuvo los mejores resultados generales con

un kappa promedio de 0.711 y una desviación estándar de 0.038. Evaluando cada caso por separado, es notorio que el modelo de LS con logvar obtuvo el mejor desempeño para la clasificación entre la combinación de brazo izquierdo con brazo derecho (LR), brazo izquierdo con lengua (LT) y brazo derecho con pies (RF); el modelo ITFE con logpower obtuvo el mejor desempeño en la clasificación de brazo izquierdo con pies (LF) y el modelo LS con logpower obtuvo el mejor desempeño en la clasificación de brazo derecho con lengua (RT) y pies con lengua (FT).

**Tabla 5.2.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom $\pm$ Std
ITFE con logpower	Kappa	0.576	<b>0.707</b>	0.668	0.797	0.681	0.597	0.671 $\pm$ 0.028
	Exactitud [%]	78.8	<b>85.37</b>	83.41	89.85	84.04	79.85	83.55 $\pm$ 1.41
	AUC [%]	86.94	<b>94.14</b>	92.95	96.09	93.05	89.61	92.13 $\pm$ 0.54
	Precisión [%]	77.54	<b>83.52</b>	87.06	91.02	90.58	81.57	85.21 $\pm$ 2.2
	Recall [%]	81.6	<b>90.34</b>	79.14	88.55	75.0	79.69	82.39 $\pm$ 2.96
ITFE con logvar	Kappa	<b>0.586</b>	0.696	0.667	<b>0.806</b>	0.683	0.595	0.672 $\pm$ 0.028
	Exactitud [%]	<b>79.31</b>	84.8	83.35	<b>90.29</b>	84.14	79.75	83.61 $\pm$ 1.38
	AUC [%]	<b>87.42</b>	94.15	92.66	<b>96.11</b>	93.07	89.93	92.22 $\pm$ 0.5
	Precisión [%]	<b>78.02</b>	84.17	87.2	<b>91.75</b>	90.71	81.09	85.49 $\pm$ 2.14
	Recall [%]	<b>81.98</b>	88.33	79.17	<b>88.77</b>	75.31	80.59	82.36 $\pm$ 3.06
LS con logpower	Kappa	0.572	0.68	<b>0.677</b>	0.791	0.713	<b>0.639</b>	<b>0.679 <math>\pm</math> 0.031</b>
	Exactitud [%]	78.61	84.0	<b>83.83</b>	89.55	85.63	<b>81.93</b>	<b>83.92 <math>\pm</math> 1.53</b>
	AUC [%]	86.68	94.76	<b>93.52</b>	96.06	93.31	<b>90.89</b>	<b>92.54 <math>\pm</math> 0.51</b>
	Precisión [%]	78.32	86.85	<b>88.73</b>	90.53	91.75	<b>85.78</b>	<b>86.99 <math>\pm</math> 2.5</b>
	Recall [%]	80.77	81.48	<b>78.21</b>	88.58	77.56	<b>78.83</b>	<b>80.9 <math>\pm</math> 2.73</b>
LS con logvar	Kappa	0.56	0.697	0.675	0.794	<b>0.718</b>	0.626	0.678 $\pm$ 0.035
	Exactitud [%]	77.98	84.84	83.75	89.72	<b>85.93</b>	81.28	83.92 $\pm$ 1.75
	AUC [%]	86.84	94.74	93.21	96.15	<b>93.28</b>	90.04	92.38 $\pm$ 0.52
	Precisión [%]	77.02	87.2	88.1	90.53	<b>92.03</b>	83.61	86.42 $\pm$ 2.58
	Recall [%]	81.42	83.52	79.07	88.92	<b>77.59</b>	79.29	81.64 $\pm$ 3.4

En la Tabla 5.2 se muestran los resultados obtenidos utilizando el clasificador SVM con las diferentes combinaciones de técnicas en el marco de trabajo WPD-CSP-FS. En ella se puede observar que, al igual que en el marco de trabajo anterior, el modelo que obtuvo el mejor desempeño fue LS con logpower logrando 0.679 de métrica kappa y una desviación estándar de 0.031.

Para ambos marcos de trabajo se cumple que en los casos LR y LF se evidencian un mayor porcentaje de exhaustividad en comparación con la precisión, indicando que clasificó

mejor las señales de la clase R y F que las señales de la clase L. Caso contrario para las combinaciones LT, RF y RT que obtuvieron mejor precisión que exhaustividad, indicando que la clase L se clasificó mejor que la clase T, y la clase R se clasificó mejor que la clase F y T. Por último en la combinación FT, la precisión y exhaustividad tienen valores muy cercanos indicando que se obtuvieron desempeños similares al momento de clasificar las señales de la clase F y la clase T, con una ligera inclinación a clasificar mejor las señales de la clase F. Para OFB-CSP-FS, el caso que presentó el mejor resultado de clasificación fue la combinación RF utilizando LS con logvar, el cual logró un total de 0.801 en la métrica kappa con 97.09 % de AUC. En cuanto a WPD-CSP-FS, el mejor resultado fue obtenido utilizando ITFE con logvar en el caso RF con 0.806 de kappa y 96.11 % de AUC.

### 5.1.1.2. Resultados con Clasificador KNN

**Tabla 5.3.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom $\pm$ Std
ITFE con logpower	Kappa	0.618	0.669	0.707	0.693	0.673	0.555	0.652 $\pm$ 0.021
	Exactitud [%]	80.88	83.47	85.37	84.68	83.66	77.73	82.63 $\pm$ 1.05
	AUC [%]	88.78	93.13	92.55	94.08	93.74	89.84	92.02 $\pm$ 0.55
	Precisión [%]	80.52	90.34	93.48	94.47	94.77	82.97	89.43 $\pm$ 1.37
	Recall [%]	82.01	75.18	75.49	74.72	71.45	75.28	75.69 $\pm$ 2.0
ITFE con logvar	Kappa	<b>0.625</b>	0.668	0.705	0.685	0.668	0.558	0.652 $\pm$ 0.022
	Exactitud [%]	<b>81.26</b>	83.4	85.26	84.26	83.41	77.92	82.58 $\pm$ 1.09
	AUC [%]	<b>88.65</b>	93.11	92.4	94.1	93.81	89.75	91.97 $\pm$ 0.65
	Precisión [%]	<b>81.35</b>	90.65	93.53	94.17	94.99	83.08	89.63 $\pm$ 1.55
	Recall [%]	<b>81.42</b>	74.97	75.09	74.26	70.68	76.08	75.42 $\pm$ 2.07
LS con logpower	Kappa	0.589	<b>0.682</b>	0.719	<b>0.703</b>	0.697	0.564	0.659 $\pm$ 0.03
	Exactitud [%]	79.44	<b>84.12</b>	85.97	<b>85.14</b>	84.85	78.23	82.96 $\pm$ 1.52
	AUC [%]	87.75	<b>93.26</b>	91.82	<b>93.7</b>	92.85	90.01	91.56 $\pm$ 1.0
	Precisión [%]	79.31	<b>90.95</b>	90.95	<b>92.56</b>	94.43	82.56	88.46 $\pm$ 1.64
	Recall [%]	81.17	<b>77.07</b>	78.64	<b>77.31</b>	74.17	76.57	77.49 $\pm$ 2.22
LS con logvar	Kappa	0.61	0.681	<b>0.732</b>	0.692	<b>0.709</b>	<b>0.572</b>	<b>0.666 <math>\pm</math> 0.031</b>
	Exactitud [%]	80.48	84.06	<b>86.62</b>	84.61	<b>85.43</b>	<b>78.61</b>	<b>83.3 <math>\pm</math> 1.53</b>
	AUC [%]	88.65	92.62	<b>92.43</b>	93.43	<b>93.07</b>	<b>90.38</b>	<b>91.76 <math>\pm</math> 0.84</b>
	Precisión [%]	80.55	91.33	<b>91.81</b>	92.83	<b>94.1</b>	<b>82.48</b>	<b>88.85 <math>\pm</math> 1.53</b>
	Recall [%]	80.37	76.39	<b>79.38</b>	76.17	<b>75.56</b>	<b>77.59</b>	<b>77.58 <math>\pm</math> 2.79</b>

Los resultados obtenidos con el marco de trabajo OFB-CSP-FS en los mismos casos de prueba pero utilizando el modelo de clasificador KNN se muestran en la Tabla 5.3. El

modelo basado en LS con logvar obtuvo los mejores resultados generales con 0.666 de métrica kappa y 0.031 de desviación estándar. Igualmente el modelo LS con logvar obtuvo los mejores desempeños de clasificación para los casos LT, RT y FT, el modelo LS con logpower superó al resto en los casos LF y RF, y el modelo ITFE con logvar obtuvo el mejor desempeño en el caso LR.

**Tabla 5.4.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom.
ITFE con logpower	Kappa	0.569	0.654	0.641	<b>0.709</b>	0.576	0.544	0.616 ± 0.02
	Exactitud [%]	78.47	82.69	82.07	<b>85.43</b>	78.79	77.19	80.77 ± 0.99
	AUC [%]	85.79	92.6	90.69	<b>94.34</b>	90.51	88.38	90.38 ± 0.61
	Precisión [%]	80.04	92.99	94.12	<b>95.29</b>	94.01	83.99	90.07 ± 1.4
ITFE con logvar	Recall [%]	72.87	70.25	67.35	<b>74.91</b>	61.11	72.44	69.82 ± 2.08
	Kappa	<b>0.571</b>	0.645	0.648	0.704	0.566	0.558	0.615 ± 0.022
	Exactitud [%]	<b>78.58</b>	82.25	82.41	85.2	78.3	77.9	80.77 ± 1.08
	AUC [%]	<b>85.81</b>	92.45	90.56	94.15	90.5	88.53	90.33 ± 0.7
LS con logpower	Precisión [%]	<b>80.16</b>	92.74	93.82	95.19	94.0	85.14	90.18 ± 1.38
	Recall [%]	<b>73.61</b>	69.35	68.12	74.72	59.75	72.75	69.72 ± 2.19
	Kappa	0.544	0.666	<b>0.696</b>	0.704	<b>0.652</b>	<b>0.58</b>	<b>0.64 ± 0.037</b>
	Exactitud [%]	77.21	83.29	<b>84.8</b>	85.21	<b>82.59</b>	<b>79.01</b>	<b>82.02 ± 1.84</b>
LS con logvar	AUC [%]	83.21	91.77	<b>90.87</b>	92.82	<b>91.2</b>	<b>88.86</b>	<b>89.79 ± 1.43</b>
	Precisión [%]	78.17	90.29	<b>92.38</b>	92.1	<b>93.62</b>	<b>81.89</b>	<b>88.08 ± 2.58</b>
	Recall [%]	73.92	73.92	<b>74.88</b>	78.02	<b>69.57</b>	<b>78.33</b>	<b>74.77 ± 3.32</b>
	Kappa	0.569	<b>0.667</b>	0.687	0.691	0.644	0.561	0.636 ± 0.036
LS con logpower	Exactitud [%]	78.46	<b>83.38</b>	84.34	84.55	82.21	78.02	81.83 ± 1.79
	AUC [%]	84.46	<b>91.35</b>	91.02	92.76	91.22	88.64	89.91 ± 1.43
	Precisión [%]	79.79	<b>90.55</b>	92.3	92.15	94.11	80.65	88.26 ± 2.5
	Recall [%]	75.53	<b>73.76</b>	74.01	76.48	68.15	78.52	74.41 ± 3.35

En cuanto a los resultados logrados con el marco de trabajo WPD-CSP-FS, descritos en la Tabla 5.4. El modelo basado en LS con logpower demostró obtener los mejores resultados generales con 0.679 de métrica kappa y 0.031 de desviación estándar. En esta oportunidad, el modelo LS con logvar obtuvo el mejor desempeño de clasificación únicamente para el caso RT, el modelo LS con logpower superó al resto en los casos FT y LT, el modelo ITFE con logvar obtuvo el mejor desempeño en los casos LR y RF, y finalmente, el modelo ITFE con logpower obtuvo los mejores resultados de clasificación para el caso LF.

Al aplicar el modelo de clasificación KNN bajo las mismas condiciones, tanto con OFB-CSP-FS como con WPD-CSP-FS, es posible identificar tendencias diferentes a las

obtenidas con el clasificador SVM. Para los casos LF y LT, la clase L obtuvo mejor índice de clasificación que las clases F y T. No obstante, se mantuvo la misma tendencia en cuanto a la mejor clasificación de la clase R con respecto a las clases F y T. Para el caso LR con OFB-CSP-FS ambas clases obtuvieron un acierto similar en su clasificación, mientras que con WPD-CSP-FS se logró clasificar mejor la clase L que la clase R. El caso que demostró el mejor desempeño en clasificación con OFB-CSP-FS fue el LT con 0.732 de métrica kappa y 92.43 % de AUC, y el que logró el mejor índice de clasificación con WPD-CSP-FS fue el caso RF con 0.709 de kappa y 94.34 % de AUC.

### 5.1.1.3. Resultados con Clasificador LDA

Los resultados obtenidos con el modelo de clasificación LDA con el marco de trabajo OFB-CSP-FS, se muestran en la Tabla 5.5, donde nuevamente el modelo con mejor resultados generales es el LS con logvar obteniendo 0.594 de métrica kappa con 0.064 de desviación estándar. Este mismo modelo obtuvo los mejores desempeños en todas los casos de estudio, excepto en el caso LR, donde el mejor desempeño fue obtenido por la combinación de LS con logpower.

En esta oportunidad, durante las pruebas realizadas con OFB-CSP-FS para los casos LR, LF, LT y RF, los resultados de precisión y exhaustividad obtenidos fueron cercanos indicando que el desempeño de clasificación para ambas clases en cada caso fue similar. En cuanto a los casos RT y FT, sí se observa una mejor clasificación de la clase R y la clase F con respecto a la clase T. El caso LT obtuvo mejor desempeño en clasificación con 0.656 de métrica kappa y 90.71 % de AUC.

Por otro lado, los resultados obtenidos por el marco de trabajo WPD-CSP-FS, descritos en la Tabla 5.6, indican que el mejor modelo de clasificación fue el LS con logvar con una métrica kappa general de 0.624 con desviación estándar de 0.067. WPD-CSP-FS describe resultados con tendencias diferentes a los obtenidos por OFB-CSP-FS. Los casos LR y FT presentan índices de precisión y exhaustividad cercanos, lo que indica que la clasificación entre cada par de clases en estos casos fue similar. En los casos LF y LT, se logra observar una tendencia a clasificar mejor las señales de la clase L con respecto a la clase F y clase T. Para el caso RF se obtuvo una mejor clasificación de la clase F, y en la clase RT, el modelo logró un mejor desempeño de clasificación en señales de la clase R.

**Tabla 5.5.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom $\pm$ Std
ITFE con logpower	Kappa	0.525	0.567	0.641	0.6	0.591	0.407	0.555 $\pm$ 0.061
	Exactitud [%]	76.24	78.36	82.04	79.98	79.54	70.32	77.75 $\pm$ 3.04
	AUC [%]	84.34	89.5	90.73	89.97	91.04	83.0	88.1 $\pm$ 2.44
	Precisión [%]	79.96	82.36	84.24	79.08	89.14	76.18	81.83 $\pm$ 3.73
	Recall [%]	71.48	76.54	82.07	85.43	69.1	68.61	75.54 $\pm$ 5.94
ITFE con logvar	Kappa	0.526	0.591	0.656	0.589	0.595	0.376	0.556 $\pm$ 0.065
	Exactitud [%]	76.33	79.57	82.81	79.43	79.75	68.83	77.79 $\pm$ 3.23
	AUC [%]	83.99	89.99	90.63	89.02	91.1	82.07	87.8 $\pm$ 2.27
	Precisión [%]	79.2	83.06	84.36	78.66	88.5	76.33	81.68 $\pm$ 3.36
	Recall [%]	72.75	77.56	83.83	85.71	69.85	67.41	76.18 $\pm$ 6.54
LS con logpower	Kappa	<b>0.57</b>	0.601	0.654	0.633	0.606	0.476	0.59 $\pm$ 0.065
	Exactitud [%]	<b>78.49</b>	80.03	82.7	81.64	80.32	73.78	79.49 $\pm$ 3.22
	AUC [%]	<b>84.32</b>	91.23	90.73	90.58	91.49	85.54	88.98 $\pm$ 2.02
	Precisión [%]	<b>79.19</b>	83.22	84.7	83.06	88.35	78.97	82.92 $\pm$ 3.87
	Recall [%]	<b>79.07</b>	80.22	81.14	82.04	72.25	72.35	77.85 $\pm$ 6.25
LS con logvar	Kappa	0.562	<b>0.606</b>	<b>0.656</b>	<b>0.651</b>	<b>0.614</b>	<b>0.476</b>	<b>0.594 <math>\pm</math> 0.064</b>
	Exactitud [%]	78.09	<b>80.29</b>	<b>82.81</b>	<b>82.53</b>	<b>80.68</b>	<b>73.8</b>	<b>79.7 <math>\pm</math> 3.19</b>
	AUC [%]	84.15	<b>90.8</b>	<b>90.71</b>	<b>90.73</b>	<b>90.72</b>	<b>86.71</b>	<b>88.97 <math>\pm</math> 2.57</b>
	Precisión [%]	78.56	<b>82.88</b>	<b>83.09</b>	<b>83.65</b>	<b>88.7</b>	<b>80.32</b>	<b>82.87 <math>\pm</math> 3.91</b>
	Recall [%]	77.84	<b>82.31</b>	<b>84.14</b>	<b>82.81</b>	<b>72.93</b>	<b>71.27</b>	<b>78.55 <math>\pm</math> 6.08</b>

#### 5.1.1.4. Análisis de Resultados

Finalizadas las pruebas utilizando el marco de trabajo OFB-CSP-FS, es posible comparar la actuación de los clasificadores propuestos para verificar cual fue el mejor modelo de acuerdo a sus resultados en prueba. En la Figura 5.1, se describen los resultados expresados en métrica kappa promedio del mejor modelo obtenido con cada clasificador durante las pruebas realizadas en combinación binaria de clases. En ella se puede observar que los modelos que utilizaron el clasificador SVM superaron al resto significativamente en todos los casos exceptuando en los casos LR y LT, donde el clasificador KNN obtuvo un desempeño ligeramente superior. Adicionalmente, se puede destacar que el mejor modelo obtenido con el clasificador SVM obtuvo su máximo desempeño en el caso RF, donde consecuentemente también se aprecia una mayor diferencia con respecto al desempeño de los demás clasificadores.

En la Tabla 5.7 se describe detalladamente los resultados obtenidos por cada sujeto de pruebas utilizando el mejor modelo con clasificador SVM, el cual contempla el uso del

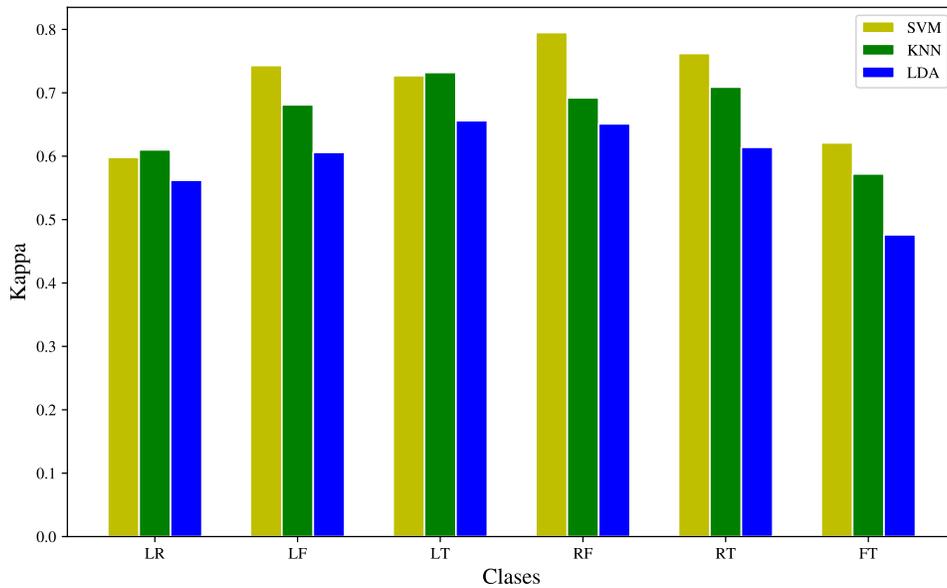
**Tabla 5.6.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación binaria de clases.

Combinación	Métrica	LR	LF	LT	RF	RT	FT	Prom.
ITFE con logpower	Kappa	<b>0.531</b>	0.606	0.599	0.735	0.56	0.467	0.583 ± 0.052
	Exactitud [%]	<b>76.53</b>	80.32	79.95	86.73	77.99	73.36	79.15 ± 2.6
	AUC [%]	<b>82.8</b>	90.0	89.51	93.24	87.98	83.26	87.8 ± 2.4
	Precisión [%]	<b>76.52</b>	80.97	85.05	85.94	86.27	76.31	81.84 ± 2.55
	Recall [%]	<b>76.94</b>	82.38	74.78	89.51	66.85	75.03	77.58 ± 4.63
ITFE con logvar	Kappa	0.514	0.595	0.586	0.706	0.573	0.466	0.573 ± 0.054
	Exactitud [%]	75.71	79.75	79.31	85.29	78.66	73.29	78.67 ± 2.7
	AUC [%]	82.55	89.79	88.49	92.19	88.88	82.75	87.44 ± 2.55
	Precisión [%]	75.56	80.66	83.97	84.63	86.31	74.51	80.94 ± 2.76
	Recall [%]	76.51	81.11	74.44	88.86	67.19	77.93	77.67 ± 5.01
LS con logpower	Kappa	0.503	<b>0.622</b>	0.657	0.735	<b>0.713</b>	0.498	0.621 ± 0.067
	Exactitud [%]	75.12	<b>81.13</b>	82.87	86.74	<b>85.63</b>	74.92	81.07 ± 3.33
	AUC [%]	79.9	<b>89.06</b>	91.37	93.61	<b>92.16</b>	83.1	88.2 ± 2.98
	Precisión [%]	75.53	<b>82.94</b>	86.38	86.38	<b>89.24</b>	77.78	83.04 ± 3.9
	Recall [%]	76.39	<b>80.56</b>	78.08	88.52	<b>80.31</b>	75.31	79.86 ± 4.99
LS con logvar	Kappa	0.507	0.612	<b>0.666</b>	<b>0.736</b>	0.704	<b>0.516</b>	<b>0.624 ± 0.067</b>
	Exactitud [%]	75.37	80.57	<b>83.3</b>	<b>86.79</b>	85.19	<b>75.8</b>	<b>81.17 ± 3.34</b>
	AUC [%]	81.68	88.04	<b>91.73</b>	<b>93.48</b>	91.54	<b>84.23</b>	<b>88.45 ± 3.07</b>
	Precisión [%]	75.45	82.36	<b>86.92</b>	<b>85.96</b>	88.32	<b>77.48</b>	<b>82.75 ± 3.85</b>
	Recall [%]	76.7	80.34	<b>78.89</b>	<b>89.11</b>	80.4	<b>77.38</b>	<b>80.47 ± 4.25</b>

logaritmo de la potencia promedio como característica y el método de valor laplaciano como selector de dichas características. En ella se observa que el sujeto A01, resaltado en color azul, obtuvo el mejor desempeño en su clasificación general, donde sus mejores resultados se pueden apreciar durante las pruebas en los casos LT y RT. El caso contrario ocurre con el sujeto A06, remarcado en color rojo, con el cual se presentaron los peores resultados de clasificación siendo el caso FT donde se obtuvo su mejor desempeño con un kappa promedio de 0.572.

Por otra parte, analizando los resultados obtenidos con el marco de trabajo WPD-CSP-FS, en la Figura 5.2 se puede observar como los resultados obtenidos por el mejor modelo con clasificador SVM superan a los obtenidos con los clasificadores KNN y LDA. Se mantiene la tendencia observada con OFB-CSP-FS, donde el caso de mayor desempeño fue el RT, pero en esta oportunidad el clasificador KNN solo logró superar ligeramente el desempeño de los demás clasificadores en el caso LT.

Durante las pruebas realizadas con la propuesta WPD-CSP-FS se observó que nuevamente el modelo que utilizó el logaritmo de la potencia promedio como característica

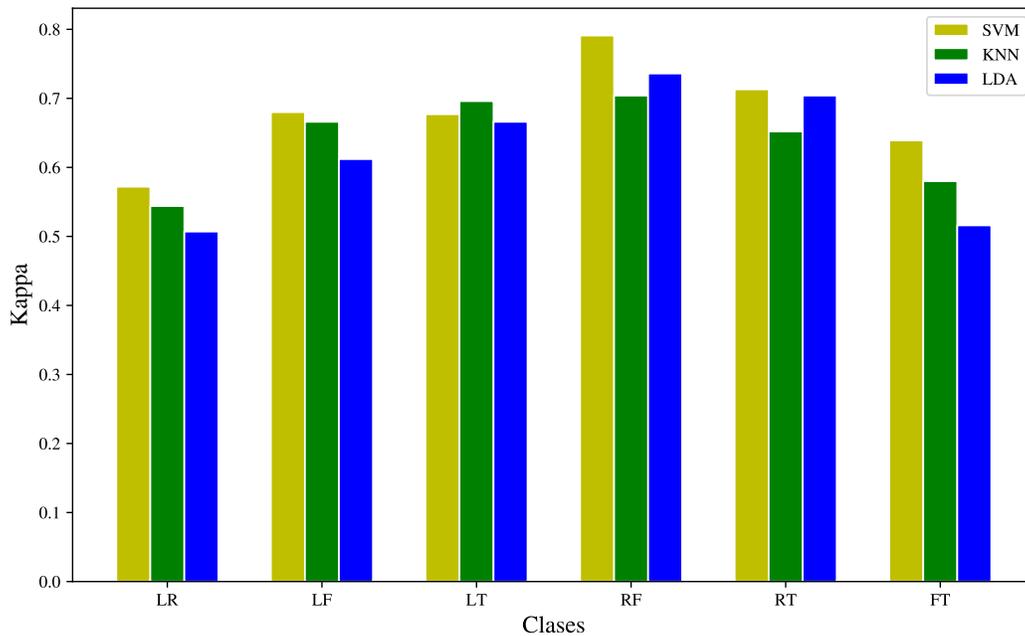


**Figura 5.1.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS.

**Tabla 5.7.** Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo OFB-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación binaria de clases.

No.	LR		LF		LT		RF		RT		FT	
	Kappa	Std										
A01	<b>0.856</b>	<b>0.020</b>	<b>0.933</b>	<b>0.006</b>	<b>0.972</b>	<b>0.0</b>	<b>0.947</b>	<b>0.006</b>	<b>0.972</b>	<b>0.0</b>	<b>0.738</b>	<b>0.026</b>
A02	0.05	0.019	0.736	0.015	0.461	0.022	0.839	0.023	0.386	0.01	0.736	0.036
A03	0.900	0.016	0.975	0.02	0.808	0.036	0.917	0.032	0.967	0.007	0.605	0.044
A04	0.408	0.031	0.606	0.044	0.853	0.014	0.892	0.006	0.817	0.024	0.480	0.039
A05	0.642	0.106	0.464	0.181	0.619	0.014	0.544	0.022	0.772	0.109	0.341	0.014
A06	<b>0.336</b>	<b>0.026</b>	<b>0.553</b>	<b>0.018</b>	<b>0.344</b>	<b>0.058</b>	<b>0.483</b>	<b>0.088</b>	<b>0.425</b>	<b>0.036</b>	<b>0.572</b>	<b>0.034</b>
A07	0.606	0.134	0.675	0.021	0.731	0.036	0.953	0.03	0.928	0.033	0.708	0.008
A08	0.925	0.006	0.881	0.042	0.772	0.017	0.856	0.017	0.706	0.014	0.666	0.030
A09	0.817	0.041	0.861	0.0	0.986	0.0	0.722	0.012	0.881	0.032	0.738	0.059
Prom.	0.615	0.017	0.743	0.039	0.727	0.022	0.795	0.026	0.761	0.029	0.621	0.033

en conjunto con un selector basado en la técnica de valor laplaciano fue el más exitoso cuando fue implementado con el clasificador SVM. En la Tabla 5.8 se describen los resultados del mejor modelo obtenido por cada sujeto de pruebas expresados en función del kappa promedio y su desviación estándar. El sujeto que obtuvo los mejores resultados de clasificación general fue el



**Figura 5.2.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS.

A03, el cual obtuvo su mejor resultado durante las pruebas realizadas en el caso LF, logrando un kappa de 0.961. Las señales del sujeto de prueba que presentaron mayor dificultad durante su clasificación fueron las de A06, obteniendo su mejor resultado en el caso FT con un kappa promedio de 0.594.

### 5.1.2. Clasificación Multiclases

A continuación se presentan los resultados obtenidos utilizando los clasificadores SVM, KNN y LDA, observando su influencia en cada una de las diferentes combinaciones multiclases junto con las técnicas de extracción y selección de características propuestas. Las métricas utilizadas para valorar el desempeño de los modelos fueron el índice kappa, la exactitud, el área bajo la curva (AUC) y el valor-F de cada una de las clases (F1). Los modelos con las mejores métricas para cada caso de estudio han sido resaltados con color azul.

**Tabla 5.8.** Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo WPD-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación binaria de clases.

No.	LR		LF		LT		RF		RT		FT	
Sujeto	Kappa	Std										
A01	0.828	0.026	0.944	0.0	0.933	0.006	0.961	0.006	0.972	0.0	0.630	0.031
A02	-0.064	0.034	0.683	0.027	0.339	0.031	0.858	0.018	0.281	0.042	0.813	0.041
<b>A03</b>	<b>0.833</b>	<b>0.045</b>	<b>0.961</b>	<b>0.014</b>	<b>0.842</b>	<b>0.014</b>	<b>0.928</b>	<b>0.034</b>	<b>0.939</b>	<b>0.011</b>	<b>0.777</b>	<b>0.00</b>
A04	0.239	0.065	0.492	0.017	0.794	0.059	0.817	0.024	0.703	0.042	0.472	0.042
A05	0.725	0.12	0.167	0.0	0.406	0.084	0.633	0.007	0.669	0.038	0.266	0.064
<b>A06</b>	<b>0.278</b>	<b>0.028</b>	<b>0.453</b>	<b>0.023</b>	<b>0.328</b>	<b>0.048</b>	<b>0.517</b>	<b>0.032</b>	<b>0.397</b>	<b>0.041</b>	<b>0.594</b>	<b>0.013</b>
A07	0.617	0.106	0.728	0.017	0.706	0.02	0.928	0.031	0.953	0.039	0.627	0.038
A08	0.906	0.01	0.869	0.03	0.778	0.012	0.847	0.03	0.672	0.007	0.663	0.018
A09	0.789	0.027	0.822	0.022	0.964	0.007	0.631	0.053	0.828	0.029	0.900	0.029
Prom.	0.572	0.051	0.68	0.017	0.677	0.031	0.791	0.026	0.713	0.028	0.639	0.031

### 5.1.2.1. Resultados con modelo SVM

Los resultados del marco de trabajo OFB-CSP-FS, con clasificador SVM sobre los casos multiclases, se muestran en la Tabla 5.9. En ella se observa que el modelo que obtuvo el mejor resultado global fue el LS con logpower con 0.661 de métrica kappa y 0.012 de desviación estándar. Este mismo modelo logró los mejores desempeños en la clasificación de los casos LRFT, LFT y LRF, y el modelo ITFE con logvar obtuvo mejores resultados que el resto en los casos RFT y LRT. Por otra parte, también es posible observar que la clase con el mejor desempeño de clasificación fue la clase R con un valor F1 promedio de 77.82 %, y el caso con mejor desempeño fue el LRF donde la clase F logró la mejor clasificación con un F1 promedio de 85.26 %.

Con respecto al marco de trabajo WPD-CSP-FS, los resultados están descritos en la Tabla 5.10, donde se observa que el modelo con mejor resultado global fue el ITFE con logvar logrando 0.623 de kappa y 0.019 de desviación estándar, aunque el modelo LS con logpower obtuvo un desempeño muy similar. De igual forma, este modelo logró los mejores desempeños en la clasificación de los casos RFT y LRF. El modelo LS con logpower mostró mejor desempeño en los casos LRFT y LRT, y el modelo LS con logvar obtuvo el mejor resultado de clasificación en el caso LFT. La clase con el mejor desempeño de clasificación fue la clase R con un F1 promedio de 76.04 %, y el caso con mejor desempeño fue el LRF donde la clase F obtuvo la mejor clasificación con un F1 promedio de 80.90 %.

**Tabla 5.9.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom $\pm$ Std
ITFE con logpower	Kappa	0.62	0.664	0.647	0.66	0.699	0.658 $\pm$ 0.021
	Exactitud [%]	71.5	77.57	76.43	77.37	79.96	76.57 $\pm$ 1.42
	AUC [%]	91.67	92.26	92.04	91.79	93.39	92.23 $\pm$ 0.37
	F1-Clase L [%]	70.84	—	80.71	74.49	77.1	75.78 $\pm$ 2.36
	F1-Clase R [%]	72.87	82.44	—	78.21	77.35	77.72 $\pm$ 1.67
	F1-Clase F [%]	69.89	76.38	74.37	—	83.57	76.05 $\pm$ 2.39
	F1-Clase T [%]	69.43	72.67	72.89	78.57	—	73.39 $\pm$ 1.96
ITFE con logvar	Kappa	0.619	<b>0.67</b>	0.643	<b>0.663</b>	0.7	0.659 $\pm$ 0.020
	Exactitud [%]	71.4	<b>77.97</b>	76.21	<b>77.55</b>	80.03	76.63 $\pm$ 1.34
	AUC [%]	91.39	<b>92.61</b>	91.76	<b>91.59</b>	93.29	92.13 $\pm$ 0.44
	F1-Clase L [%]	71.24	—	80.86	<b>75.07</b>	76.65	75.96 $\pm$ 2.11
	F1-Clase R [%]	71.81	<b>83.51</b>	—	<b>78.14</b>	77.15	77.65 $\pm$ 1.54
	F1-Clase F [%]	69.69	<b>75.2</b>	72.74	—	84.16	75.45 $\pm$ 2.54
	F1-Clase T [%]	69.32	<b>73.43</b>	73.5	<b>78.62</b>	—	73.72 $\pm$ 2.04
LS con logpower	Kappa	<b>0.629</b>	0.667	<b>0.662</b>	0.635	<b>0.71</b>	<b>0.661 <math>\pm</math> 0.012</b>
	Exactitud [%]	<b>72.18</b>	77.81	<b>77.45</b>	75.69	<b>80.69</b>	<b>76.76 <math>\pm</math> 0.82</b>
	AUC [%]	<b>91.64</b>	92.55	<b>92.45</b>	91.33	<b>93.53</b>	<b>92.3 <math>\pm</math> 0.28</b>
	F1-Clase L [%]	<b>72.23</b>	—	<b>81.98</b>	71.36	<b>78.16</b>	<b>75.93 <math>\pm</math> 1.35</b>
	F1-Clase R [%]	<b>74.22</b>	83.05	—	76.61	<b>77.42</b>	<b>77.82 <math>\pm</math> 1.15</b>
	F1-Clase F [%]	<b>70.82</b>	75.79	<b>75.55</b>	—	<b>85.26</b>	<b>76.86 <math>\pm</math> 1.51</b>
	F1-Clase T [%]	<b>69.25</b>	73.2	<b>73.97</b>	77.49	—	<b>73.48 <math>\pm</math> 1.08</b>
LS con logvar	Kappa	0.629	0.657	0.655	0.649	0.698	0.658 $\pm$ 0.020
	Exactitud [%]	72.17	77.1	77.0	76.62	79.87	76.55 $\pm$ 1.33
	AUC [%]	91.64	92.26	91.94	91.64	93.28	92.15 $\pm$ 0.40
	F1-Clase L [%]	72.68	—	81.6	73.08	77.52	76.22 $\pm$ 1.97
	F1-Clase R [%]	73.49	82.52	—	77.55	76.21	77.44 $\pm$ 1.76
	F1-Clase F [%]	70.47	75.28	75.3	—	84.1	76.29 $\pm$ 1.69
	F1-Clase T [%]	69.74	71.94	73.04	77.96	—	73.17 $\pm$ 1.83

### 5.1.2.2. Resultados con modelo KNN

En las pruebas realizadas con la propuesta OFB-CSP-FS utilizando el clasificador KNN, mostradas en la Tabla 5.11, se puede observar que el modelo ITFE con logpower obtuvo el mejor resultado global con 0.616 de métrica kappa y 0.013 de desviación estándar. Este mismo modelo logró los mejores desempeños en la clasificación en todos los casos de prueba exceptuando la combinación LFT, en la cual, el modelo que obtuvo el mejor desempeño fue el ITFE con logvar.

Los resultados con la propuesta WPD-CSP-FS presentes en la Tabla 5.12, también demuestran que el modelo ITFE con logpower obtuvo el mejor resultado global con 0.577 de métrica kappa y 0.017 de desviación estándar. El modelo ITFE con logpower logró los mejores

**Tabla 5.10.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom $\pm$ Std
ITFE con logpower	Kappa	0.595	0.628	0.601	0.625	0.649	0.62 $\pm$ 0.021
	Exactitud [%]	69.66	75.2	73.42	75.02	76.62	73.98 $\pm$ 1.43
	AUC [%]	89.93	90.82	90.27	89.63	91.69	90.47 $\pm$ 0.44
	F1-Clase L [%]	68.59	—	77.56	73.51	71.45	72.78 $\pm$ 2.06
	F1-Clase R [%]	72.65	81.6	—	75.85	75.62	76.43 $\pm$ 1.86
	F1-Clase F [%]	66.1	73.67	71.4	—	79.62	72.7 $\pm$ 2.67
	F1-Clase T [%]	66.85	67.83	69.59	74.67	—	69.74 $\pm$ 2.33
ITFE con logvar	Kappa	0.588	<b>0.636</b>	0.604	0.624	<b>0.663</b>	<b>0.623 <math>\pm</math> 0.019</b>
	Exactitud [%]	69.1	<b>75.73</b>	73.58	74.93	<b>77.55</b>	<b>74.18 <math>\pm</math> 1.31</b>
	AUC [%]	89.79	<b>90.5</b>	90.03	89.56	<b>91.64</b>	<b>90.3 <math>\pm</math> 0.43</b>
	F1-Clase L [%]	69.81	—	78.29	72.9	<b>73.61</b>	<b>73.65 <math>\pm</math> 2.06</b>
	F1-Clase R [%]	71.51	<b>81.69</b>	—	75.74	<b>75.2</b>	<b>76.04 <math>\pm</math> 1.77</b>
	F1-Clase F [%]	64.82	<b>73.91</b>	72.11	—	<b>80.9</b>	<b>72.94 <math>\pm</math> 2.48</b>
	F1-Clase T [%]	65.82	<b>69.66</b>	68.51	75.11	—	<b>69.78 <math>\pm</math> 2.11</b>
LS con logpower	Kappa	<b>0.603</b>	0.633	0.597	<b>0.632</b>	0.646	0.622 $\pm$ 0.023
	Exactitud [%]	<b>70.19</b>	75.5	73.14	<b>75.49</b>	76.42	74.15 $\pm$ 1.55
	AUC [%]	<b>89.9</b>	90.63	89.86	<b>89.75</b>	91.1	90.25 $\pm$ 0.5
	F1-Clase L [%]	<b>70.11</b>	—	77.94	<b>73.35</b>	71.59	73.25 $\pm$ 1.99
	F1-Clase R [%]	<b>71.24</b>	81.74	—	<b>75.8</b>	73.45	75.56 $\pm$ 1.87
	F1-Clase F [%]	<b>70.58</b>	73.07	72.53	—	80.67	74.21 $\pm$ 2.76
	F1-Clase T [%]	<b>66.43</b>	69.59	67.36	<b>75.98</b>	—	69.84 $\pm$ 2.17
LS con logvar	Kappa	0.598	0.632	<b>0.607</b>	0.616	0.654	0.621 $\pm$ 0.021
	Exactitud [%]	69.84	75.49	<b>73.8</b>	74.39	76.91	74.09 $\pm$ 1.39
	AUC [%]	89.76	90.78	<b>90.06</b>	89.72	91.12	90.29 $\pm$ 0.54
	F1-Clase L [%]	70.43	—	<b>78.83</b>	71.87	72.08	73.3 $\pm$ 2.13
	F1-Clase R [%]	70.6	81.55	—	74.14	74.96	75.31 $\pm$ 2.09
	F1-Clase F [%]	69.16	73.37	<b>72.85</b>	—	80.63	74.0 $\pm$ 2.33
	F1-Clase T [%]	66.7	69.07	<b>68.5</b>	75.39	—	69.91 $\pm$ 2.04

desempeños en la clasificación de los casos RFT y LRF, el modelo ITFE con logvar obtuvo los mejores desempeños en los casos LRFT y LFT, y finalmente el modelo LS con logvar obtuvo el mejor desempeño para el caso LRT. A diferencia de las pruebas anteriores, la clase L obtuvo la mejor clasificación con un F1 promedio de 72.68%, y el caso con mejores resultados de clasificación general fue el LRF con un kappa promedio de 0.6.

### 5.1.2.3. Resultados con modelo LDA

Los resultados del marco de trabajo OFB-CSP-FS, con clasificador LDA sobre los casos multiclases, se muestran en la Tabla 5.13. En ella se observa que el modelo que obtuvo el mejor

**Tabla 5.11.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom [ $\pm$ ]	Std
ITFE con logpower	Kappa	<b>0.574</b>	<b>0.615</b>	0.605	<b>0.634</b>	<b>0.651</b>	<b>0.616 <math>\pm</math> 0.013</b>	
	Exactitud [%]	<b>68.07</b>	<b>74.36</b>	73.63	<b>75.62</b>	<b>76.71</b>	<b>73.68 <math>\pm</math> 0.89</b>	
	AUC [%]	<b>89.15</b>	<b>90.31</b>	89.7	<b>89.12</b>	<b>91.29</b>	<b>89.91 <math>\pm</math> 0.37</b>	
	F1-Clase L [%]	<b>68.81</b>	—	78.39	<b>73.58</b>	<b>74.48</b>	<b>73.82 <math>\pm</math> 1.42</b>	
	F1-Clase R [%]	<b>69.82</b>	<b>79.21</b>	—	<b>75.0</b>	<b>73.98</b>	<b>74.5 <math>\pm</math> 1.23</b>	
	F1-Clase F [%]	<b>60.92</b>	<b>69.4</b>	64.75	—	<b>79.37</b>	<b>68.61 <math>\pm</math> 1.28</b>	
	F1-Clase T [%]	<b>66.98</b>	<b>71.72</b>	73.16	<b>77.51</b>	—	<b>72.34 <math>\pm</math> 1.08</b>	
ITFE con logvar	Kappa	0.572	0.614	<b>0.605</b>	0.627	0.65	0.614 $\pm$ 0.014	
	Exactitud [%]	67.91	74.3	<b>73.64</b>	75.12	76.69	73.53 $\pm$ 0.98	
	AUC [%]	89.05	90.25	<b>89.75</b>	88.95	91.15	89.83 $\pm$ 0.4	
	F1-Clase L [%]	68.74	—	<b>78.71</b>	73.45	74.55	73.86 $\pm$ 1.48	
	F1-Clase R [%]	69.65	79.31	—	74.13	73.99	74.27 $\pm$ 1.31	
	F1-Clase F [%]	60.3	69.65	<b>64.65</b>	—	79.22	68.46 $\pm$ 1.6	
	F1-Clase T [%]	67.16	71.36	<b>73.08</b>	76.89	—	72.12 $\pm$ 1.23	
LS con logpower	Kappa	0.545	0.603	0.583	0.615	0.622	0.594 $\pm$ 0.018	
	Exactitud [%]	65.89	73.49	72.17	74.31	74.79	72.13 $\pm$ 1.26	
	AUC [%]	88.2	89.22	89.0	88.25	90.29	88.99 $\pm$ 0.62	
	F1-Clase L [%]	67.68	—	77.04	71.85	71.78	72.09 $\pm$ 1.99	
	F1-Clase R [%]	66.76	78.9	—	72.37	70.7	72.18 $\pm$ 1.71	
	F1-Clase F [%]	59.7	68.21	67.03	—	79.47	68.6 $\pm$ 2.11	
	F1-Clase T [%]	64.37	70.48	70.17	77.3	—	70.58 $\pm$ 1.62	
LS con logvar	Kappa	0.54	0.604	0.576	0.622	0.635	0.595 $\pm$ 0.022	
	Exactitud [%]	65.53	73.61	71.76	74.8	75.67	72.27 $\pm$ 1.48	
	AUC [%]	88.06	89.22	88.37	88.51	90.61	88.95 $\pm$ 0.65	
	F1-Clase L [%]	67.37	—	76.96	72.42	73.56	72.58 $\pm$ 2.21	
	F1-Clase R [%]	65.74	79.16	—	73.15	71.88	72.48 $\pm$ 1.94	
	F1-Clase F [%]	59.75	67.6	66.33	—	79.05	68.18 $\pm$ 2.45	
	F1-Clase T [%]	64.19	71.03	69.57	77.5	—	70.57 $\pm$ 1.71	

resultado global fue el ITFE con logvar logrando 0.545 de métrica kappa y 0.037 de desviación estándar. El modelo ITFE con logvar obtuvo los mejores desempeños en la clasificación de los casos LFT, LRT y LRF, mientras que el modelo ITFE con logpower logró mejores resultados en los casos LRFT y RFT. Por otra parte, también es posible observar que la clase con el mejor desempeño de clasificación nuevamente fue la clase R con un F1 promedio de 69.94 %, y el caso con mejor desempeño fue el LRF donde la clase F obtuvo la mejor clasificación con un F1 promedio de 74.24 %.

Por otro lado, los resultados obtenidos por el marco de trabajo WPD-CSP-FS, descritos en la Tabla 5.14, indican que el mejor modelo de clasificación fue el LS con logpower obteniendo una métrica kappa general de 0.593 con desviación estándar de 0.024. WPD-CSP-FS describe

**Tabla 5.12.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom $\pm$ Std
ITFE con logpower	Kappa	0.535	<b>0.584</b>	0.576	0.591	<b>0.6</b>	<b>0.577 <math>\pm</math> 0.017</b>
	Exactitud [%]	65.17	<b>72.27</b>	71.73	72.72	<b>73.36</b>	<b>71.05 <math>\pm</math> 1.15</b>
	AUC [%]	87.49	<b>87.82</b>	88.61	87.5	<b>89.77</b>	<b>88.24 <math>\pm</math> 0.46</b>
	F1-Clase L [%]	67.57	—	78.04	71.84	<b>73.25</b>	<b>72.68 <math>\pm</math> 1.37</b>
	F1-Clase R [%]	65.15	<b>76.63</b>	—	70.87	<b>69.09</b>	<b>70.44 <math>\pm</math> 1.94</b>
	F1-Clase F [%]	57.24	<b>69.33</b>	63.08	—	<b>74.25</b>	<b>65.97 <math>\pm</math> 1.6</b>
	F1-Clase T [%]	64.67	<b>68.36</b>	69.42	74.37	—	<b>69.2 <math>\pm</math> 1.76</b>
ITFE con logvar	Kappa	<b>0.541</b>	0.569	<b>0.583</b>	0.593	0.599	0.577 $\pm$ 0.015
	Exactitud [%]	<b>65.56</b>	71.3	<b>72.17</b>	72.85	73.25	71.03 $\pm$ 1.04
	AUC [%]	<b>87.63</b>	87.59	<b>88.75</b>	87.45	89.68	88.22 $\pm$ 0.44
	F1-Clase L [%]	<b>67.99</b>	—	<b>78.1</b>	71.92	73.18	72.8 $\pm$ 1.23
	F1-Clase R [%]	<b>66.3</b>	75.7	—	71.63	68.49	70.53 $\pm$ 1.52
	F1-Clase F [%]	<b>57.5</b>	68.21	<b>63.6</b>	—	74.32	65.91 $\pm$ 1.77
	F1-Clase T [%]	<b>64.21</b>	67.43	<b>70.47</b>	73.82	—	68.98 $\pm$ 1.76
LS con logpower	Kappa	0.531	0.569	0.558	0.596	0.582	0.567 $\pm$ 0.024
	Exactitud [%]	64.82	71.29	70.53	73.06	72.12	70.36 $\pm$ 1.63
	AUC [%]	87.08	87.36	87.36	87.39	88.18	87.47 $\pm$ 0.75
	F1-Clase L [%]	67.6	—	76.74	72.38	72.1	72.2 $\pm$ 2.13
	F1-Clase R [%]	65.2	77.77	—	70.71	65.67	69.84 $\pm$ 2.56
	F1-Clase F [%]	59.58	67.2	65.65	—	75.42	66.96 $\pm$ 2.66
	F1-Clase T [%]	62.2	65.01	66.87	74.77	—	67.21 $\pm$ 2.07
LS con logvar	Kappa	0.532	0.572	0.57	<b>0.601</b>	0.583	0.572 $\pm$ 0.023
	Exactitud [%]	64.92	71.46	71.31	<b>73.37</b>	72.18	70.65 $\pm$ 1.54
	AUC [%]	86.93	87.52	87.75	<b>87.69</b>	88.48	87.67 $\pm$ 0.69
	F1-Clase L [%]	67.72	—	77.96	<b>72.72</b>	71.89	72.57 $\pm$ 2.34
	F1-Clase R [%]	65.19	77.18	—	<b>70.99</b>	67.43	70.2 $\pm$ 2.14
	F1-Clase F [%]	59.58	67.74	65.07	—	74.68	66.77 $\pm$ 2.72
	F1-Clase T [%]	62.65	65.6	68.07	<b>75.22</b>	—	67.88 $\pm$ 2.11

resultados con tendencias diferentes a los obtenidos por OFB-CSP-FS. Durante estas pruebas, los mejores desempeños de clasificación generales se obtuvieron en el caso LRT con un kappa promedio de 0.617. El modelo LS con logpower obtuvo los mejores resultados de clasificación en los casos LRFT y LRF, mientras que el modelo LS con logvar superó a los demás modelos en la clasificación de los casos RFT y LRT.

#### 5.1.2.4. Análisis de Resultados

De acuerdo con los resultados obtenidos con OFB-CSP-FS empleados en la clasificación multiclases, el desempeño de los mejores modelos con cada clasificador se observan en la

**Tabla 5.13.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom $\pm$ Std
ITFE con logpower	Kappa	<b>0.551</b>	<b>0.559</b>	0.49	0.562	0.551	0.543 $\pm$ 0.038
	Exactitud [%]	<b>66.31</b>	<b>70.61</b>	65.98	70.8	70.08	68.76 $\pm$ 2.6
	AUC [%]	<b>88.22</b>	<b>87.06</b>	85.73	85.7	85.43	86.43 $\pm$ 1.96
	F1-Clase L [%]	<b>62.37</b>	—	70.33	68.82	66.95	67.12 $\pm$ 4.38
	F1-Clase R [%]	<b>68.6</b>	<b>75.29</b>	—	69.2	64.57	69.41 $\pm$ 3.47
	F1-Clase F [%]	<b>65.51</b>	<b>65.89</b>	59.95	—	73.18	66.13 $\pm$ 2.67
	F1-Clase T [%]	<b>65.37</b>	<b>68.55</b>	61.56	72.83	—	67.08 $\pm$ 3.12
ITFE con logvar	Kappa	0.542	0.555	<b>0.499</b>	<b>0.562</b>	<b>0.567</b>	<b>0.545 <math>\pm</math> 0.037</b>
	Exactitud [%]	65.65	70.32	<b>66.63</b>	<b>70.8</b>	<b>71.14</b>	<b>68.91 <math>\pm</math> 2.55</b>
	AUC [%]	87.48	86.89	<b>86.26</b>	<b>85.92</b>	<b>86.25</b>	<b>86.56 <math>\pm</math> 1.98</b>
	F1-Clase L [%]	61.1	—	<b>70.44</b>	<b>68.9</b>	<b>68.1</b>	<b>67.13 <math>\pm</math> 3.48</b>
	F1-Clase R [%]	68.54	74.72	—	<b>69.78</b>	<b>66.7</b>	<b>69.94 <math>\pm</math> 3.39</b>
	F1-Clase F [%]	65.62	66.45	<b>60.71</b>	—	<b>74.24</b>	<b>66.76 <math>\pm</math> 2.74</b>
	F1-Clase T [%]	63.6	67.61	<b>62.13</b>	<b>72.45</b>	—	<b>66.45 <math>\pm</math> 3.24</b>
LS con logpower	Kappa	0.497	0.509	0.464	0.553	0.528	0.51 $\pm$ 0.058
	Exactitud [%]	62.24	67.27	64.24	70.23	68.57	66.51 $\pm$ 3.95
	AUC [%]	83.86	84.52	82.39	84.92	84.27	83.99 $\pm$ 2.84
	F1-Clase L [%]	61.48	—	68.93	68.28	65.63	66.08 $\pm$ 5.81
	F1-Clase R [%]	62.26	70.62	—	66.45	63.53	65.72 $\pm$ 5.54
	F1-Clase F [%]	59.55	65.28	61.28	—	70.82	64.23 $\pm$ 5.52
	F1-Clase T [%]	60.33	62.05	59.12	73.42	—	63.73 $\pm$ 5.27
LS con logvar	Kappa	0.477	0.53	0.462	0.552	0.545	0.513 $\pm$ 0.059
	Exactitud [%]	60.8	68.68	64.15	70.12	69.69	66.69 $\pm$ 4.03
	AUC [%]	82.37	84.81	82.84	84.55	84.43	83.8 $\pm$ 3.3
	F1-Clase L [%]	59.93	—	68.47	68.9	67.91	66.3 $\pm$ 6.3
	F1-Clase R [%]	61.55	73.38	—	65.02	64.87	66.2 $\pm$ 5.33
	F1-Clase F [%]	58.2	66.98	61.37	—	72.33	64.72 $\pm$ 5.1
	F1-Clase T [%]	58.51	63.45	58.04	73.74	—	63.44 $\pm$ 5.04

Figura 5.3. Es notorio que el mejor modelo con el clasificador SVM superó al resto en todos los casos de combinación multiclases.

En esta oportunidad, al igual que durante la clasificación binaria de clases, el mejor modelo de la propuesta OFB-CSP-FS se logró utilizando el logaritmo de la potencia promedio como característica y un selector basado en el cálculo del valor laplaciano. Este modelo, implementado en conjunto con un clasificador SVM, obtuvo su máximo desempeño en el caso LRF con un kappa promedio de 0.71.

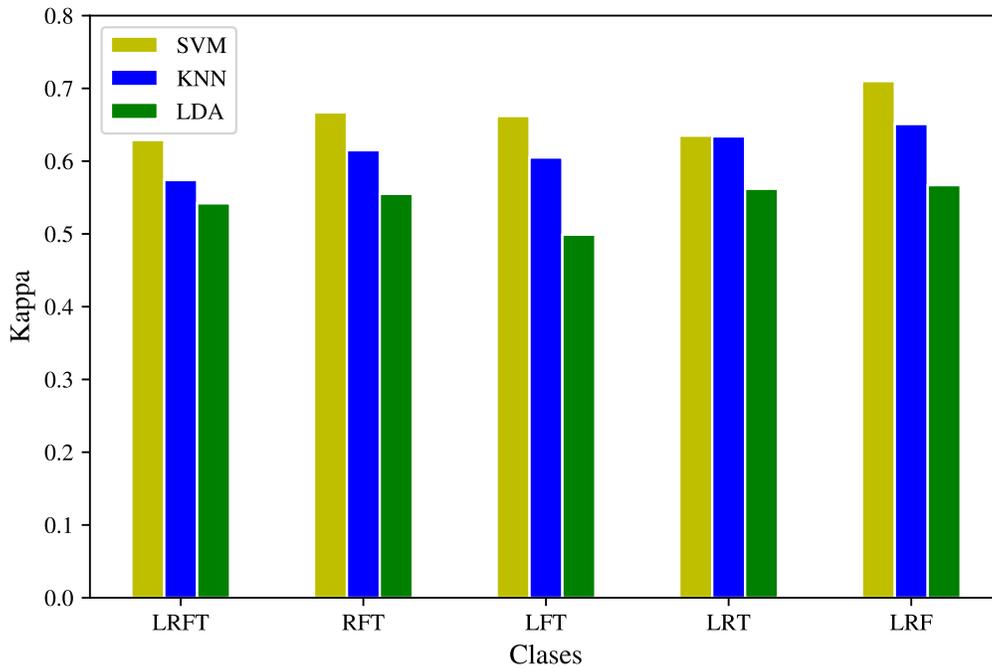
En la Tabla 5.15 se describe detalladamente los resultados obtenidos por el mejor modelo utilizando la propuesta OFB-CSP-FS sobre cada uno de los sujetos de prueba. En ella se observa

**Tabla 5.14.** Métricas de desempeño en clasificación de datos de prueba, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas con combinación multiclases.

Combinación	Métrica	LRFT	RFT	LFT	LRT	LRF	Prom $\pm$ Std
ITFE con logpower	Kappa	0.557	0.6	<b>0.59</b>	0.587	0.614	0.59 $\pm$ 0.014
	Exactitud [%]	66.81	73.34	<b>72.68</b>	72.5	74.24	71.91 $\pm$ 0.96
	AUC [%]	88.05	89.34	<b>89.73</b>	87.64	89.14	88.78 $\pm$ 0.48
	F1-Clase L [%]	67.42	—	<b>79.21</b>	70.27	69.56	71.61 $\pm$ 1.03
	F1-Clase R [%]	68.79	78.42	—	73.61	74.05	73.72 $\pm$ 1.49
	F1-Clase F [%]	61.9	72.54	<b>64.9</b>	—	75.3	68.66 $\pm$ 1.72
	F1-Clase T [%]	64.7	67.95	<b>69.64</b>	72.63	—	68.73 $\pm$ 1.89
ITFE con logvar	Kappa	0.553	0.595	0.573	0.584	0.597	0.58 $\pm$ 0.017
	Exactitud [%]	66.44	72.99	71.55	72.28	73.1	71.27 $\pm$ 1.19
	AUC [%]	87.84	89.07	89.21	87.67	88.91	88.54 $\pm$ 0.58
	F1-Clase L [%]	68.48	—	78.8	69.91	69.41	71.65 $\pm$ 1.52
	F1-Clase R [%]	68.16	77.81	—	73.25	71.83	72.76 $\pm$ 1.67
	F1-Clase F [%]	61.99	71.47	62.53	—	74.53	67.63 $\pm$ 2.3
	F1-Clase T [%]	63.44	68.56	68.87	72.71	—	68.4 $\pm$ 1.65
LS con logpower	Kappa	<b>0.576</b>	0.597	0.565	0.613	<b>0.615</b>	<b>0.593 <math>\pm</math> 0.024</b>
	Exactitud [%]	<b>68.17</b>	73.12	70.99	74.22	<b>74.34</b>	<b>72.17 <math>\pm</math> 1.66</b>
	AUC [%]	<b>88.33</b>	89.28	87.83	88.73	<b>88.6</b>	<b>88.55 <math>\pm</math> 0.98</b>
	F1-Clase L [%]	<b>70.5</b>	—	76.28	72.89	<b>68.88</b>	<b>72.14 <math>\pm</math> 2.93</b>
	F1-Clase R [%]	<b>68.62</b>	78.88	—	74.94	<b>74.42</b>	<b>74.22 <math>\pm</math> 2.12</b>
	F1-Clase F [%]	<b>63.71</b>	70.32	67.04	—	<b>75.45</b>	<b>69.13 <math>\pm</math> 3.47</b>
	F1-Clase T [%]	<b>65.66</b>	67.37	66.76	73.15	—	<b>68.24 <math>\pm</math> 3.22</b>
LS con logvar	Kappa	0.567	<b>0.601</b>	0.57	<b>0.617</b>	0.608	0.593 $\pm$ 0.023
	Exactitud [%]	67.51	<b>73.43</b>	71.31	<b>74.47</b>	73.84	72.11 $\pm$ 1.56
	AUC [%]	87.96	<b>89.34</b>	88.12	<b>88.49</b>	88.27	88.44 $\pm$ 0.95
	F1-Clase L [%]	68.97	—	74.78	<b>72.99</b>	67.96	71.18 $\pm$ 2.99
	F1-Clase R [%]	67.78	<b>79.64</b>	—	<b>74.83</b>	73.84	74.02 $\pm$ 2.13
	F1-Clase F [%]	63.58	<b>70.13</b>	67.65	—	75.2	69.14 $\pm$ 3.45
	F1-Clase T [%]	65.23	<b>67.91</b>	69.37	<b>74.06</b>	—	69.14 $\pm$ 2.45

que el sujeto *A01*, resaltado en color azul, presentó el mejor desempeño general durante todas las pruebas con casos multi-clases, donde el mejor resultado se presentó en el caso LRT. Los peores resultados de clasificación se obtuvieron con el sujeto *A06*, resaltado en color rojo, logrando su mejor resultado en el caso LFT con un kappa de 0.493.

Visualizando el rendimiento de los clasificadores con la propuesta WPD-CSP-FS, descritos en la Figura 5.4, se puede ver que en todas las pruebas los resultados son peores que los obtenidos durante las pruebas con la propuesta OFB-CSP-FS. Igualmente en esta oportunidad el desempeño clasificador SVM superó el desempeño de los demás clasificadores, siendo el caso LRF el que presentó el mejor resultado promedio.

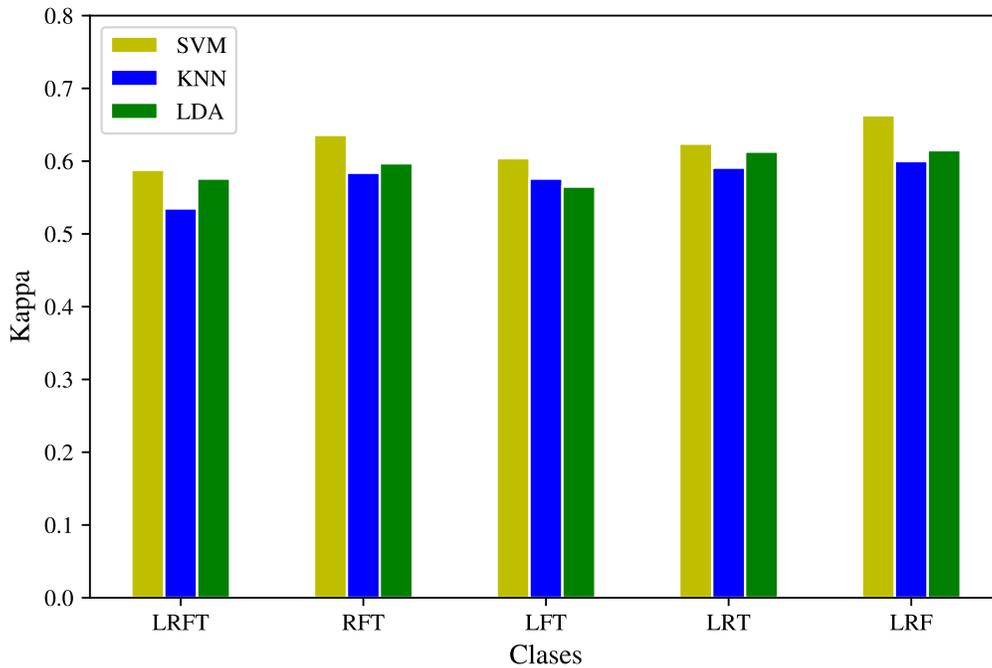


**Figura 5.3.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS en combinación multi-clases.

**Tabla 5.15.** Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo OFB-CSP-FS con Valor Laplaciano (LS) como selector de característica, el logaritmo de potencia promedio como característica (logpower) y clasificador SVM con combinación multi-clases.

No.	LRFT		RFT		LFT		LRT		LRF	
	Kappa	Std								
A01	<b>0.821</b>	<b>0.009</b>	<b>0.854</b>	<b>0.023</b>	<b>0.812</b>	<b>0.009</b>	<b>0.917</b>	<b>0.004</b>	<b>0.903</b>	<b>0.004</b>
A02	0.408	0.006	0.569	0.015	0.599	0.01	0.165	0.007	0.46	0.045
A03	0.749	0.001	0.8	0.01	0.703	0.005	0.854	0.004	0.896	0.008
A04	0.597	0.006	0.703	0.007	0.601	0.01	0.494	0.008	0.667	0.008
A05	0.523	0.012	0.519	0.013	0.444	0.015	0.71	0.011	0.592	0.012
A06	<b>0.330</b>	<b>0.011</b>	<b>0.453</b>	<b>0.009</b>	<b>0.493</b>	<b>0.012</b>	<b>0.346</b>	<b>0.005</b>	<b>0.464</b>	<b>0.016</b>
A07	0.698	0.016	0.711	0.017	0.685	0.016	0.656	0.049	0.814	0.036
A08	0.760	0.008	0.664	0.006	0.735	0.017	0.765	0.005	0.854	0.012
A09	0.774	0.009	0.731	0.012	0.883	0.01	0.811	0.003	0.744	0.011
Prom.	0.629	0.009	0.667	0.012	0.662	0.012	0.635	0.011	0.71	0.017

El mejor modelo obtenido utilizando la propuesta WPD-CSP-FS con multi-clases corresponde al que implementó el cálculo del logaritmo de la varianza de la señal como



**Figura 5.4.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS en combinación multi-clases.

característica y el algoritmo de información mutua como selector. Los resultados obtenidos se muestran en la Tabla 5.16, donde se puede observar que el sujeto *A01* logró los mejores resultados de clasificación general, siendo el caso LRT su mejor resultado. Por otro lado, el sujeto *A06* presentó el peor desempeño en clasificación con un resultado máximo de 0.393 de kappa promedio en los casos LRFT y RFT.

## 5.2. Resultados con base de datos BCI Competition IV 2b

De acuerdo con [Tangermann et al. \[2012\]](#), esta segunda base de datos denominada *transferencia de sesión a sesión de intenciones motoras para Interfaces Cerebro-Computadora bajo presencia de artefactos oculares* plantea como principal reto la clasificación de señales EEG afectadas por artefactos del movimiento ocular, en conjunto con la capacidad de adaptación de los algoritmos entre una sesión de adquisición de datos a otra, debido a que los conjuntos de datos de entrenamiento y prueba fueron registrados en cinco días diferentes.

**Tabla 5.16.** Resultados de kappa promedio por cada sujeto en prueba utilizando la mejor propuesta del marco de trabajo WPD-CSP-FS con información mutua (ITFE) como selector de característica, el logaritmo de la varianza promedio como característica (logvar) y clasificador SVM con combinación multi-clases.

No.	LRFT		RFT		LFT		LRT		LRF	
	Kappa	Std								
<b>A01</b>	<b>0.778</b>	<b>0.014</b>	<b>0.778</b>	<b>0.014</b>	<b>0.761</b>	<b>0.014</b>	<b>0.899</b>	<b>0.02</b>	<b>0.868</b>	<b>0.034</b>
A02	0.551	0.024	0.551	0.024	0.478	0.008	0.168	0.02	0.476	0.018
A03	0.812	0.0	0.812	0.0	0.736	0.014	0.81	0.008	0.875	0.015
A04	0.599	0.029	0.599	0.029	0.518	0.027	0.532	0.029	0.508	0.031
A05	0.453	0.03	0.453	0.03	0.374	0.044	0.681	0.0	0.562	0.0
<b>A06</b>	<b>0.393</b>	<b>0.018</b>	<b>0.393</b>	<b>0.018</b>	<b>0.306</b>	<b>0.053</b>	<b>0.242</b>	<b>0.029</b>	<b>0.332</b>	<b>0.017</b>
A07	0.782	0.01	0.782	0.01	0.719	0.007	0.733	0.033	0.797	0.051
A08	0.657	0.008	0.657	0.008	0.725	0.008	0.753	0.011	0.836	0.003
A09	0.699	0.033	0.699	0.033	0.817	0.025	0.799	0.009	0.714	0.01
Prom.	0.588	0.017	0.636	0.018	0.604	0.022	0.624	0.018	0.663	0.02

A diferencia de la base de datos BCI Competition IV 2a, en esta oportunidad durante el proceso de entrenamiento de los modelos se utilizó una partición de 120 muestras sin retroalimentación visual y 80 muestras con retroalimentación visual por clase, y para la prueba de los modelos se utilizó otra partición de 160 muestras con retroalimentación visual por clase. Ambas particiones de datos fueron adquiridas en sesiones realizadas en cinco días diferentes con nueve sujetos de prueba distintos, y están previamente definidas por la BCI Competition IV. Adicionalmente, la base de datos contempla un total de dos clases, determinadas por pensamientos de movilidad en brazo izquierdo (L) y brazo derecho (R).

A continuación se presentan los resultados obtenidos utilizando los marcos de trabajo propuestos OFB-CSP-FS y WPD-CSP-FS en conjunto con clasificadores SVM, KNN y LDA, observando su influencia en la clasificación junto con las técnicas de extracción y selección de características propuestas. Las métricas utilizadas para valorar el desempeño de los modelos fueron el índice kappa, la exactitud, el área bajo la curva (AUC), la precisión y la exhaustividad (Recall). Los resultados expuestos corresponden a los valores promedio, obtenidos después de realizar cinco ejecuciones de cada modelo clasificando la porción de datos de prueba, donde la selección y validación de las mejores características se realizó con datos escogidos aleatoriamente de la porción de datos de entrenamiento. Los modelos con las mejores métricas para cada caso de estudio han sido resaltados con color azul.

### 5.2.1. Resultados con Clasificador SVM

Los resultados obtenidos utilizando OFB-CSP-FS con el clasificador SVM se describen en la Tabla 5.17. Es posible observar que el modelo que utilizó ITFE con logvar obtuvo los mejores resultados generales con un kappa promedio de 0.565 y una desviación estándar de 0.264. Verificando cada sujeto de pruebas por separado, es notorio que el modelo de ITFE con logvar obtuvo el mejor desempeño durante la clasificación de los sujetos B01, B02, B05 y B08; el modelo ITFE con logpower obtuvo el mejor desempeño de clasificación para los sujetos B06 y B09; el modelo LS con logpower obtuvo el mejor desempeño en la clasificación los sujetos B03 y B04, y finalmente el modelo LS con logvar sólo superó al resto durante la clasificación del sujeto B07.

**Tabla 5.17.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas.

Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	0.353	0.237	0.102	0.901	0.809	<b>0.669</b>	0.494	0.802	<b>0.694</b>	0.562 $\pm$ 0.264
	Exactitud [%]	67.62	61.86	55.12	95.06	90.44	<b>83.44</b>	74.69	90.13	<b>84.69</b>	78.12 $\pm$ 13.21
	AUC [%]	87.43	63.11	59.83	99.35	97.82	<b>94.86</b>	85.14	97.03	<b>94.55</b>	86.57 $\pm$ 14.14
	Precision [%]	93.19	61.39	54.65	94.56	97.63	<b>75.85</b>	75.48	86.31	<b>88.81</b>	80.87 $\pm$ 14.27
	Recall [%]	38.0	64.0	60.38	95.62	82.88	<b>98.12</b>	73.12	95.38	<b>79.38</b>	76.32 $\pm$ 18.7
ITFE con logvar	Kappa	<b>0.375</b>	<b>0.243</b>	0.095	0.901	<b>0.82</b>	0.661	0.494	<b>0.806</b>	0.686	<b>0.565<math>\pm</math>0.264</b>
	Exactitud [%]	<b>68.75</b>	<b>62.14</b>	54.75	95.06	<b>91.0</b>	83.06	74.69	<b>90.31</b>	84.31	<b>78.23<math>\pm</math>13.19</b>
	AUC [%]	<b>88.29</b>	<b>63.43</b>	59.61	99.35	<b>98.03</b>	94.85	85.14	<b>97.06</b>	94.54	<b>86.7<math>\pm</math>14.16</b>
	Precision [%]	<b>94.12</b>	<b>61.97</b>	54.23	94.56	<b>98.09</b>	75.56	75.48	<b>86.44</b>	89.39	<b>81.09<math>\pm</math>14.48</b>
	Recall [%]	<b>40.0</b>	<b>62.86</b>	61.12	95.62	<b>83.62</b>	97.75	73.12	<b>95.62</b>	77.88	<b>76.4<math>\pm</math>18.26</b>
LS con logpower	Kappa	0.261	0.184	<b>0.106</b>	<b>0.919</b>	0.792	0.548	0.523	0.793	0.676	0.534 $\pm$ 0.275
	Exactitud [%]	63.06	59.21	<b>55.31</b>	<b>95.94</b>	89.62	77.38	76.12	89.62	83.81	76.67 $\pm$ 13.76
	AUC [%]	85.18	62.04	<b>58.99</b>	<b>99.42</b>	97.44	89.96	84.88	97.06	93.12	85.34 $\pm$ 14.15
	Precision [%]	89.43	58.27	<b>54.93</b>	<b>96.23</b>	98.33	70.28	77.21	86.53	86.12	79.7 $\pm$ 14.8
	Recall [%]	29.62	65.0	<b>59.13</b>	<b>95.62</b>	80.62	94.88	74.13	93.88	80.62	74.83 $\pm$ 20.14
LS con logvar	Kappa	0.258	0.171	0.102	0.919	0.8	0.548	<b>0.534</b>	0.795	0.672	0.533 $\pm$ 0.279
	Exactitud [%]	62.87	58.57	55.12	95.94	90.0	77.38	<b>76.69</b>	89.75	83.62	76.66 $\pm$ 13.95
	AUC [%]	85.14	62.21	58.9	99.36	97.61	89.96	<b>85.25</b>	97.1	92.64	85.35 $\pm$ 14.12
	Precision [%]	89.3	57.72	54.72	96.0	98.2	70.28	<b>78.6</b>	86.98	85.88	79.74 $\pm$ 14.86
	Recall [%]	29.25	64.14	59.38	95.88	81.5	94.88	<b>73.38</b>	93.5	80.5	74.71 $\pm$ 20.28

En cuanto a los resultados obtenidos utilizando WPD-CSP-FS con el clasificador SVM, descritos en la Tabla 5.18, puede observarse que nuevamente el modelo ITFE con logvar obtuvo los mejores resultados generales con un kappa promedio de 0.541 y una desviación estándar de 0.269. El mismo modelo de ITFE con logvar obtuvo mejores resultados que el resto durante la clasificación de los sujetos B05, B06, B07 y B09; el modelo LS con logpower obtuvo los mejores resultados para los sujetos B01 y B02; el modelo LS con logvar obtuvo el mejor desempeño en la clasificación los sujetos B03 y B08, y el modelo ITFE con logpower sólo superó al resto

durante la clasificación del sujeto B04.

Aunque todos los sujetos utilizados en las pruebas eran diestros, tanto en los resultados obtenidos con OFB-CSP-FS como con WPD-CSP-FS se observa la tendencia a favorecer la clasificación de pensamientos de brazo izquierdo (L) sobre los pensamientos de brazo derecho (R), logrando una precisión general mayor a la exhaustividad. Esta tendencia también puede observarse en los resultados de cada sujeto de pruebas, donde los sujetos B01, B04, B05, y B07, lograron mejor clasificación para la clase L.

**Tabla 5.18.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador SVM con las diferentes técnicas de selección y extracción de características propuestas.

Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	0.376	0.174	0.094	<b>0.879</b>	0.798	0.596	0.431	0.762	0.746	0.54 $\pm$ 0.268
	Exactitud [%]	68.81	58.71	54.69	<b>93.94</b>	89.88	79.81	71.56	88.12	87.31	76.98 $\pm$ 13.41
	AUC [%]	83.52	61.95	58.15	<b>98.23</b>	97.06	92.25	81.09	96.21	94.97	84.83 $\pm$ 14.4
	Precision [%]	87.13	59.1	54.72	<b>95.36</b>	94.93	72.53	73.79	83.52	85.94	78.56 $\pm$ 13.78
	Recall [%]	44.12	56.57	54.37	<b>92.37</b>	84.25	96.0	66.88	95.0	89.25	75.42 $\pm$ 18.9
ITFE con logvar	Kappa	0.387	0.166	0.09	0.878	<b>0.798</b>	<b>0.596</b>	<b>0.444</b>	0.762	<b>0.749</b>	<b>0.541<math>\pm</math>0.269</b>
	Exactitud [%]	69.38	58.29	54.5	93.88	<b>89.88</b>	<b>79.81</b>	<b>72.19</b>	88.12	<b>87.44</b>	<b>77.05<math>\pm</math>13.45</b>
	AUC [%]	83.82	61.81	58.4	98.21	<b>97.07</b>	<b>92.25</b>	<b>81.87</b>	96.21	<b>94.82</b>	<b>84.94<math>\pm</math>14.34</b>
	Precision [%]	87.8	58.78	54.56	95.59	<b>94.93</b>	<b>72.53</b>	<b>74.48</b>	83.52	<b>86.67</b>	<b>78.76<math>\pm</math>13.96</b>
	Recall [%]	45.0	55.43	53.75	92.0	<b>84.25</b>	<b>96.0</b>	<b>67.5</b>	95.0	<b>88.5</b>	<b>75.27<math>\pm</math>18.82</b>
LS con logpower	Kappa	<b>0.388</b>	<b>0.176</b>	0.1	0.869	0.782	0.534	0.401	0.77	0.736	0.528 $\pm$ 0.264
	Exactitud [%]	<b>69.38</b>	<b>58.79</b>	55.0	93.44	89.12	76.69	70.06	88.5	86.81	76.42 $\pm$ 13.18
	AUC [%]	<b>83.87</b>	<b>62.06</b>	58.04	98.41	96.79	91.33	78.99	96.24	95.19	84.55 $\pm$ 14.44
	Precision [%]	<b>87.44</b>	<b>59.09</b>	55.38	94.29	94.84	69.44	72.26	84.78	86.77	78.25 $\pm$ 13.89
	Recall [%]	<b>45.25</b>	<b>57.14</b>	51.38	92.5	82.75	95.38	65.12	93.88	86.88	74.48 $\pm$ 18.67
LS con logvar	Kappa	0.374	0.173	<b>0.104</b>	0.869	0.781	0.528	0.4	<b>0.77</b>	0.729	0.525 $\pm$ 0.264
	Exactitud [%]	68.69	58.64	<b>55.19</b>	93.44	89.06	76.38	70.0	<b>88.5</b>	86.44	76.26 $\pm$ 13.18
	AUC [ $\pm$ ]	83.78	62.06	<b>58.08</b>	98.39	96.74	91.55	79.46	<b>96.3</b>	95.0	84.6 $\pm$ 14.41
	Precision [%]	86.22	58.91	<b>55.6</b>	94.17	94.2	68.91	71.99	<b>85.1</b>	85.8	77.88 $\pm$ 13.69
	Recall [%]	44.5	57.14	<b>51.62</b>	92.63	83.25	96.12	65.5	<b>93.38</b>	87.38	74.61 $\pm$ 18.86

## 5.2.2. Resultados con Clasificador KNN

Así como se describe en la Tabla 5.19, el marco de trabajo OFB-CSP-FS con clasificador KNN utilizando ITFE con logpower, obtuvo los mejores resultados generales con un kappa promedio de 0.535 y una desviación estándar de 0.263. En este caso, el modelo LS con logpower obtuvo el mejor desempeño clasificando las señales de los sujetos B01, B04, B05 y B07; el modelo ITFE con logpower obtuvo los mejores resultados en los sujetos B02, B03 y B09; y el modelo LS con logvar logró un mejor desempeño con los sujetos B06 y B08.

En los resultados de nuevo se puede evidenciar una tendencia a clasificar mejor las

señales correspondientes a la clase L. Aunque clase R logró resultados superiores en los sujetos B03, B6 y B8.

**Tabla 5.19.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas.

Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	0.279	<b>0.143</b>	<b>0.191</b>	0.894	0.75	0.568	0.475	0.802	<b>0.714</b>	<b>0.535<math>\pm</math>0.263</b>
	Exactitud [%]	63.94	<b>57.14</b>	<b>59.56</b>	94.69	87.5	78.38	73.75	90.12	<b>85.69</b>	<b>76.75<math>\pm</math>13.13</b>
	AUC [%]	81.55	<b>61.81</b>	<b>59.34</b>	99.45	95.59	91.62	82.87	97.24	<b>94.99</b>	<b>84.94<math>\pm</math>14.25</b>
	Precision [%]	89.21	<b>57.24</b>	<b>58.45</b>	93.87	92.86	71.48	80.78	89.24	<b>89.98</b>	<b>80.35<math>\pm</math>13.67</b>
	Recall [%]	31.75	<b>57.29</b>	<b>66.0</b>	95.62	81.25	94.5	62.38	91.25	<b>80.38</b>	<b>73.38<math>\pm</math>19.83</b>
ITFE con logvar	Kappa	0.276	0.129	0.182	0.896	0.758	0.562	0.469	0.81	0.709	0.532 $\pm$ 0.268
	Exactitud [%]	63.81	56.43	59.13	94.81	87.87	78.12	73.44	90.5	85.44	76.62 $\pm$ 13.41
	AUC [%]	81.36	61.39	59.15	99.44	95.88	92.05	82.95	97.2	94.76	84.91 $\pm$ 14.39
	Precision [%]	87.52	56.38	58.31	94.1	93.29	71.03	80.0	89.9	89.17	79.97 $\pm$ 13.84
	Recall [%]	32.25	57.71	63.87	95.62	81.62	95.0	62.5	91.25	80.75	73.4 $\pm$ 19.86
LS con logpower	Kappa	<b>0.282</b>	0.129	0.144	<b>0.905</b>	<b>0.774</b>	0.539	<b>0.503</b>	0.831	0.683	0.532 $\pm$ 0.276
	Exactitud [%]	<b>64.12</b>	56.43	57.19	<b>95.25</b>	<b>88.69</b>	76.94	<b>75.12</b>	91.56	84.12	76.6 $\pm$ 13.81
	AUC	<b>81.57</b>	59.46	57.81	<b>99.43</b>	<b>95.51</b>	84.65	<b>82.04</b>	97.67	93.9	83.56 $\pm$ 14.73
	Precision [%]	<b>87.17</b>	55.39	56.74	<b>95.36</b>	<b>93.0</b>	72.22	<b>81.71</b>	89.74	81.08	79.16 $\pm$ 13.97
	Recall [%]	<b>33.12</b>	66.57	60.75	<b>95.13</b>	<b>83.87</b>	87.62	<b>64.75</b>	93.88	89.13	74.98 $\pm$ 19.24
LS con logvar	Kappa	0.278	0.143	0.128	0.905	0.756	<b>0.571</b>	0.492	<b>0.834</b>	0.669	0.531 $\pm$ 0.275
	Exactitud [%]	63.88	57.14	56.38	95.25	87.81	<b>78.56</b>	74.62	<b>91.69</b>	83.44	76.53 $\pm$ 13.76
	AUC [%]	81.67	59.95	58.15	99.38	95.61	<b>86.68</b>	81.91	<b>97.68</b>	93.37	83.82 $\pm$ 14.57
	Precision [%]	87.0	56.04	55.84	95.36	92.78	<b>73.61</b>	81.09	<b>90.24</b>	80.27	79.14 $\pm$ 13.93
	Recall [%]	32.63	66.71	61.12	95.13	82.12	<b>89.12</b>	64.25	<b>93.5</b>	88.75	74.81 $\pm$ 19.32

Por otro lado, al estudiar los resultados obtenidos con la propuesta WPD-CSP-FS descritos en la Tabla 5.20, se demuestra que el modelo LS con logpower logró los mejores resultados generales con un kappa promedio de 0.497 y una desviación estándar de 0.265. Este resultado es inferior al logrado por la propuesta OFB-CSP-FS con el mismo clasificador KNN y describe comportamientos ligeramente diferentes en sus resultados. Durante la clasificación en prueba, el modelo LS con logpower generó mejores resultados para los sujetos B03, B04 y B6; mientras que el modelo ITFE con logpower fue sobresaliente con los sujetos B07, B08 y B09; el modelo LS con logvar logró mejores resultados que los demás modelos con los sujetos B02 y B05; y finalmente el modelo ITFE con logvar solo logró superar el desempeño de los otros modelos durante la clasificación con el sujeto B01.

Los resultados obtenidos con la propuesta WPD-CSP-FS logran evidenciar una diferencia importante entre la clasificación de la clase L y la clase R, reflejada en la diferencia existente entre la precisión general de 77.82% y la exhaustividad de 69.41%. Sólo los sujetos B06, B08 y B09, obtuvieron una mejor clasificación de la clase R. En los sujetos restantes la clase L logró una mejor clasificación.

**Tabla 5.20.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador KNN con las diferentes técnicas de selección y extracción de características propuestas.

Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	0.316	0.117	0.101	0.884	0.702	0.438	<b>0.376</b>	<b>0.764</b>	<b>0.734</b>	0.492 $\pm$ 0.273
	Exactitud [%]	65.81	55.86	55.06	94.19	85.12	71.88	<b>68.81</b>	<b>88.19</b>	<b>86.69</b>	74.62 $\pm$ 13.65
	AUC [%]	76.55	59.06	53.6	98.47	94.28	88.87	<b>78.71</b>	<b>96.04</b>	<b>94.89</b>	82.27 $\pm$ 15.65
	Precision [%]	78.28	58.03	55.63	93.59	90.91	65.35	<b>78.66</b>	<b>85.73</b>	<b>85.49</b>	76.85 $\pm$ 13.22
	Recall [%]	43.88	42.43	49.88	94.87	78.12	93.12	<b>51.62</b>	<b>91.62</b>	<b>88.38</b>	70.44 $\pm$ 21.62
ITFE con logvar	Kappa	<b>0.324</b>	0.114	0.086	0.881	0.7	0.438	0.376	0.754	0.732	0.489 $\pm$ 0.273
	Exactitud [%]	<b>66.19</b>	55.71	54.31	94.06	85.0	71.88	68.81	87.69	86.62	74.47 $\pm$ 13.67
	AUC [%]	<b>77.78</b>	58.68	55.59	98.43	93.92	88.87	78.71	96.05	94.91	82.55 $\pm$ 15.24
	Precision [%]	<b>80.5</b>	57.69	55.19	93.25	90.0	65.35	78.66	85.26	85.47	76.82 $\pm$ 13.22
	Recall [%]	<b>42.88</b>	42.86	45.75	95.0	78.75	93.12	51.62	91.13	88.25	69.93 $\pm$ 22.14
LS con logpower	Kappa	0.309	0.141	<b>0.109</b>	<b>0.896</b>	0.71	<b>0.525</b>	0.351	0.726	0.702	<b>0.497<math>\pm</math>0.265</b>
	Exactitud [%]	65.44	57.07	<b>55.44</b>	<b>94.81</b>	85.5	<b>76.25</b>	67.56	86.31	85.12	<b>74.83<math>\pm</math>13.27</b>
	AUC [%]	78.47	59.98	<b>56.34</b>	<b>98.53</b>	94.62	<b>88.74</b>	76.07	95.2	94.19	<b>82.46<math>\pm</math>14.87</b>
	Precision [%]	81.91	60.08	<b>56.57</b>	<b>95.21</b>	92.8	<b>68.47</b>	78.8	83.53	82.99	<b>77.82<math>\pm</math>12.73</b>
	Recall [%]	39.63	42.14	<b>46.88</b>	<b>94.38</b>	77.0	<b>97.5</b>	48.12	90.5	88.5	<b>69.41<math>\pm</math>23.26</b>
LS con logvar	Kappa	0.285	<b>0.143</b>	0.091	0.895	<b>0.734</b>	0.506	0.35	0.724	0.705	0.493 $\pm$ 0.272
	Exactitud [%]	64.25	<b>57.14</b>	54.56	94.75	<b>86.69</b>	75.31	67.5	86.19	85.25	74.63 $\pm$ 13.6
	AUC [%]	78.58	<b>59.69</b>	56.0	98.57	<b>94.84</b>	87.45	76.58	95.29	94.32	82.37 $\pm$ 14.95
	Precision [%]	79.96	<b>60.0</b>	55.52	95.78	<b>90.96</b>	67.86	79.74	83.33	83.22	77.37 $\pm$ 12.8
	Recall [%]	38.0	<b>42.86</b>	46.0	93.62	<b>81.62</b>	96.25	47.0	90.5	88.38	69.36 $\pm$ 23.57

### 5.2.3. Resultados con Clasificador LDA

En la Tabla 5.21 se describen los resultados obtenidos utilizando la propuesta OFB-CSP-FS en conjunto con el clasificador LDA. En ella se indica que el modelo ITFE con logpower logró los mejores resultados generales con un kappa promedio de 0.564 y una desviación estándar de 0.259. Este mismo modelo obtuvo el mejor desempeño clasificando las señales de los sujetos B01, B02, B03 y B06; el modelo LS con logpower logró los mejores resultados en los sujetos B04, B05 y B07; y el modelo LS con logvar logró un mejor desempeño con los sujetos B08 y B09.

Por otra parte, estudiando los resultados mostrados en la Tabla 5.22 obtenidos utilizando la propuesta WPD-CSP-FS, se evidencia que el modelo ITFE con logpower logró los mejores resultados generales con un kappa promedio de 0.552 y una desviación estándar de 0.264. Este modelo logró superar al resto de los modelos clasificando las señales de los sujetos B03, B05, B06 y B09; el modelo ITFE con logvar obtuvo el mejor desempeño con el sujeto B07; el modelo LS con logpower logró los mejores resultados en los sujetos B01 y B04; y el modelo LS con logvar obtuvo una mejor clasificación con los sujetos B02 y B08.

**Tabla 5.21.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta OFB-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas..

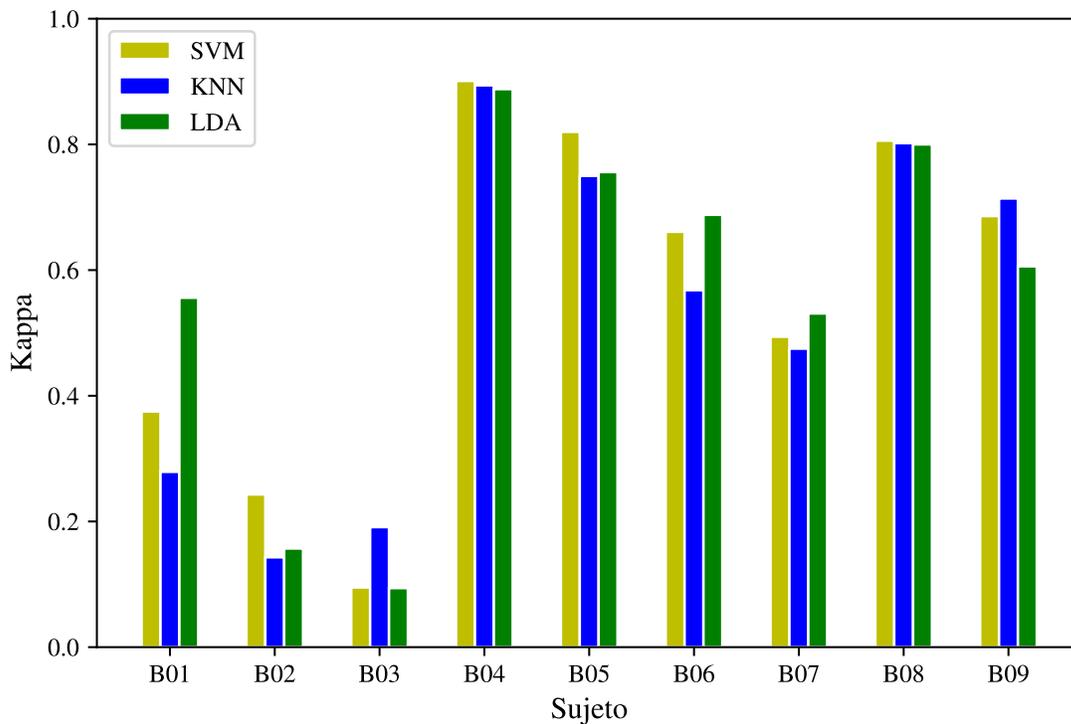
Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	<b>0.556</b>	<b>0.157</b>	<b>0.094</b>	0.888	0.756	<b>0.688</b>	0.531	0.8	0.606	<b>0.564<math>\pm</math>0.259</b>
	Exactitud [%]	<b>77.81</b>	<b>57.86</b>	<b>54.69</b>	94.38	87.81	<b>84.38</b>	76.56	90.0	80.31	<b>78.2<math>\pm</math>12.93</b>
	AUC [%]	<b>85.2</b>	<b>59.02</b>	<b>57.89</b>	99.03	96.1	<b>91.88</b>	84.42	96.49	93.11	<b>84.79<math>\pm</math>14.81</b>
	Precision [%]	<b>85.04</b>	<b>59.87</b>	<b>54.66</b>	96.1	90.07	<b>78.95</b>	81.95	85.56	88.8	<b>80.11<math>\pm</math>13.11</b>
	Recall [%]	<b>67.5</b>	<b>47.71</b>	<b>55.0</b>	92.5	85.0	<b>93.75</b>	68.12	96.25	69.38	<b>75.02<math>\pm</math>16.63</b>
ITFE con logvar	Kappa	0.556	0.157	0.081	0.888	0.756	0.688	0.531	0.8	0.606	0.563 $\pm$ 0.261
	Exactitud [%]	77.81	57.86	54.06	94.38	87.81	84.38	76.56	90.0	80.31	78.13 $\pm$ 13.06
	AUC [%]	85.2	59.02	58.11	99.03	96.1	91.88	84.42	96.49	93.11	84.82 $\pm$ 14.77
	Precision [%]	85.04	59.87	54.05	96.1	90.07	78.95	81.95	85.56	88.8	80.04 $\pm$ 13.24
	Recall [%]	67.5	47.71	54.25	92.5	85.0	93.75	68.12	96.25	69.38	74.94 $\pm$ 16.73
LS con logpower	Kappa	0.482	0.133	0.064	<b>0.895</b>	<b>0.794</b>	0.606	<b>0.532</b>	0.844	0.635	0.554 $\pm$ 0.277
	Exactitud [%]	74.13	56.64	53.19	<b>94.75</b>	<b>89.69</b>	80.31	<b>76.62</b>	92.19	81.75	77.7 $\pm$ 13.85
	AUC [%]	85.05	58.94	56.57	<b>99.24</b>	<b>96.58</b>	89.67	<b>84.98</b>	97.84	92.05	84.55 $\pm$ 15.12
	Precision [%]	88.28	57.13	53.04	<b>96.02</b>	<b>93.93</b>	74.59	<b>81.51</b>	90.04	89.31	80.43 $\pm$ 14.86
	Recall [%]	55.62	53.0	55.88	<b>93.38</b>	<b>84.88</b>	92.12	<b>68.88</b>	94.88	72.12	74.53 $\pm$ 16.3
LS con logvar	Kappa	0.485	0.134	0.056	0.893	0.79	0.564	0.518	<b>0.845</b>	<b>0.655</b>	0.549 $\pm$ 0.278
	Exactitud [%]	74.25	56.71	52.81	94.62	89.5	78.19	75.87	<b>92.25</b>	<b>82.75</b>	77.44 $\pm$ 13.9
	AUC [%]	84.81	59.14	56.58	99.22	96.91	87.49	84.47	<b>97.85</b>	<b>92.38</b>	84.32 $\pm$ 15.06
	Precision [%]	88.34	57.43	52.84	96.25	93.64	72.72	79.52	<b>90.15</b>	<b>88.66</b>	79.95 $\pm$ 14.9
	Recall [%]	55.88	51.86	52.5	92.88	84.88	90.5	69.88	<b>94.87</b>	<b>75.12</b>	74.26 $\pm$ 16.6

**Tabla 5.22.** Métricas de desempeño en clasificación de datos de prueba de la BCI Competition IV dataset 2b, utilizando la propuesta WPD-CSP-FS y clasificador LDA con las diferentes técnicas de selección y extracción de características propuestas..

Combinación	Métrica	B01	B02	B03	B04	B05	B06	B07	B08	B09	Prom $\pm$ std
ITFE con logpower	Kappa	0.454	0.16	<b>0.09</b>	0.884	<b>0.766</b>	<b>0.588</b>	0.501	0.782	<b>0.744</b>	<b>0.552<math>\pm</math>0.264</b>
	Exactitud [%]	72.69	58.0	<b>54.5</b>	94.19	<b>88.31</b>	<b>79.38</b>	75.06	89.12	<b>87.19</b>	<b>77.6<math>\pm</math>13.19</b>
	AUC [%]	82.16	61.74	<b>57.27</b>	98.19	<b>95.58</b>	<b>91.39</b>	83.07	95.94	<b>93.88</b>	<b>84.36<math>\pm</math>14.31</b>
	Precision [%]	81.72	59.88	<b>54.88</b>	95.97	<b>91.81</b>	<b>73.04</b>	76.84	86.83	<b>88.89</b>	<b>78.87<math>\pm</math>13.36</b>
	Recall [%]	58.5	48.71	<b>51.13</b>	92.25	<b>84.12</b>	<b>93.12</b>	71.75	92.25	<b>85.0</b>	<b>75.2<math>\pm</math>17.16</b>
ITFE con logvar	Kappa	0.449	0.149	0.07	0.884	0.766	0.588	<b>0.506</b>	0.782	0.732	0.547 $\pm$ 0.269
	Exactitud [%]	72.44	57.43	53.5	94.19	88.31	79.38	<b>75.31</b>	89.12	86.62	77.37 $\pm$ 13.44
	AUC [%]	82.06	61.45	56.45	98.19	95.58	91.39	<b>83.47</b>	95.94	93.98	84.28 $\pm$ 14.54
	Precision [%]	81.03	60.24	53.67	95.97	91.81	73.04	<b>77.18</b>	86.83	88.75	78.72 $\pm$ 13.52
	Recall [%]	58.62	43.86	52.12	92.25	84.12	93.12	<b>71.88</b>	92.25	83.87	74.68 $\pm$ 17.81
LS con logpower	Kappa	<b>0.475</b>	0.164	0.074	<b>0.885</b>	0.766	0.521	0.467	0.785	0.675	0.535 $\pm$ 0.262
	Exactitud [%]	<b>73.75</b>	58.21	53.69	<b>94.25</b>	88.31	76.06	73.38	89.25	83.75	76.74 $\pm$ 13.09
	AUC [%]	<b>83.51</b>	61.79	56.08	<b>98.34</b>	95.57	88.99	81.46	96.24	94.25	84.03 $\pm$ 14.52
	Precision [%]	<b>82.65</b>	59.14	53.74	<b>95.86</b>	92.28	69.48	74.61	88.52	88.48	78.31 $\pm$ 14.11
	Recall [%]	<b>60.13</b>	53.14	53.5	<b>92.5</b>	83.62	93.25	70.88	90.25	77.63	74.99 $\pm$ 15.39
LS con logvar	Kappa	0.468	<b>0.174</b>	0.075	0.879	0.764	0.526	0.474	<b>0.789</b>	0.676	0.536 $\pm$ 0.259
	Exactitud [%]	73.37	<b>58.71</b>	53.75	93.94	88.19	76.31	73.69	<b>89.44</b>	83.81	76.8 $\pm$ 12.96
	AUC [%]	83.57	<b>62.07</b>	56.4	98.47	95.66	88.95	81.34	<b>96.21</b>	94.32	84.11 $\pm$ 14.43
	Precision [%]	82.37	<b>59.5</b>	53.84	95.25	92.38	69.73	74.71	<b>87.92</b>	88.5	78.24 $\pm$ 13.89
	Recall [%]	59.5	<b>54.57</b>	53.25	92.5	83.25	93.25	71.62	<b>91.5</b>	77.75	75.24 $\pm$ 15.38

### 5.2.4. Análisis de Resultados

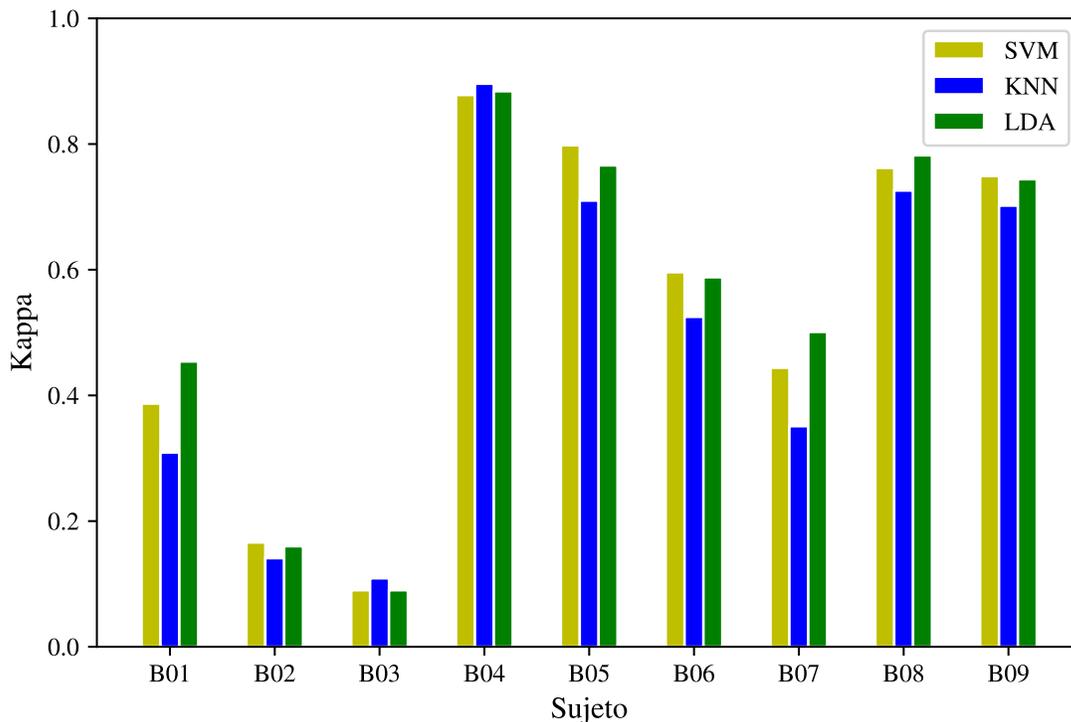
Analizando el comportamiento de cada uno de los clasificadores utilizando la propuesta OFB-CSP-FS, mostrado en la Figura 5.5, se puede observar que existen diferentes tendencias. En cuanto al clasificador KNN, los mejores resultados se obtuvieron con el sujeto B09 y el sujeto B03, donde además se superó ampliamente el desempeño de los demás clasificadores. Por otro lado, el clasificador LDA logró un resultado sobresaliente con el sujeto B01, y un desempeño superior con los sujetos B06 y B07. Pero el clasificador SVM logró el mejor desempeño promedio general concluidas todas las pruebas, superando los resultados de los otros clasificadores con los sujetos B02, B04, B05 y B08.



**Figura 5.5.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo OFB-CSP-FS por cada sujeto de prueba de la base de datos BCI Competition IV 2b.

De la misma forma, de acuerdo a los resultados que se muestran en la Figura 5.6, el comportamiento de los clasificadores utilizando la propuesta WPD-CSP-FS presenta algunas diferencias con respecto a los resultados de la propuesta OFB-CSP-FS. El clasificador KNN logró obtener resultados ligeramente superiores con los sujetos B03 y B04, pero fue superado

en el resto de pruebas. En esta oportunidad, aunque el clasificador SVM obtuvo resultados ligeramente superiores en las pruebas con los sujetos B02, B05, B06 y B09, este mismo no demostró una superioridad contundente sobre los demás clasificadores. Por esta razón, al evaluar la totalidad de las pruebas con WPD-CSP-FS, el clasificador LDA demostró un desempeño general promedio superior al resto de clasificadores con resultados destacados en los sujetos B01, B07 y B08.



**Figura 5.6.** Comparación de desempeño de la mejor combinación por cada clasificador utilizando el marco de trabajo WPD-CSP-FS por cada sujeto de prueba de la base de datos BCI Competition IV 2b.

En ambas propuestas también es posible identificar que los sujetos B01, B02, B03 y B04, recibieron resultados de clasificación notoriamente inferiores al resto de sujetos de prueba. El sujeto B03 representó el elemento más difícil de generalizar para los modelos propuestos mientras que el sujeto B04 logró obtener resultados de clasificación destacados.

### 5.3. Comparativa con la literatura

Para analizar los rendimientos de las propuestas OFB-CSP-FS y WPD-CSP-FS, y los aportes de la investigación, es importante contrastar los resultados obtenidos con la contribución de trabajos previos bajo las mismas condiciones de prueba. En la Tabla 5.23 se presentan los resultados expresados en métrica de Exactitud y Kappa promedio, obtenidos por 11 investigaciones publicadas desde el 2009 hasta el 2021 durante la clasificación de las cuatro clases L,R,F y T, de la base de datos BCI Competition IV 2a. Cada una de estas propuestas utilizó el método CSP, o un enfoque similar, como base para sus algoritmos en conjunto con otras técnicas que buscaban aumentar el desempeño original del filtro espacial. Los mejores resultados registrados por cada usuario están remarcados con el color azul. En ella es notorio observar como la propuesta OFB-CSP-FS superó al resto con una Exactitud promedio global de 72.18 % y un Kappa de 0.629, representando una mejora de 1.3 % con respecto al resultado obtenido por CWT-CSP-CNN en el 2021.

**Tabla 5.23.** Comparación entre métricas de Exactitud [%] y Kappa promedio publicadas en la literatura, y las obtenidas por los marcos de trabajo propuestos durante clasificación de cuatro clases con los datos de prueba de la BCI Competition VI 2a.

	Sujeto									Exact. (Kappa)
	A01	A02	A03	A04	A05	A06	A07	A08	A09	
CCSP [2009]	84.72	52.78	80.90	59.38	54.51	49.31	88.54	71.88	56.60	66.51 (0.553)
FBCSP [2012]	76.00	56.50	81.25	61.00	55.00	45.25	82.75	81.25	70.75	67.75 (0.570)
SSCSP [2013]	76.74	58.68	81.25	57.64	38.54	48.26	76.39	79.17	78.82	66.17 (0.548)
CSP-CNN [2015]	77.14	49.82	80.41	53.88	<b>65.47</b>	48.70	81.37	<b>84.39</b>	82.29	69.27 (ND)
SMM [2015]	81.94	<b>59.38</b>	81.60	62.85	59.03	49.36	86.11	77.78	78.47	70.35 (0.610)
FBCSP-CNN [2015]	80.55	53.82	84.72	64.58	59.03	44.10	84.03	86.8	77.77	70.60 (ND)
BO [2016]	82.12	44.86	86.60	66.28	48.72	53.30	72.64	82.33	76.35	68.13 (0.575)
CSP-SRIT2NFIS [2018]	74.65	45.48	74.31	39.58	32.99	37.90	54.17	66.32	66.31	54.63 (ND)
HSS-ELM [2019]	81.14	49.86	78.02	63.33	44.03	49.44	81.11	81.49	78.82	70.22 (0.570)
NSL-EEGNet [2020]	82.29	51.39	85.07	67.01	58.33	<b>56.25</b>	83.33	73.96	78.47	70.68 (ND)
CWT-CSP-CNN [2021]	<b>87.07</b>	56.17	<b>92.97</b>	68.67	39.85	52.00	<b>89.85</b>	72.14	82.56	71.25 (ND)
OFB-CSP-FS	86.60	55.63	81.18	<b>69.79</b>	64.23	49.79	77.36	82.01	<b>83.06</b>	<b>72.18 (0.629)</b>
WPD-CSP-FS	84.65	52.63	86.80	61.94	59.86	47.84	80.14	80.07	77.78	70.19 (0.603)

Abreviaciones: ND, No Disponible.

También cabe destacar que la propuesta OFB-CSP-FS logró superar al resto de algoritmos durante la clasificación de los sujetos A04 y A09, con 69.79 % y 83.06 %, respectivamente. Por otro lado, la propuesta WPD-CSP-FS también logró resultados competitivos superando a 6 de los 11 algoritmos de la comparación con una Exactitud promedio de 70.19 % y Kappa de 0.603.

En cuanto al rendimiento mostrado en clasificación binaria entre las clases L y R con las bases de datos BCI Competition IV 2a y 2b, en la Tabla 5.24 se muestran los resultados de Kappa máxima publicados en la literatura por diferentes investigadores desde 2009 hasta 2019. Todos los algoritmos en la comparación están basados igualmente en el método CSP en conjunto de otras etapas de procesamiento y clasificadores que buscaban maximizar su desempeño. Los sujetos identificados con la letra *A* corresponden a los de la base de datos BCI Competition IV 2a, y los identificados con la letra *B* corresponden a los de la base de datos 2b.

**Tabla 5.24.** Comparación entre máximo kappa obtenido por los marcos de trabajo propuestos y los algoritmos en la literatura durante la clasificación de las clases L y R con datos de prueba de la BCI Competition VI 2a y 2b.

	Algoritmos							
	DCSP [2009]	FBCSP [2012]	WPD-CSP [2016]	IDFB-CSP [2017b]	CSP-TSM [2017a]	CDF-CSP [2019]	OFB-CSP-FS	WPD-CSP-FS
A01	0.46	0.60	0.56	0.56	0.60	0.62	<b>0.89</b>	0.88
A02	<b>0.24</b>	0.20	0.23	0.21	0.17	0.20	0.13	-0.02
A03	0.20	0.82	0.77	0.87	0.81	0.87	<b>0.94</b>	0.90
A04	0.09	0.38	0.24	0.34	<b>0.45</b>	0.40	0.44	0.36
A05	0.17	0.38	0.14	0.21	0.13	0.47	0.79	<b>0.85</b>
A06	0.31	0.35	0.34	0.35	0.33	0.37	<b>0.38</b>	0.32
A07	0.51	0.47	0.20	0.46	0.37	0.50	<b>0.82</b>	0.76
A08	0.60	0.87	0.70	0.87	0.85	0.78	<b>0.94</b>	0.92
A09	0.74	0.62	0.59	0.82	0.69	0.66	0.79	<b>0.82</b>
B01	0.31	0.39	0.31	0.30	0.36	<b>0.48</b>	0.27	0.41
B02	0.20	0.22	0.29	0.21	<b>0.33</b>	0.29	0.19	0.18
B03	0.30	<b>0.32</b>	0.29	0.25	0.16	0.25	0.13	0.11
B04	0.90	0.80	0.92	0.90	0.93	<b>0.97</b>	0.93	0.89
B05	0.83	0.90	0.70	0.83	0.54	<b>0.98</b>	0.80	0.80
B06	0.60	<b>0.75</b>	0.36	0.66	0.65	0.70	0.59	0.56
B07	0.67	0.66	0.56	<b>0.67</b>	0.58	0.67	0.54	0.41
B08	0.86	0.86	0.79	<b>0.88</b>	0.86	0.84	0.80	0.79
B09	0.67	0.62	0.50	0.73	0.72	<b>0.79</b>	0.69	0.74
Prom.	0.48	0.57	0.47	0.56	0.53	0.60	<b>0.61</b>	0.59

Es posible observar que la propuesta OFB-CSP-FS nuevamente presenta el mejor desempeño promedio entre todos los algoritmos con un Kappa de 0.61, superando en 1.6% el resultado del modelo CDF-CSP publicado en el 2019. De esta misma manera, se evidencia resultados de clasificación superiores en los sujetos A01, A03, A06, A07 y A08, marcados en color azul. En esta oportunidad la propuesta WPD-CSP-FS obtuvo un Kappa máximo promedio de 0.59, superior a cinco de los seis algoritmos en la comparación, logrando resultados sobresalientes con los sujetos A05 y A09.

También es importante destacar que aunque se observa superioridad de los demás algoritmos en la clasificación de los sujetos de la base de datos 2b, cada uno de estos

modelos publicados en la literatura utilizó un esquema de selección de datos diferente, donde se seleccionaron las sesiones de datos de entrenamiento que maximizara el desempeño de clasificación en cada sujeto. Por el contrario, el esquema de entrenamiento utilizado para las propuestas de esta investigación contempla el uso de todas las sesiones de datos de entrenamiento establecidas originalmente en las bases de datos, con la finalidad de someter a los modelos a la compleja tarea de buscar patrones que permitieran lograr una correcta generalización entre sesiones de días diferentes. Sin embargo, el correcto desempeño generalizado entre ambas bases de datos demuestra que la propuesta OFB-CSP-FS y WPD-CSP-FS ofrecen desempeños superiores o cercanos al resto trabajando con los datos de todas las sesiones disponibles en ambas bases de datos.

## CAPÍTULO 6

---

### Conclusiones Generales

---

En este trabajo se planteó el desarrollo de un marco de trabajo robusto utilizando técnicas de aprendizaje automático que generen un mayor desempeño de los algoritmos existentes en la literatura. De este modo, la metodología establecida para la implementación de las propuestas OFB-CSP-FS y WPD-CSP-FS se basaron en la técnica de filtrado espacial CSP, creando modelos que incorporaran diferentes métodos de procesamiento, extracción, selección de características y clasificación, para maximizar el desempeño de dicha técnica.

Los resultados obtenidos en la fase de pruebas sugieren que la propuesta más exitosa para la clasificación de las señales de intenciones motoras corresponde al OFB-CSP-FS, utilizando el logaritmo de la potencia promedio de la señal como extractor de características, el cálculo del valor laplaciano como herramienta de selección de los mejores atributos y finalmente un clasificador SVM.

La etapa correspondiente al banco de filtros pasa-banda con solapamiento en frecuencia demostró tener mayor capacidad para extraer información diferenciadora entre clases que la descomposición en paquetes de ondas, evitando de esta manera el problema de variabilidad existente en las bandas de frecuencia relacionadas con la imaginación motora, producto de las variables biológicas que existen entre cada sujeto. Al mismo tiempo, tomando en cuenta el

efecto que produce la Desincronización Relacionada a Eventos (ERD) sobre la potencia de las señales EEG al ejecutar una imaginación motora, el logaritmo de la potencia promedio de la señal logró representar correctamente la diferencia entre las clases actuando como un mejor atributo que el logaritmo de la varianza. Además el criterio utilizado por el cálculo del valor laplaciano demostró ser más efectivo que el cálculo de la información mutua como selector de los mejores atributos para la clasificación. Finalmente la implementación de la función RBF como kernel de transformación en el modelo SVM aplicado bajo el esquema OVR, generó mejores resultados que el modelo de clasificación lineal y el modelo basado en vecinos cercanos.

En cuanto a los rendimientos generales obtenidos, se puede apreciar que el desempeño logrado con la base de datos BCI Competition IV 2a en la clasificación binaria entre las clases L y R, es ligeramente superior al desempeño obtenido con la base de datos BCI Competition IV 2b. Las diferencias existentes pueden estar asociadas, además de las dificultades propias de cada sujeto de prueba, a la disponibilidad de más canales de adquisición en la base de datos 2a y al efecto negativo que genera la divergencia de las señales adquiridas en sesiones de días diferentes. Tomando en cuenta que las señales de la base de datos 2a se tomaron en dos días distintos con 22 canales de adquisición mientras que los de la base de datos 2b se tomaron en 5 días diferentes con 3 canales, ambos con presencia de artefactos de ruido, igualmente los modelos planteados mantuvieron un rendimiento general bastante competitivo. De esta manera se demuestra que los modelos propuestos logran mejorar el potencial del método CSP, en comparación con los resultados del estado del arte.

Se debe agregar que durante las pruebas en clasificación binaria, las señales correspondientes al pensamiento del brazo derecho (R) y el pensamiento de los pies (F) resultaron los mejores clasificados por los modelos propuestos. Misma tendencia que se observó durante la clasificación multi-clases donde la clase R logró los mejores resultados de clasificación con respecto a las demás clases. Además, es importante destacar que en los escenarios con múltiples clases se evidencia un efecto de solapamiento que afecta el desempeño de los modelos. Un ejemplo de esto se puede observar en el caso RFT donde se obtuvo la mejor clasificación para la clase R cuando no se incluyeron los datos de la clase L, y viceversa. El mismo comportamiento se evidenció entre las clases F y T, donde la mejor clasificación de la clase F se obtuvo en el caso LRF cuando la clase T no fue incluida. Este comportamiento se repitió en todos los modelos planteados aunque las zonas de la corteza motora del cerebro relacionadas con estas partes del cuerpo están distanciadas. Caso contrario ocurrió en la clasificación entre las clases RF, que reportó los mejores resultados de clasificación aunque las zonas de la corteza motora relacionada con estas partes del cuerpo están más cercanas. Esto sugiere que las propuestas planteadas no están aprovechando completamente la información espacial que ofrece la localización de los

electrodos en el cráneo, lo que representa una oportunidad de mejora para los modelos.

## 6.1. Trabajo futuro

Con respecto a la investigación desarrollada hasta el momento, se proponen los siguientes temas como posibles trabajos futuros:

- Implementar una etapa de pre-procesamiento previa basada en Aprendizaje Profundo [Zhang et al., 2021b], con la finalidad de corregir la señal original antes de ser procesada por el marco de trabajo cuando exista presencia intensa de artefactos de ruido.
- Utilizar algoritmos de optimización avanzados para determinar la mejor combinación de parámetros que maximice el desempeño de cada etapa del marco de trabajo que tome en cuenta información temporal y espacial de las señales.
- Sustituir el algoritmo de CSP clásico por versiones más actualizadas del mismo, tales como el CCSP-L21 [Gu et al., 2023], e-CSP [Sun et al., 2022] o tCSP [Ma et al., 2023], que aseguran resolver deficiencias relacionadas con la sensibilidad al ruido y la obtención de características más robustas.
- Implementar una metodología punto a punto con algoritmos de Aprendizaje Profundo que permitan clasificar de manera robusta las señales originales EEG.

---

## Referencias

---

- Abdulkader, S. N., Atia, A., and Mostafa, M.-S. M. (2015). Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2):213–230.
- Abdullah, Faye, I., and Islam, M. R. (2022). Eeg channel selection techniques in motor imagery applications: A review and new perspectives. *Bioengineering*, 9(12).
- Ablin, P. A., Cardoso, J.-F., and Gramfort, A. (2019). Beyond Pham’s algorithm for joint diagonalization. In *ESSAN 2019 - 27th European symposium on artificial neural networks*, Bruges, Belgium.
- Akram, F., Alwakeel, A., Alwakeel, M., Hijji, M., and Masud, U. (2022). A symbols based bci paradigm for intelligent home control using p300 event-related potentials. *Sensors*, 22(24).
- Aljalal, M., Djemal, R., AlSharabi, K., and Ibrahim, S. (2018). Feature extraction of eeg based motor imagery using csp based on logarithmic band power, entropy and energy. In *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*, pages 1–6.
- Ambardar, A. (1999). *Analog and Digital Signal Processing*. Brooks/Cole Publishing Company.
- Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., and Zhang, H. (2012). Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in Neuroscience*, 6.

- Austerlitz, H. (2003). Chapter 10 - data processing and analysis. In Austerlitz, H., editor, *Data Acquisition Techniques Using PCs (Second Edition)*, pages 222–250. Academic Press, San Diego, second edition edition.
- Bashashati, H., Ward, R. K., and Bashashati, A. (2016). User-customized brain computer interfaces using bayesian optimization. *Journal of Neural Engineering*, 13(2):026001.
- Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., and Muller, K.-r. (2008). Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal Processing Magazine*, 25(1):41–56.
- Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27.
- Donoho, D. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627.
- Dornhege, G., Blankertz, B., Curio, G., and Muller, K.-R. (2004). Boosting bit rates in noninvasive eeg single-trial classifications by feature combination and multiclass paradigms. *IEEE Transactions on Biomedical Engineering*, 51(6):993–1002.
- Duda, R., Hart, P., and Stork, D. (2006). *Pattern Classification*. Wiley-India.
- Feng, G., Hao, L., and Nuo, G. (2019). Feature extraction algorithm based on csp and wavelet packet for motor imagery eeg signals. In *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, pages 798–802.
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*. Academic Press, Boston, second edition edition.
- Fumanal-Idocin, J., Wang, Y.-K., Lin, C.-T., Fernández, J., Sanz, J. A., and Bustince, H. (2022). Motor-imagery-based brain-computer interface using signal derivation and aggregation functions. *IEEE Transactions on Cybernetics*, 52(8):7944–7955.
- Garcia-Ramos, D. and Cruces-Álvarez, S. (2020). *EEG signal classification for MI-BCI applications*. thesis, Universidad de Sevilla. Departamento de Teoría de la Señal y Comunicaciones.
- Goncharova, I., McFarland, D., Vaughan, T., and Wolpaw, J. (2003). Emg contamination of eeg: spectral and topographical characteristics. *Clinical Neurophysiology*, 114(9):1580–1593.

- Grosse-Wentrup, M. and Buss, M. (2008). Multiclass common spatial patterns and information theoretic feature extraction. *IEEE Transactions on Biomedical Engineering*, 55(8):1991–2000.
- Gu, J., Jiang, J., Ge, S., and Wang, H. (2023). Capped l21-norm-based common spatial patterns for eeg signals classification applicable to bci systems. *Medical and Biological Engineering and Computing*, 61(5):1083 – 1092.
- Hayta, d., Irimia, D. C., Guger, C., Erkutlu, d., and Güzelbey, d. H. (2022). Optimizing motor imagery parameters for robotic arm control by brain-computer interface. *Brain Sciences*, 12(7).
- He, X., Cai, D., and Niyogi, P. (2005). Laplacian score for feature selection. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press.
- INEGI (2020). Censo y conteo de población y vivienda.
- Jiang, X., Bian, G.-B., and Tian, Z. (2019). Removal of artifacts from eeg signals: A review. *Sensors*, 19(5).
- Kang, H., Nam, Y., and Choi, S. (2009). Composite common spatial pattern for subject-to-subject transfer. *IEEE Signal Processing Letters*, 16(8):683–686.
- Kawala-Sterniuk, A., Browarska, N., Al-Bakri, A., Pelc, M., Zygarlicki, J., Sidikova, M., Martinek, R., and Gorzelanczyk, E. J. (2021). Summary of over fifty years with brain-computer interfaces—a review. *Brain Sciences*, 11(1).
- Klem, G. H., Lüders, H. O., Jasper, H. H., and Elger, C. (1999). The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*, 52:3–6.
- Koles, Z. J., Lazar, M. S., and Zhou, S. Z. (1990). Spatial patterns underlying population differences in the background eeg. *Brain Topography*, 2:275–284.
- Kotsiantis, S. (2007). Supervised machine learning: A review of classification techniques. *Informatica (Eslovenia)*, 31:249–268.
- Kumar, S., Mamun, K., and Sharma, A. (2017a). Csp-tsm: Optimizing the performance of riemannian tangent space mapping using common spatial pattern for mi-bci. *Computers in Biology and Medicine*, 91:231–242.

- Kumar, S., Sharma, A., and Tsunoda, T. (2017b). An improved discriminative filter bank selection approach for motor imagery eeg signal classification using mutual information. *BMC Bioinformatics*, 18.
- Lee, K. J., Park, C., and Lee, B. (2015). Elimination of eeg artifacts from a single-channel eeg using sparse derivative method. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 2384–2389.
- Lim, J., Lin, D., Sohn, W. J., McCrimmon, C. M., Wang, P. T., Nenadic, Z., and Do, A. H. (2022). *BCI-Based Neuroprostheses and Physiotherapies for Stroke Motor Rehabilitation*, pages 509–524. Springer International Publishing, Cham.
- Luo, J., Wang, J., Xu, R., and Xu, K. (2019). Class discrepancy-guided sub-band filter-based common spatial pattern for motor imagery classification. *Journal of Neuroscience Methods*, 323:98–107.
- Luo, L., Xie, Y., Zhang, Z., and Li, W.-J. (2015). Support matrix machines. In Bach, F. and Blei, D., editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 938–947, Lille, France. PMLR.
- Ma, Z., Wang, K., Xu, M., Yi, W., Xu, F., and Ming, D. (2023). Transformed common spatial pattern for motor imagery-based brain-computer interfaces. *Frontiers in Neuroscience*, 17.
- Mahamune, R. and Laskar, S. H. (2021). Classification of the four-class motor imagery signals using continuous wavelet transform filter bank-based two-dimensional images. *International Journal of Imaging Systems and Technology*, 31(4):2237 – 2248.
- Marieb, E. and Hoehn, K. (2010). *Human Anatomy & Physiology*. Pearson Education, San Francisco, 6 edition.
- Meisheri, H., Ramrao, N., and Mitra, S. (2018). Multiclass common spatial pattern for EEG based brain computer interface with adaptive learning classifier. *CoRR*, abs/1802.09046.
- Michel, C. M. and Brunet, D. (2019). Eeg source imaging: A practical review of the analysis steps. *Frontiers in Neurology*, 10.
- Mohammadi, M. and Mosavi, M. (2021). Comparison of two methods of removing eeg artifacts for use in a motor imagery-based brain computer interface. *Evolving Systems*, 12:1–14.
- Mwata-Velu, T., Avina-Cervantes, J., Ruiz-Pinales, J., Garcia-Calva, T., González-Barbosa, E., Hurtado-Ramos, J., and González-Barbosa, J. (2022a). Improving motor imagery eeg

- classification based on channel selection using a deep learning architecture. *Mathematics*, 10(13).
- Mwata-Velu, T., Ruiz-Pinales, J., Avina-Cervantes, J. G., Gonzalez-Barbosa, J.-J., and Contreras-Hernandez, J. (2022b). Empirical mode decomposition and a bidirectional lstm architecture used to decode individual finger mi-eeeg signals. *Journal of Advances in Applied & Computational Mathematics*, 9:32–48.
- Mwata-Velu, T., Ruiz-Pinales, J., Rostro-Gonzalez, H., Ibarra-Manzano, M. A., Cruz-Duarte, J. M., and Avina-Cervantes, J. G. (2021). Motor imagery classification based on a recurrent-convolutional architecture to control a hexapod robot. *Mathematics*, 9(6).
- Nicolas-Alonso, L. F. and Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279.
- Organization, W. H. and Bank, W. (2011). World report on disability 2011.
- Pham, D. T. (2001). Joint approximate diagonalization of positive definite hermitian matrices. *SIAM Journal on Matrix Analysis and Applications*, 22(4):1136–1152.
- Radüntz, T., Tahoun, M. A., A-Megeed, M., and Meffert, B. (2017). Artifact elimination in neurosciences. In Hassanien, A. E., Shaalan, K., Gaber, T., Azar, A. T., and Tolba, M. F., editors, *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016*, pages 751–759, Cham. Springer International Publishing.
- Raza, H., Cecotti, H., and Prasad, G. (2015). Optimising frequency band selection with forward-addition and backward-elimination algorithms in eeg-based brain-computer interfaces. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7.
- Raza, H., Chowdhury, A., Bhattacharyya, S., and Samothrakis, S. (2020). Single-trial eeg classification with eegnet and neural structured learning for improving bci performance.
- Sakhavi, S., Guan, C., and Yan, S. (2015). Parallel convolutional-linear neural network for motor imagery classification. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 2736–2740.
- Samek, W., Meinecke, F. C., and Müller, K.-R. (2013). Transferring subspaces between subjects in brain-computer interfacing. *IEEE Transactions on Biomedical Engineering*, 60(8):2289–2298.

- Schalk, G., McFarland, D., Hinterberger, T., Birbaumer, N., and Wolpaw, J. (2004). Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043.
- Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., and Pfurtscheller, G. (2007). A fully automated correction method of eeg artifacts in eeg recordings. *Clinical Neurophysiology*, 118(1):98–104.
- Shawe-Taylor, J. and Sun, S. (2014). Chapter 16 - kernel methods and support vector machines. In Diniz, P. S., Suykens, J. A., Chellappa, R., and Theodoridis, S., editors, *Academic Press Library in Signal Processing: Volume 1*, volume 1 of *Academic Press Library in Signal Processing*, pages 857–881. Elsevier.
- She, Q., Hu, B., Luo, Z., Nguyen, T., and Zhang, Y. (2019). A hierarchical semi-supervised extreme learning machine method for eeg recognition. *Medical & Biological Engineering & Computing*, 57(1):147–157.
- Shmilovici, A. (2005). *Support Vector Machines*, pages 257–276. Springer US, Boston, MA.
- Smith, J. O. (2007). *Introduction to Digital Filters with Audio Applications*. W3K Publishing.
- Smith, S. W. (1997). *The Scientist and Engineer’s Guide to Digital Signal Processing*. California Technical Publishing, USA.
- Sun, J., Wei, M., Luo, N., Li, Z., and Wang, H. (2022). Euler common spatial patterns for eeg classification. *Medical and Biological Engineering and Computing*, 60(3):753 – 767.
- Sörnmo, L. and Laguna, P. (2005). Chapter 2 - the electroencephalogram—a brief background. In Sörnmo, L. and Laguna, P., editors, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Biomedical Engineering, pages 25–53. Academic Press, Burlington.
- Tandon, N. and Allison, B. (2018). *Noninvasive Brain-Computer Interfaces*, pages 357–377. Academic Press.
- Tangemann, M., Müller, K.-R., Aertsen, A., Birbaumer, N., Braun, C., Brunner, C., Leeb, R., Mehring, C., Miller, K., Mueller-Putz, G., Nolte, G., Pfurtscheller, G., Preissl, H., Schalk, G., Schlögl, A., Vidaurre, C., Waldert, S., and Blankertz, B. (2012). Review of the bci competition iv. *Frontiers in Neuroscience*, 6.

- Thomas, K. P., Lau, C. T., Vinod, A., Guan, C., and Ang, K. K. (2009). A new discriminative common spatial pattern method for motor imagery brain—computer interfaces. *IEEE Transactions on Biomedical Engineering*, 56(11):2730 – 2733.
- Tian, G. and Liu, Y. (2018). Study on classification of left-right hands motor imagery eeg signals based on cnn. In *2018 IEEE 17th International Conference on Cognitive Informatics and Cognitive Computing (ICCI\*CC)*, pages 324–329.
- Usakli, A. B. (2010). Improvement of EEG signal acquisition: an electrical aspect for state of the art of front end. *Comput Intell Neurosci*, 2010:630–649.
- Wang, Y., Gao, S., and Gao, X. (2005). Common spatial pattern method for channel selection in motor imagery based brain-computer interface. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 5392–5395.
- Yang, B., Li, H., Wang, Q., and Zhang, Y. (2016). Subject-based feature extraction by using fisher wpd-csp in brain–computer interfaces. *Computer Methods and Programs in Biomedicine*, 129:21–28.
- Yang, H., Sakhavi, S., Ang, K. K., and Guan, C. (2015). On the use of convolutional neural networks and augmented csp features for multi-class motor imagery of eeg signals classification. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2620–2623.
- Zahid, S., Aqil, M., Tufail, M., and Nazir, M. (2020). Online classification of multiple motor imagery tasks using filter bank based maximum-a-posteriori common spatial pattern filters. *IRBM*, 41(3):141–150.
- Zhang, C., Kim, Y.-K., and Eskandarian, A. (2021a). Eeg-inception: An accurate and robust end-to-end neural network for eeg-based motor imagery classification. *Journal of Neural Engineering*, 18(4):046014.
- Zhang, H., Zhao, M., Wei, C., Mantini, D., Li, Z., and Liu, Q. (2021b). Eegdenoisenet: a benchmark dataset for deep learning solutions of eeg denoising. *Journal of Neural Engineering*, 18(5):056057.