



**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA**



**CLASIFICACIÓN DE EMOCIONES EN ELECTROENCEFALOGRAMA CON WAVELET
SCATTERING Y ALGORITMOS DE MACHINE LEARNING**

POR

Alonso Valderrama Hickmann

Informe Final Memoria de Título presentada a la Facultad de Ingeniería de la Universidad de
Concepción para optar al grado académico de Ingeniero/a Civil Biomédica

Profesores Guía:

Esteban Pino

Mabel Urrutia

Comisión:

Pamela Guevara

Agosto

Concepción (Chile)

© 2024 Alonso Valderrama Hickmann

© 2024 Alonso Valderrama Hickmann

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

Agradecimientos

Dra. Mabel Urrutia Profesora Asociada Facultad de Educación Universidad de Concepción

Neurociencia Cognitiva UdeC

Procedural Memory Deficits in Preschool Children with Developmental Language Disorder in a Spanish-Speaking Population

Agradecimientos a: Fondecyt Exploración N°13220040: “Emocrea, descubre y recrea tus emociones en entornos de realidad virtual mediante técnicas neurocientíficas y de inteligencia artificial”.

National Center for Artificial Intelligence CENIA FB210017, Basal ANID

Quiero expresar mi más sincero agradecimiento a mis padres y mi hermana, quienes han sido mi mayor apoyo a lo largo de mis estudios. Sin su incondicional respaldo y motivación, no habría llegado hasta aquí.

Agradezco también a mi profesor guía de Memoria y Tesis, Esteban Pino, por su valiosa orientación y apoyo durante todo el proceso.

Asimismo, agradezco a mi guía de memoria y jefa del proyecto FONDECYT 13220040, Mabel Urrutia, por permitirme ser parte de un equipo de investigación y proporcionar el apoyo económico necesario para la realización de este proyecto.

Agradezco especialmente a Francisca Jara por estar conmigo durante los momentos más difíciles y frustrantes de este proyecto.

Por último, deseo hacer una mención también a mis compañeros de laboratorio, quienes hicieron que las largas horas de trabajo en la universidad fueran más llevaderas y divertidas.

Resumen

La clasificación de emociones a partir de señales de electroencefalograma (EEG) es de gran importancia para el desarrollo de tecnologías avanzadas en neurociencia afectiva e interfaces cerebro-computadora. Este campo enfrenta desafíos considerables debido a la complejidad y variabilidad de las señales EEG, lo que dificulta la creación de algoritmos robustos y generalizables. En esta memoria ligada al proyecto FONDECYT 13220040 se evaluó el uso de la Transformada Wavelet Scattering (WS) como técnica para la extracción de descriptores de señales EEG, con el objetivo de mejorar el rendimiento de los clasificadores de emociones. El enfoque propuesto busca superar las limitaciones de los métodos tradicionales de extracción de características, ofreciendo una mayor capacidad de generalización.

La metodología fue evaluada utilizando la base de datos DEAP, un referente en el estudio de reconocimiento de emociones con EEG. Los resultados muestran que el clasificador basado en WS, combinado con un Perceptrón Multicapa (MLP) y PCA para la reducción de dimensionalidad, alcanzó un alto rendimiento con precisiones superiores al 95 % en la clasificación de las dimensiones emocionales de Valence, Arousal, Dominance y Liking al utilizar validación cruzada con 5 k-folds. Sin embargo, al aplicar el método Leave-One-Subject-Out (LOSO), se observó una disminución en la precisión, lo que sugiere la necesidad de ajustar la metodología para abordar problemas como la correlación entre canales y la variabilidad intersujeto. En conclusión, la Transformada Wavelet Scattering (WS) podría mostrar un potencial significativo en la clasificación de emociones, aunque es crucial mejorar las estrategias de evaluación para asegurar la robustez y generalización del modelo en aplicaciones reales.

Abstract

The classification of emotions from electroencephalogram (EEG) signals holds great importance for the development of advanced technologies in affective neuroscience and brain-computer interfaces. This field faces considerable challenges due to the complexity and variability of EEG signals, making it difficult to create robust and generalizable algorithms. In this thesis, associated with the FONDECYT project 13220040, the use of the Wavelet Scattering Transform (WS) was evaluated as a technique for extracting descriptors from EEG signals. This was done with the objective of improving the performance of emotion classifiers. The proposed approach aims to overcome the limitations of traditional feature extraction methods, offering greater generalization capacity.

The methodology was evaluated using the DEAP database, a benchmark in the study of emotion recognition using EEG. The results show that the WS-based classifier, combined with a Multilayer Perceptron (MLP) and PCA for dimensionality reduction, achieved high performance with accuracies exceeding 95 % in the classification of the emotional dimensions of Valence, Arousal, Dominance, and Liking using 5-fold cross-validation. However, when applying the Leave-One-Subject-Out (LO-SO) method, a decrease in accuracy was observed, suggesting the need to adjust the methodology to address issues such as channel correlation and inter-subject variability. In conclusion, the WS shows significant potential in emotion classification, although it is crucial to improve evaluation strategies to ensure the robustness and generalization of the model in real-world applications.

Tabla de Contenidos

Agradecimientos	II
Resumen	V
Abstract	VI
Índice de Tablas	X
Índice de Figuras	XI
1. Introducción General	1
1.1. Objetivos	2
1.1.1. Objetivo general	2
1.1.2. Objetivos específicos	2
1.1.3. Alcances	2
1.1.4. Limitaciones	3
1.2. Temario	3
2. Marco Teórico	4
2.1. Introducción	4
2.2. Experimentos para detección de emociones	4
2.3. Adquisición señal EEG	5
2.4. Teoría clasificación emociones	6
2.4.1. Enfoque categórico	6
2.4.2. Enfoque dimensional	6
2.5. Bases de datos	8
2.6. Métodos de extracción de descriptores de EEG para entrenamiento de algoritmos . .	8
2.6.1. Short-time Fourier Transform (STFT)	9
2.6.2. Power Spectral density (PSD)	9
2.6.3. Differential Entropy (DE)	9
2.6.4. Wavelet Transform (WT)	10
2.6.5. Empirical Mode Decomposition (EMD)	10
2.6.6. Functional Connectivity Maps (FC)	11
2.7. Machine Learning	11

2.7.1. Support Vector Machine (SVM)	11
2.7.2. Random Forest Regression (RFR)	12
2.7.3. K-means	12
2.7.4. Adaboost	12
2.8. Redes Neuronales	12
2.8.1. Definición red neuronal	13
2.8.2. Multilayer Perceptron (MLP)	14
2.8.3. Deep belief networks (DBNs)	14
2.8.4. Convolutional neural networks (CNNs)	14
2.8.5. Recurrent neural networks (RNNs)	14
2.9. Wavelet Scattering (WS)	15
3. Estado del arte	18
3.1. Introducción	18
3.2. Algoritmos que dividen datos de forma generalizada.	18
3.3. Algoritmos que utilizan método LOSO	19
3.4. Discusión	21
4. Desarrollo y resultados	23
4.1. Introducción	23
4.2. Metodología	24
4.2.1. Preparación set de datos	24
4.2.2. Normalización y extracción de descriptores con WS	25
4.2.3. Creación de conjuntos de entrenamiento/validación y reducción de dimensio- nalidad	25
4.2.4. Arquitectura del clasificador	27
4.2.5. Entrenamiento y evaluación de clasificadores	28
4.3. Resultados	30
4.3.1. Curva de Accuracy	32
4.3.2. Comportamiento de la Pérdida en Entrenamiento en las 4 dimensiones emo- cionales:	34
4.4. Discusión	35
4.4.1. Comparación con estado del arte:	35
4.4.2. Fortalezas y debilidades del clasificador propuesto	36
5. Conclusión	38

	IX
6. Trabajo futuro	39
Glosario	40
Anexo A	43

Índice de Tablas

3.1. Revisión algoritmos de Deep Learning utilizados en la detección de emociones . . .	21
4.1. Contenido video de cada participante	25
4.2. Dimensiones datos al aplicar WS	25
4.3. Dimensiones de datos al aplicar PCA. nt denota la cantidad de muestras utilizadas para entrenamiento, dependiendo de si se usa método de 5 k-folds o LOSO. nv denota la cantidad de muestras utilizadas para validación, dependiendo de si se usa método de 5 k-folds o LOSO	27
4.4. Arquitectura clasificador de emociones	27
4.5. Hiperparámetros del MLP	29
4.6. Resultados Promedio de Clasificación en 5 K-Folds para las 4 Dimensiones Emocionales	31
4.7. Resultados Promedio de Clasificación con LOSO para las 4 Dimensiones Emocionales	31
4.8. Algoritmo propuesto en comparación con estado del arte	36
4.9. Algoritmo propuesto en comparación con estado del arte método LOSO	37
6.1. Orden de Geneva para canales de EEG [1].	43

Índice de Figuras

2.1. Etapas entrenamiento algoritmo detección de emociones	4
2.2. Sistemas configuración EEG	5
2.3. Enfoque categórico	6
2.4. Enfoque Dimensional	7
2.5. Línea temporal desarrollo redes neuronales, bases de datos y modelos de clasificación emociones.	13
2.6. Arquitectura WS	17
4.1. Esquema de entrenamiento	28
4.2. Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Valence. . .	31
4.3. Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Arousal. . .	32
4.4. Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Dominance. .	32
4.5. Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Liking. . . .	33
4.6. Accuracy y pérdida media al aplicar método LOSO en Valence.	33
4.7. Accuracy y pérdida media al aplicar método LOSO en Arousal.	34
4.8. Accuracy y pérdida media al aplicar método LOSO en dominance.	34
4.9. Accuracy y pérdida media al aplicar método LOSO en Liking.	35

1 Introducción General

El desarrollo de habilidades sociales y emocionales es crucial para la toma de decisiones flexibles, la resistencia al estrés, el aprendizaje y la innovación. Estas competencias son la base para gestionar eficazmente la atención, mantener la motivación, colaborar con otros, enfrentar la frustración, responder a los desafíos y evitar comportamientos riesgosos o violentos [2]. En este contexto, la comunidad científica ha explorado métodos para comprender la formación y desarrollo de emociones, con el fin de crear herramientas computacionales en pro de la salud mental y la educación emocional [3], o el tratamiento de trastornos como la hiperactividad [4] y la depresión [5]. El desarrollo de este tipo de plataformas y herramientas requiere de una etapa de clasificación de emociones, que se puede realizar a partir de señales externas, como las expresiones faciales, las posiciones corporales y las voces. Sin embargo, confiar únicamente en estas señales puede ser insuficiente para identificar con precisión las emociones humanas, especialmente cuando están ocultas. En contraste, las señales fisiológicas, entre ellas el electroencefalograma (EEG), destacan por su objetividad, independencia de la apariencia y el comportamiento externos para la detección de emociones. Varios trabajos han demostrado una fuerte correlación entre las señales EEG y los estados emocionales, logrando una gran precisión en la clasificación de emociones [6]. Por lo tanto, el análisis de las señales EEG se presenta como un método confiable y efectivo para el reconocimiento de emociones.

Para la clasificación de emociones con EEG, se utilizan algoritmos de inteligencia artificial (IA), entre ellos, las Redes Neuronales. Dichas redes pueden entrenarse en conjuntos de datos de EEG etiquetados, lo que les permite aprender a reconocer y clasificar las señales eléctricas del cerebro que corresponden a emociones específicas. Al utilizar técnicas de aprendizaje profundo, las redes neuronales pueden capturar relaciones complejas entre las características de los datos de EEG y las emociones experimentadas, lo que lleva a una clasificación más precisa y en tiempo real de las emociones [7]. No obstante, las señales EEG tienen una baja relación señal-ruido por lo que, a menudo, se mezclan con ruido generado durante el proceso de recolección de datos. Además, son no estacionarias, lo que significa que las propiedades de las señales EEG (media, varianza y covarianza) varían con el tiempo, lo que dificulta asociarlas con un estado emocional único. Por último, a estas dificultades se suma la gran variabilidad en las señales de EEG que existe entre diferentes sujetos, lo que complica el entrenamiento de algoritmos de clasificación universales [6]. Todos estos inconvenientes afectan negativamente a la precisión de los algoritmos de clasificación de emociones. El Wavelet Scattering (WS) puede abordar estos problemas proporcionando una representación robusta y estable de las señales EEG. A diferencia de otros métodos de análisis de señales, el WS captura información tanto en el dominio del tiempo

como en el de la frecuencia, lo que le permite identificar patrones relevantes en las señales eléctricas. Además, WS es resistente al ruido y a pequeñas variaciones en las señales, lo que puede ayudar a mitigar los efectos de artefactos y variabilidad inter e intra-sujeto [8]. Al proporcionar características invariantes a traslaciones y estables frente a deformaciones, el WS facilita la extracción de descriptores útiles que pueden mejorar significativamente el rendimiento, la generalización de los algoritmos de IA en la detección de emociones y otras tareas basadas en EEG.

El propósito de este estudio es entrenar un algoritmo de clasificación de emociones utilizando WS en la base de datos DEAP, que se utiliza como banco de pruebas para la clasificación de emociones. Se empleará como clasificador un Multilayer perceptron (MLP), utilizando señales de EEG como entrada, a las cuales se les extraerán características mediante WS. Finalmente, se evaluará el rendimiento del clasificador con 5 k-folds aleatorios y el método Leave One Subject Out (LOSO).

1.1 Objetivos

1.1.1 Objetivo general

Evaluar la posibilidad de mejorar los algoritmos de clasificación de emociones a partir de señales electroencefalograma (EEG) mediante la extracción de descriptores utilizando Wavelet Scattering (WS).

1.1.2 Objetivos específicos

- Preprocesar señales de EEG y extraer Wavelet scattering de sujetos de base de datos DEAP.
- Implementar y entrenar un MLP para la clasificación de emociones en EEG a partir de los descriptores de la señal extraídos con WS.
- Aplicar validación cruzada con 5 k-folds aleatorios y el método Leave-one-subject-out (LOSO) para generar conjuntos de entrenamiento y validación, y calcular métricas como precisión, sensibilidad, especificidad y F1-score para conocer rendimiento del algoritmo.

1.1.3 Alcances

- Se realizará el entrenamiento de los algoritmos en la base de datos DEAP.

1.1.4 Limitaciones

- Algoritmos serán entrenados en Python, a través del entorno de google colab con una GPU T4.

1.2 Temario

(i) Capítulo 1: Introducción en forma general del trabajo, indicando los objetivos y alcances del proyecto.

(ii) Capítulo 2: En esta sección se tiene el marco teórico, en el cual se entrega contexto sobre el tema y las tecnologías que se han desarrollado en el área.

(iii) Capítulo 3: En esta sección se realiza el estudio bibliográfico, en el que se muestra el estado del arte en la clasificación de emociones con EEG.

(iv) Capítulo 4: En esta sección se presentan los resultados obtenidos al aplicar el algoritmo de clasificación de emociones en EEG con WS junto con una discusión que presenta un análisis crítico sobre estos.

(v) Capítulo 5: Se concluye el trabajo realizado.

(vi) Capítulo 6: En este capítulo se abordan los cambios que se deben añadir a la metodología para mejorar resultados a futuro e implementar este tipo de algoritmos en entornos reales.

2 Marco Teórico

2.1 Introducción

En esta sección se proporciona una explicación detallada sobre el entrenamiento y la evaluación de algoritmos para la detección de emociones a partir de señales EEG. Se inicia con una descripción del proceso de diseño de un experimento de detección de emociones, incluyendo la adquisición de las señales EEG. Luego, se exploran las principales teorías actuales para la clasificación de emociones, así como las diferentes bases de datos que se utilizan para comparar la efectividad de los algoritmos de clasificación.

A continuación, se presenta el proceso de extracción de descriptores a partir de las señales EEG, necesarios para el entrenamiento de algoritmos de Deep Learning y Machine Learning. Además, se ofrece una explicación de cómo funcionan estos algoritmos y cómo se aplican en el contexto de la detección de emociones.

Por último, se ofrece se explica cómo aplicar WS para mejorar la representación de las señales EEG y facilitar el entrenamiento de algoritmos de Deep Learning y Machine Learning.

2.2 Experimentos para detección de emociones

El desarrollo de algoritmos de detección de emociones con EEG sigue una serie de etapas específicas. La primera consiste en la creación de un protocolo experimental en el cual se definen los estímulos y la forma en que se medirá la señal de EEG. Luego de esto, se adquieren las señales y se pre procesan para eliminar su ruido y artefactos. Se procede de ahí a extraer las diferentes características de la señal con las cuales se entrenará el modelo. Por último, se selecciona uno o varios algoritmos de clasificación para desarrollar modelos de reconocimiento de patrones capaces de clasificar emociones. En la figura 2.1 se puede observar este proceso.

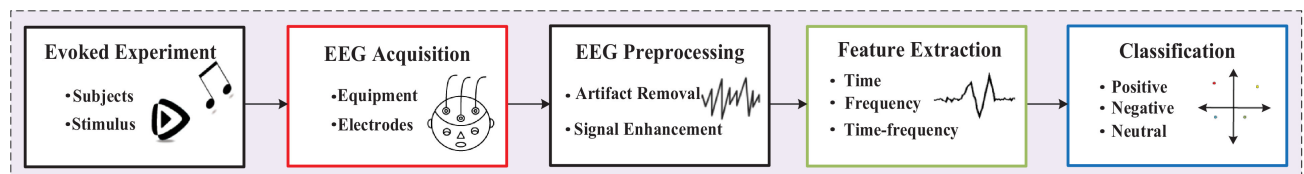


Fig. 2.1: Etapas entrenamiento algoritmo detección de emociones [9].

Fig. 2.2: Sistemas configuración EEG. Negro:10-20, gris: 10-10, blanco: 10-5 [10].

2.4 Teoría clasificación emociones

La detección de emociones se refiere al proceso de identificar, clasificar y comprender las emociones en individuos o grupos de personas. Hay varios enfoques y métodos para la detección de emociones, en este trabajo presentaremos los dos métodos en los que se suelen basar la gran mayoría de trabajos [11].

2.4.1 Enfoque categórico

En el enfoque categórico se clasifican las emociones en las Emociones Básicas de Ekman, las cuales son la felicidad, la tristeza, el miedo, la sorpresa, el enojo y el disgusto. Según el modelo categórico con el que se esté trabajando, estas emociones se subdividen en otros grupos. En la Figura 2.3, se pueden observar las distintas emociones adoptadas por los cinco principales modelos categóricos.

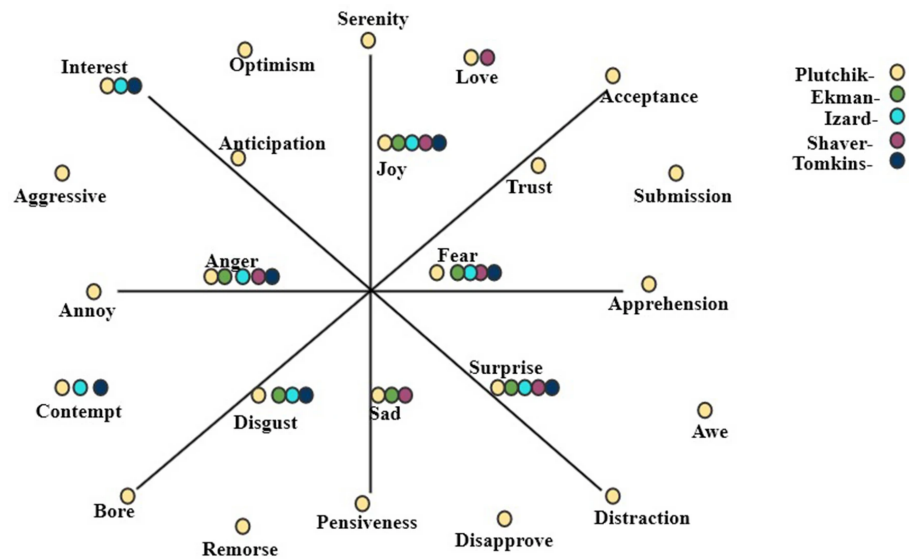


Fig. 2.3: Enfoque categórico [6].

2.4.2 Enfoque dimensional

El enfoque Dimensional describe las emociones por medio de una serie de dimensiones independientes que oscilan entre diferentes escalas dependiendo de la investigación (ejemplo: -1 a 1). Un ejemplo de enfoque dimensional sería el modelo Valence/Arousal/Dominance. Valence determina si una emoción es positiva o negativa, Arousal corresponde al nivel de excitación que genera una emoción, y Dominance, determina el nivel de control que siente la persona sobre su emoción. En la figura

2.4 podemos observar un ejemplo de modelo dimensional tipo Arousal/Valence, aunque en la imagen Valence ha sido reemplazado por positivo/negativo. Además de las 3 dimensiones emocionales ya mencionadas, esta memoria utilizó la dimensión emocional Liking, que determina que tan agradable/desagradable es una emoción.

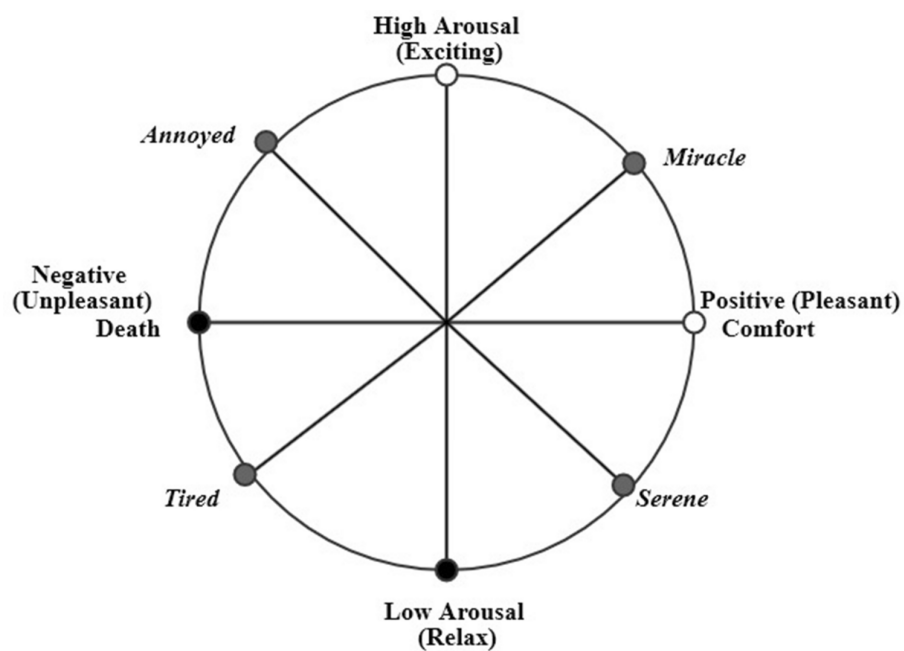


Fig. 2.4: Enfoque Dimensional [6].

2.5 Bases de datos

En el ámbito de la detección de emociones, es esencial contar con bases de datos estandarizadas que faciliten el entrenamiento y la comparación de algoritmos. Estas bases de datos se generan mediante la inducción de estímulos audiovisuales a los participantes, quienes posteriormente, expresan el impacto emocional percibido siguiendo métricas detalladas en la sección de teoría de clasificación de emociones.

Las bases de datos utilizadas con este propósito incluyen [6]:

- DEAPdataset
- MANHOB-HCI
- SEED-IV
- MPED
- DREAMER

En este trabajo, nos centramos en la base de datos DEAP. Esta base de datos consiste en señales EEG y señales fisiológicas periféricas recopiladas de 32 participantes adultos, todas ellas muestreadas a 512 Hz. La obtención de estos datos se llevó a cabo mientras los participantes observaban 40 extractos de vídeos musicales, cada uno con una duración de 1 minuto. La evaluación de los vídeos se realizó mediante el método de clasificación dimensional, donde los participantes categorizaron cada vídeo en términos de Arousal, Valence, Dominance y Like/Dislike. Además, para 22 de los 32 participantes, se registraron vídeos frontales de sus rostros [1].

2.6 Métodos de extracción de descriptores de EEG para entrenamiento de algoritmos

Existen diversos métodos para extraer los descriptores de las señales de EEG. En esta sección se resumirán los que se han utilizado la mayor cantidad de trabajos de clasificación de emociones. Los métodos de extracción de descriptores más utilizados son: Short-time Fourier Transform; Power Spectral Density; Wavelet Transform; Differential Entropy; Wavelet Entropy y Empirical Mode Decomposition, Functional Connectivity Maps [9].

2.6.1 Short-time Fourier Transform (STFT)

La STFT es una técnica utilizada en el procesamiento de señales para analizar la evolución de las frecuencias de una señal a lo largo del tiempo. A diferencia de la Transformada de Fourier clásica, que analiza una señal en su totalidad, la STFT divide la señal en segmentos temporales más pequeños y aplica la Transformada de Fourier discreta (DFT) a cada segmento individualmente. Esto permite capturar cambios en las frecuencias de la señal a medida que evoluciona en el tiempo, proporcionando así una representación más detallada de las características en tiempo y frecuencia de la señal. El uso de esta técnica es útil en el análisis de señales no estacionarias y estocásticas como lo son las señales de EEG. Este método se puede utilizar para extraer bandas de frecuencia de la señal de EEG como las ondas alpha, betha, gamma, theta y delta que pueden servir posteriormente para entrenar modelos de clasificación de emociones o pasarse directamente como dato a un clasificador.

2.6.2 Power Spectral density (PSD)

PSD es una medida en el procesamiento de señales que describe la distribución de la potencia de una señal a través de diferentes frecuencias [12]. En lugar de proporcionar información sobre cómo cambian las frecuencias a lo largo del tiempo, como en el caso de la STFT, PSD se centra en la intensidad de la señal en diferentes bandas de frecuencia específicas. Se obtiene aplicando un método de extracción de bandas de frecuencia (como la DFT o la Wavelet transform) a la señal en toda su duración o en ventanas de tiempo y calculando la magnitud cuadrada de las componentes frecuenciales, proporcionando de esta forma información sobre la distribución de la energía de la señal en el dominio de la frecuencia.

En el análisis de EEG este método puede proporcionar información sobre la potencia de bandas de frecuencia específicas como lo son las de las ondas alfa, beta, delta, theta y gamma. Esta información puede luego ser utilizada por el algoritmo de clasificación para extraer información de la señal.

2.6.3 Differential Entropy (DE)

La entropía es un concepto fundamental en la teoría de la información y la estadística, que mide la incertidumbre o el grado de sorpresa asociado con un conjunto de eventos o una distribución de probabilidad. Cuanto mayor sea la entropía, mayor será la incertidumbre. Si todos los eventos son equiprobables, la entropía alcanza su máximo, indicando máxima incertidumbre. Por otro lado, si un evento es completamente predecible, la entropía es mínima. A diferencia de la fórmula clásica de

entropía, que se calcula de manera discreta, la DE se determina de manera continua. En la clasificación de emociones, este método se aplica generalmente extrayendo primero diferentes bandas de frecuencia de las señales de EEG, como lo son las ondas alfa, beta, delta, theta y gamma, para luego calcular su nivel de DE en ventanas de tiempo [9]. El nivel de DE de la señal nos muestra que tan regular o irregular es la banda de frecuencia de la señal de EEG que estamos estudiando.

2.6.4 Wavelet Transform (WT)

Se utiliza para analizar señales y fenómenos en el dominio del tiempo y la frecuencia de manera simultánea. A diferencia de la Transformada de Fourier, que descompone una señal en sinusoides de diferentes frecuencias, la WT utiliza funciones de onda conocidas como "wavelets" que tienen una duración finita tanto en el dominio del tiempo como en el de la frecuencia. Este método se suele utilizar para separar la señal de EEG en diferentes bandas de frecuencia y luego calcular descriptores de la señal como componentes estadísticos, DE, PSD, etc.

El método de WE corresponde específicamente a utilizar la WT para descomponer la señal en bandas de frecuencia y luego extraer DE.

2.6.5 Empirical Mode Decomposition (EMD)

EMD es una técnica de análisis de señales que descompone una señal compleja en un conjunto de componentes oscilatorios llamados Intrinsic Mode Functions (IMFs) [13]. El proceso comienza identificando las envolventes superior e inferior de la señal original mediante interpolación de sus máximos y mínimos locales. Luego, se resta la media de estas envolventes de la señal original para obtener el primer IMF. Este proceso se repite iterativamente para extraer múltiples IMFs, cada uno representando diferentes escalas de oscilación dentro de la señal. Finalmente, la señal original se descompone en la suma de estos IMFs y un residuo de tendencia.

La EMD es especialmente útil para analizar señales no lineales y no estacionarias, como las ondas de EEG, permitiendo una representación adaptativa y más precisa de la dinámica de la señal. En el contexto de EEG, la EMD se puede utilizar para extraer características relevantes al descomponer las señales en IMFs que capturan diferentes frecuencias y patrones temporales de forma similar a como se utiliza la WT. Estas características pueden luego ser analizadas para identificar correlaciones con estados emocionales, cognitivos, o patológicos, y utilizadas como input en algoritmos de clasificación o detección de eventos específicos en las señales cerebrales.

2.6.6 Functional Connectivity Maps (FC)

FC maps se utilizan en la clasificación de emociones en EEG para capturar la relación y sincronización entre diferentes regiones del cerebro durante la actividad emocional [14]. Estos mapas representan la conectividad entre electrodos, calculada mediante métricas como la coherencia, la correlación o el análisis de fase. Una vez contruidos, los mapas se utilizan como entradas en modelos de clasificación, ya que encapsulan patrones de interacción cerebral que son distintivos de diferentes estados emocionales. Estos patrones permiten a los algoritmos identificar y clasificar emociones al reconocer las características específicas de la conectividad cerebral asociadas a cada emoción.

2.7 Machine Learning

El Machine Learning es un campo de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos computacionales que permiten a las máquinas mejorar su rendimiento en una tarea específica a través de la experiencia. En lugar de ser programadas explícitamente para realizar una tarea, las máquinas utilizan datos y patrones para aprender y tomar decisiones de manera autónoma. El proceso de aprendizaje automático implica la construcción de modelos matemáticos basados en datos de entrenamiento, con el objetivo de hacer predicciones o tomar decisiones sin intervención humana directa. Las aplicaciones del aprendizaje automático son diversas e incluyen tareas como reconocimiento de patrones, clasificación, regresión y clustering. En el contexto de esta memoria el Machine Learning se utiliza para la clasificación de emociones por medio del análisis de señales de EEG.

A continuación se presenta el funcionamiento de 4 algoritmos de Machine Learning que serán explicados en base a lo presentado en el libro [15].

2.7.1 Support Vector Machine (SVM)

SVM es un algoritmo de aprendizaje supervisado utilizado para tareas de clasificación y regresión. La idea principal detrás de SVM es encontrar el hiperplano óptimo que mejor separa las clases en el espacio de características. SVM es conocido por su capacidad para manejar eficientemente conjuntos de datos de alta dimensionalidad y por su versatilidad tanto en problemas de clasificación como de regresión.

2.7.2 Random Forest Regression (RFR)

RFR es una técnica de aprendizaje automático supervisado que se basa en la construcción de múltiples árboles de decisión durante el entrenamiento del algoritmo y utiliza el promedio de las predicciones de estos árboles para obtener un resultado más robusto y preciso. Esta técnica es eficaz para manejar conjuntos de datos complejos y ruidosos, y suele dar buenos resultados en problemas de regresión donde se busca predecir valores numéricos en lugar de clasificar en categorías discretas.

2.7.3 K-means

K-Means es un algoritmo de agrupamiento utilizado en aprendizaje no supervisado. La tarea principal del algoritmo K-Means es dividir un conjunto de datos en k grupos o clusters basándose en características similares. Este proceso de agrupamiento ayuda a identificar patrones y relaciones en los datos sin la necesidad de ejemplos previos como es el caso de los algoritmos de aprendizaje supervisado.

2.7.4 Adaboost

AdaBoost es un algoritmo de aprendizaje automático que construye un clasificador fuerte (modelo que tiene un rendimiento muy alto en la tarea de clasificación) combinando iterativamente clasificadores débiles(modelo que tiene un rendimiento ligeramente mejor que el azar en la tarea de clasificación). Inicia asignando pesos a cada instancia en el conjunto de entrenamiento y entrena un clasificador débil para minimizar el error ponderado total. Luego, ajusta los pesos para dar más importancia a las instancias mal clasificadas, repitiendo este proceso para construir múltiples clasificadores débiles. Los clasificadores se combinan mediante una suma ponderada para formar un clasificador fuerte. AdaBoost es eficaz en mejorar la precisión de modelos débiles y es resistente al sobreajuste, aunque puede ser sensible a ruido y outliers en los datos.

2.8 Redes Neuronales

En esta sección se explican las diferentes redes neuronales que existen actualmente. En el contexto de la detección de emociones, las Redes neuronales se utilizan como clasificadores de emociones. En esta sección también se utiliza la información contenida en el libro [15].

2.8.1 Definición red neuronal

Una red neuronal es un modelo computacional inspirado en la estructura y función del cerebro humano. Consiste en capas de nodos interconectados, llamados neuronas artificiales, que procesan información de entrada, aprenden patrones y generan resultados. Cada conexión entre nodos tiene un peso que ajusta la influencia de una neurona sobre otra. Durante el entrenamiento, la red ajusta estos pesos para mejorar su capacidad para realizar tareas específicas, como reconocimiento de patrones, clasificación o predicción. En esencia, las redes neuronales buscan imitar la capacidad del cerebro para aprender y adaptarse a partir de datos [16]. Los algoritmos que utilizan Redes neuronales con múltiples capas en su entrenamiento, son parte de un grupo específico de algoritmos de Machine Learning, denominados algoritmos de tipo Deep Learning. En la figura 2.5 se observa la aparición en el tiempo de las diferentes Redes Neuronales existentes, las bases de datos utilizadas para el entrenamiento de algoritmos de detección de emociones y los diferentes enfoques psicológicos utilizados para la clasificación de emociones.

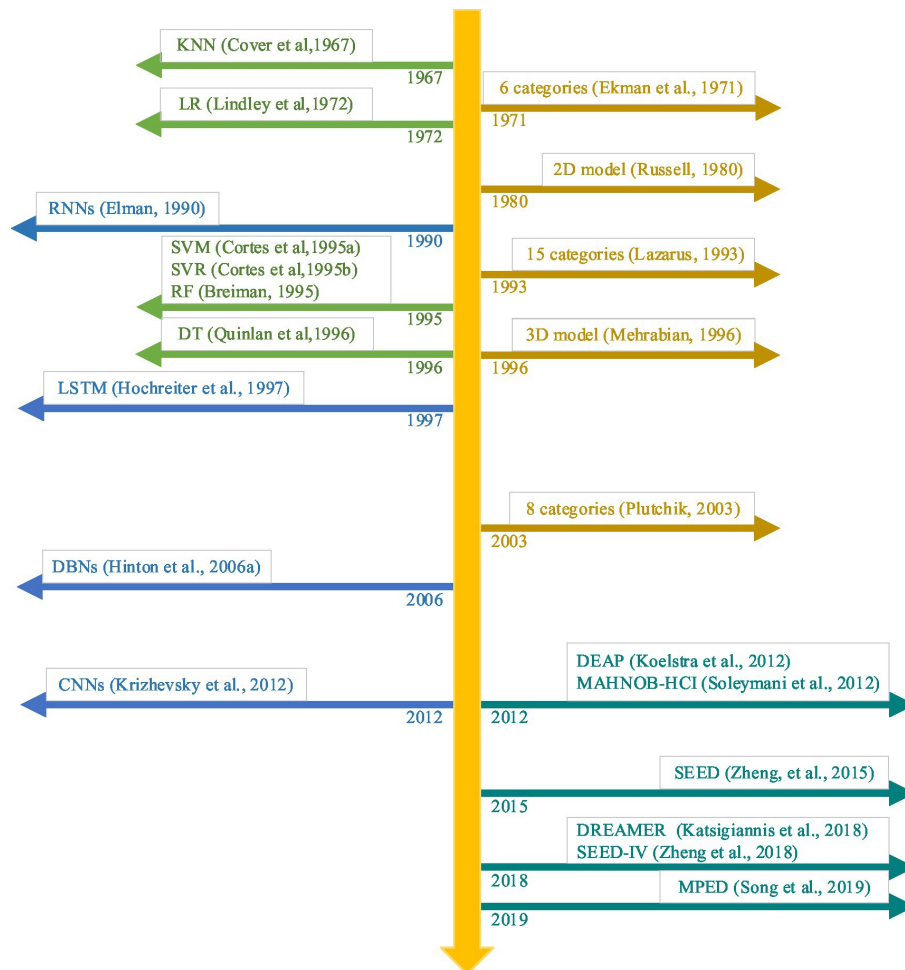


Fig. 2.5: Línea temporal [6].

2.8.2 Multilayer Perceptron (MLP)

Un MLP es un tipo de red neuronal artificial compuesta por múltiples capas de nodos, organizadas en una capa de entrada, una o más capas ocultas, y una capa de salida. Cada nodo (o neurona) en una capa está conectado a los nodos de la siguiente capa mediante pesos sinápticos, y cada conexión tiene un peso ajustable. Los nodos aplican una función de activación para transformar las entradas recibidas antes de pasar la salida a la siguiente capa. El MLP se entrena mediante retropropagación, ajustando los pesos para minimizar el error entre las predicciones de la red y los valores reales. Esta estructura le permite al MLP modelar relaciones complejas y no lineales, siendo ampliamente utilizado en tareas de clasificación, regresión, y reconocimiento de patrones. En problemas de clasificación, se suele combinar con otras estructuras de redes neuronales como las CNN.

2.8.3 Deep belief networks (DBNs)

Las DBNs son un tipo de modelo de aprendizaje profundo que utiliza múltiples capas de nodos interconectados. Estas redes son capaces de aprender representaciones jerárquicas de datos, lo que las hace eficaces para tareas complejas de reconocimiento de patrones. Las DBNs se han aplicado con éxito en diversas áreas, como reconocimiento de voz, gráficos y tareas de clasificación de datos visuales, logrando buenos resultados de reconocimiento de patrones.

2.8.4 Convolutional neural networks (CNNs)

Las Redes Neuronales Convolucionales (CNNs) son un tipo de modelo de aprendizaje profundo diseñado específicamente para procesar datos estructurados, como imágenes y vídeos. Utilizan capas de convolución para extraer características importantes de manera automática. Las CNNs son ampliamente utilizadas en tareas de visión por computadora, como reconocimiento de objetos, segmentación de imágenes y clasificación de imágenes. También se aplican en otras áreas, como procesamiento de señales, traducción automática y análisis de texto, donde la estructura jerárquica de las CNNs permite aprender patrones complejos de manera eficiente.

2.8.5 Recurrent neural networks (RNNs)

En comparación con las CNN, las RNN están diseñadas para trabajar con datos secuenciales, como texto, audio o series temporales. Los investigadores han utilizado ampliamente las RNN en procesa-

miento del lenguaje natural, incluida la traducción automática y la clasificación de texto. Las RNN han logrado un rendimiento excepcional en campos relacionados con el audio, como el reconocimiento y la síntesis del habla. Este tipo de red puede ser ideal para trabajar con señales y por lo tanto es una muy buena opción a la hora de entrenar algoritmos de clasificación en señales de EEG.

2.9 Wavelet Scattering (WS)

En esta memoria se propone utilizar WS como descriptor de las señales de EEG. El WS es una técnica de procesamiento de señales que combina conceptos de las CNN con la WT para extraer características invariantes y estables de señales temporales o espaciales. Esta técnica fue desarrollada para superar algunas limitaciones de los métodos tradicionales de análisis de señales, como la Transformada de Fourier y las transformadas wavelets convencionales, al proporcionar una representación más robusta frente al ruido y las variaciones en la señal [17].

1. **Invariancia y Estabilidad:** Una de las principales motivaciones detrás del desarrollo del WS es obtener características que sean invariantes a traslaciones y estables frente a deformaciones de la señal. La invariancia a traslaciones significa que las características no cambian si la señal se desplaza en el tiempo, mientras que la estabilidad implica que las características son robustas frente a pequeñas perturbaciones o ruidos en la señal.

2. Arquitectura del Wavelet Scattering:

- Primera capa: La señal original se convoluciona con una familia de wavelets de diferentes escalas. Las respuestas de estas convoluciones se someten a una operación de módulo para obtener invariancia a traslaciones locales. El módulo de una señal elimina la información de fase, reteniendo solo la magnitud de las oscilaciones. Esta eliminación de la fase es clave para la invariancia a traslaciones porque la fase es la parte de la señal que cambia con los desplazamientos temporales.
- Capas Subsecuentes: Las salidas de la primera capa se vuelven a convolucionar con wavelets, y nuevamente se aplica el operador de módulo. Este proceso se repite para varias capas, cada vez capturando características de mayor nivel.
- Pooling: Después de cada convolución y operación de módulo, se aplica un promedio a lo largo del tiempo para lograr invariancia a traslaciones globales y reducir la dimensionalidad de las características.

3. Comparación con CNN:

El WS y las CNN comparten similitudes en su capacidad para extraer características jerárquicas y robustas de datos complejos, como señales e imágenes. Ambos enfoques utilizan convoluciones para capturar patrones locales y aplicar operaciones de pooling para reducir la dimensionalidad y lograr invariancia a traslaciones. Sin embargo, el WS ofrece ventajas específicas, como la estabilidad frente a deformaciones y el ruido, y no requiere entrenamiento previo de los filtros, lo que lo hace menos dependiente de grandes cantidades de datos etiquetados. Esta característica es bastante importante a la hora de entrenar clasificadores con señales de EEG, porque no se suele disponer de bases de datos grandes de este tipo de señales.

Las CNN, por otro lado, son más flexibles y potentes para aprender características específicas y complejas cuando se dispone de suficiente información de entrenamiento, permitiendo ajustes finos a través del proceso de aprendizaje profundo.

4. Aplicación en EEG:

En el contexto del análisis de señales EEG (Electroencefalografía), el Wavelet Scattering se utiliza para extraer características que capturan tanto la estructura temporal como la frecuencia de las señales cerebrales. Esto es especialmente útil porque las señales EEG son intrínsecamente ruidosas y varían significativamente entre sujetos y sesiones. Dado que el WS es estable frente a pequeñas perturbaciones, puede manejar mejor el ruido presente en las señales EEG. Además, capturan información en múltiples resoluciones, desde detalles finos hasta patrones globales.

5. Implementación en python:

Existen varias librerías que permiten implementar WS en python. En este trabajo se escogió Kymatio [18] ya que permite la aceleración del cómputo con GPU. Esta biblioteca permite computar el WS dejando modificar 3 parámetros principales:

- J: Número de escalas de la transformación. Este parámetro define la profundidad de la descomposición de la señal en el dominio de la frecuencia. En términos prácticos, representa cuántas veces la señal original será descompuesta en componentes de frecuencia más bajas. Un valor mayor de J permite capturar características de la señal a escalas más grandes (frecuencias más bajas), pero aumenta el tamaño de los coeficientes resultantes y puede requerir más procesamiento. Sin embargo, si J es muy bajo, se pueden perder detalles importantes en escalas más grandes.
- Max order: El orden máximo de los coeficientes de dispersión que se deben calcular, puede ser 1 o 2. El primer orden (order 1) captura las características primarias de la señal, como las variaciones de energía a través de diferentes escalas y posiciones. El segundo orden

(order 2) captura interacciones más complejas entre esas variaciones, analizando cómo cambian las características del primer orden a diferentes escalas.

- Q: De forma predeterminada, Q es el número de wavelets por octava para el primer orden y que para el segundo orden tiene un wavelet por octava. En otras palabras, controla la resolución en el dominio de la frecuencia.

Al introducir un batch de datos bidimensionales de C canales y señales de largo T en la transformada de WS se obtiene una matriz tridimensional de las siguientes dimensiones mostradas en la ecuación 2.9.1. La figura 2.6 muestra como se descomponen los datos con WS cuando se calculan características de orden 2.

$$(B, P, \frac{T}{2^J}) \quad (2.9.1)$$

$$P = 1 + JQ + J(J - 1)Q/2 \quad (2.9.2)$$

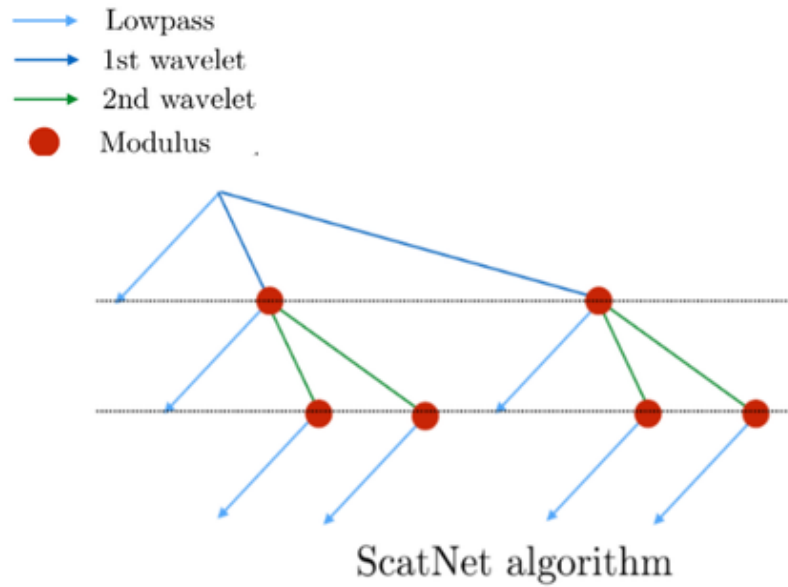


Fig. 2.6: Arquitectura WS[18].

3 Estado del arte

3.1 Introducción

En esta sección, se destacan los algoritmos más relevantes desarrollados hasta la fecha para la detección de emociones. La fase de entrenamiento de estos algoritmos se lleva a cabo tras la extracción de características de la señal. En la tabla 3.1, se presenta una recopilación de algoritmos entrenados utilizando bases de datos estándar, junto con la precisión obtenida en la detección de emociones. Todos estos algoritmos emplean un enfoque dimensional para la clasificación emocional. Los algoritmos se agrupan según el método que utilizan para la partición de los datos: aquellos que utilizan un enfoque generalizado, mezclando los datos de todos los participantes, y aquellos que emplean el método "Leave-one-subject-out"(LOSO), donde los datos se dividen por sujetos, dejando un sujeto en el conjunto de prueba y el resto en el conjunto de entrenamiento.

3.2 Algoritmos que dividen datos de forma generalizada.

Estos algoritmos comparten la característica de dividir los datos originales en segmentos más pequeños (ya sea por ventanas o por canales), los cuales se asignan de manera aleatoria a los conjuntos de entrenamiento y evaluación del clasificador. Este enfoque conlleva el riesgo de contaminación de los datos, ya que es posible que partes de la misma muestra original se incluyan tanto en el conjunto de entrenamiento como en el de prueba.

En [19] se realiza la evaluación del algoritmo en los conjuntos de datos DEAP y DREAMER. La señal se divide en ventanas más pequeñas y no se realiza la extracción de descriptores antes del entrenamiento de la red. Como clasificador emocional, este trabajo propone una Red Neuronal Convolutiva Regional-Asimétrica (RACNN) que consta de extractores de características temporales, regionales y asimétricas. Utiliza capas de convolución 1D continuas para aprender representaciones de tiempo-frecuencia y capas de convolución 2D para capturar información regional entre canales adyacentes físicamente. Además, introduce una Capa Diferencial Asimétrica (ADL) para capturar información discriminativa entre los hemisferios izquierdo y derecho del cerebro. El modelo propuesto obtiene precisiones de reconocimiento superiores al 95 % para las tareas de clasificación de valencia y arousal tanto en DEAP como DREAMER.

[13] utiliza el conjunto de datos DEAP y se enfoca primero en la extracción de 5 bandas de frecuencia por medio de EMD. Luego, se extraen características utilizando la DE y la dimensión fractal de Higuchi (HFD). En la etapa de clasificación, se utilizan cuatro métodos de clasificación: Naïve Bayes, k-nearest neighbor (k-NN), red neuronal convolucional (CNN) y árbol de decisión (DT). Los experimentos muestran una precisión del 95.20 % utilizando el método basado en CNN.

[20] emplea el conjunto de datos DEAP para evaluar su modelo propuesto, una red neuronal profunda basada en una CNN 2D. Esta red utiliza dos filtros de convolución de diferentes tamaños para extraer características emocionales en las dimensiones temporal y espacial a partir de señales de EEG crudas segmentadas en ventanas de 3 segundos. Los resultados obtenidos demuestran precisiones de hasta 99.99 % en la clasificación binaria de arousal y 99.98 % en la de valencia.

[8] utiliza el conjunto de datos DEAP y realiza la extracción de características por canal utilizando WS de orden 2. Luego, se aplica el método de análisis de componentes principales (PCA) para la reducción de dimensionalidad antes de alimentar con datos a los clasificadores. El enfoque propuesto, que combina WS, PCA y un algoritmo SVM, logra más del 98 % de accuracy para dos clases emocionales y más del 97 % de accuracy para tres. Este trabajo logra una precisión en clasificación de emociones alta, demostrando que WS puede ser útil en la clasificación de emociones. Sin embargo, la clasificación de los datos se realiza de forma individual para cada sujeto y no generaliza para el resto de conjunto de datos. En esta memoria se busca ampliar este método al conjunto de datos completo, logrando un algoritmo de clasificación que sea independiente del sujeto. Por esto, se conservó la metodología de extracción de descriptores propuesta pero se cambió el clasificador de los datos por un MLP capaz de trabajar con una mayor cantidad de datos que un algoritmo de SVM.

3.3 Algoritmos que utilizan método LOSO

Los siguientes algoritmos utilizan el método "Leave-one-subject-out"(LOSO) para la evaluación de los clasificadores de emociones. Este método garantiza que el algoritmo de clasificación no dependa del sujeto en el que se entrena, y además, elimina el riesgo de contaminación de datos.

[21] Introduce un enfoque novedoso utilizando mapas FEC combinados con una red CNN para la clasificación. Este método, denominado FEC-CNN, utiliza medidas de Partial directed coherence (PDC) para calcular mapas de Conectividad Efectiva a partir de señales EEG en diferentes bandas de frecuencia. El estudio ajusta seis modelos CNN preentrenados y logra alta precisión y puntuaciones F, particularmente con ResNet-18, que alcanzó hasta un 95.05 % de precisión en la base de datos DEAP.

[22] Utiliza las bases de datos DEAP y MAHNOB-HCI. Las señales EEG se preprocesan y se convierten en representaciones de tiempo-frecuencia mediante la WT. Al igual que [21] emplea una combinación de CNN preentrenadas (AlexNet, ResNet-18, VGG-19 e Inception-v3) y SVM. Las características profundas extraídas de la capa Res2a de ResNet-18 combinadas con SVM fueron evaluadas con el método LOSO y alcanzaron una precisión media del 87.45 % para DEAP y del 77.47 % para MAHNOB-HCI.

[14] Calcula mapas FEC, utilizando Phase Locking Value (PLV) para el estudio del dinamismo en redes cerebrales. El estudio calcula coeficientes de club-rico y utiliza el algoritmo ReliefF para filtrar características y reducir dimensionalidad, logrando un 86.11 % y un 87.92 % de precisión en la dimensión de valencia en los conjuntos de datos DEAP y SEED, respectivamente.

Por último, [23] propone la Red Residual de Desplazamiento Temporal (TSM-ResNet) que convierte los datos EEG en secuencias de imágenes de características 3D para que preserven sus estructuras espaciales, espectrales y temporales. El Módulo de Desplazamiento Temporal (TSM) facilita el intercambio de información entre cuadros adyacentes, mejorando el modelado temporal sin aumentar los costos computacionales. Este enfoque se evalúa utilizando validación cruzada dependiente del sujeto (diez pliegues) y la validación cruzada dejando un sujeto fuera (LOSO). Los resultados muestran una precisión media del 93.43 % para la valencia y del 93.26 % para la excitación en experimentos dependientes del sujeto, y una precisión del 64.91 % para la valencia y del 62.52 % para la excitación en experimentos con el método LOSO.

References	Database	Classification methods	Acc . (%)
[13] (2020)	DEAP	CNN , K - NN , NB , DT	Overall accuracy : 95.20 (using CNN)
[19] (2020)	DEAP and DREAMR	Regional- Asymmetric Convolutional Neural Network (RACNN)	Overall accuracy : 96.65 (Valence) , 97.11 (Arousal)
[20] (2022)	DEAP	2D CNN	Average accuracy : Arousal : 99.99 , valence : 99.98
[21] (2022)	DEAP	Resnet-18	Overall accuracy LOSO: 95.05 %
[14] (2023)	DEAP, SEED	SVM	Overall accuracy LOSO : DEAP: 86.11 % , SEED: 87.92 %
[22] (2023)	DEAP, MANHOB-HCI	Resnet-18 + SVM	Overall accuracy LOSO : DEAP: 87.45 % , MANHOB-HCI : 77.47 %
[23] (2024)	DEAP	TSM-CNN	Overall accuracy: Arousal: 93.43 % Valence: 93.26 % Overall accuracy LOSO: Arousal: 64.91 % Valence: 62.52 %
[8] (2024)	DEAP	SVM	Overall accuracy: 2 categorías: 98 % 3 categorías: 97 %

Tabla 3.1: Revisión algoritmos de Deep Learning utilizados en la detección de emociones

3.4 Discusión

Los datos presentados revelan que los algoritmos que utilizan el método de evaluación LOSO tienen una precisión de clasificación inferior en comparación con aquellos que no lo emplean. Esto sugiere la posibilidad de una contaminación entre los conjuntos de entrenamiento y prueba debido a la metodología de división de los datos. Por esta razón, se evaluará el algoritmo de clasificación utilizando tanto el método general como el método LOSO.

Además, se observa que métodos de extracción de descriptores como EMD, DE, WT y medidas de conectividad han sido ampliamente utilizados en el campo de la detección de emociones. Sin embargo, la Transformada WS no ha sido lo suficientemente estudiada, aunque ha demostrado ser efectiva en la clasificación de emociones de manera dependiente del sujeto, como se muestra en [8]. Este trabajo se propone estudiar el desempeño de WS en la clasificación de datos de manera independiente del sujeto.

Para ello, se plantea: 1) la extracción de descriptores y la reducción de dimensionalidad siguiendo la metodología propuesta en [8], 2) la sustitución del clasificador por un MLP, y 3) la evaluación de la clasificación del algoritmo utilizando tanto el método general como el LOSO.

4 Desarrollo y resultados

En esta sección se muestra de forma detallada la metodología junto con los resultados obtenidos durante el entrenamiento del algoritmo de clasificación de emociones con WS.

4.1 Introducción

El enfoque propuesto en esta tesis, que combina WS como método de extracción de características y un Multilayer Perceptron (MLP) como clasificador, representa una estrategia innovadora para la detección de emociones en señales EEG. La elección de WS se justifica por su capacidad para capturar tanto la estructura temporal como frecuencial de las señales EEG, lo que es esencial para modelar adecuadamente las complejas dinámicas involucradas en la respuesta emocional.

La metodología planteada, que incluye un preprocesamiento de las señales EEG, la aplicación de WS, y la validación cruzada con 5-folds y LOSO, está diseñada para evaluar que el modelo entrenado no solo sea preciso, sino también generalizable a nuevos sujetos. Este enfoque permite evaluar la robustez del modelo y su potencial para ser aplicado en entornos reales donde la variabilidad entre individuos es alta. Todos los códigos utilizados en este trabajo fueron ejecutados en el entorno de Google colab con una GPU T4 y 52.8 GB de memoria RAM. Los resultados que se presentan en esta sección corresponden tanto a los del entrenamiento del algoritmo en 5 k-folds aleatorios como a los del entrenamiento del algoritmo con el método LOSO.

4.2 Metodología

El entrenamiento del modelo de clasificación de emociones siguió las siguientes etapas:

1. Se ordenaron los datos de forma adecuada para entrenar el clasificador.
2. Se utilizó WS para extraer descriptores de las señales de EEG.
3. Se realizó una reducción de la dimensionalidad de los datos con PCA.
4. Se dividió la base de datos en diferentes conjuntos de entrenamiento y validación utilizando 2 métodos: creación de 5 k-folds aleatorios y método LOSO.
5. Entrenamiento del algoritmo en cada uno de estos conjuntos y evaluación de su rendimiento con el cálculo de varias métricas.

A continuación, se explica cada uno de estos pasos.

4.2.1 Preparación set de datos

Se trabajó con las señales preprocesadas entregadas en la base de datos DEAP [1]. El procesamiento que siguieron los creadores de la base de datos siguió el siguiente procedimiento:

- Los datos se redujeron a una frecuencia de muestreo de 128Hz.
- Los artefactos de EOG se eliminaron.
- Se aplicó un filtro de frecuencia de paso de banda de 4.0-45.0Hz.
- Los datos se promediaron con respecto a la referencia común.
- Los canales de EEG se reordenaron para seguir el orden de Geneva (este orden se puede observar en la tabla 6.1 del anexo).
- Los datos se segmentaron en ensayos de 63 segundos.

Se trabajó con 4 dimensiones emocionales (Valence, Arousal, Dominance y Liking), que se encuentran en una escala del 1 al 9. Para facilitar la clasificación, se estableció como 0 (LOW) los valores entre 1 y 4, y como 1 (HIGHT) los valores entre 5 y 9. En la tabla 4.1 se muestran las dimensiones de los datos preprocesados.

Data	Dimensiones	Contenido
Datos EEG	40 x 32 x 8064	video x canales x data
Etiquetas emocionales	40 x 4	video x etiqueta

Tabla 4.1: Contenido video de cada participante

4.2.2 Normalización y extracción de descriptores con WS

Se comienza normalizando las señales de EEG usando la función Minmaxscaler de Sklearn. De esta manera, las señales quedan con media 0 y desviación estándar 1, lo que facilita la clasificación.

Luego, se realizó la extracción de descriptores de la señal siguiendo la metodología utilizada por [8]. Para esto, se comenzó concatenando los datos de todos los participantes he utilizando como **datos independientes** cada uno de los canales originales de 8064 puntos. El conjunto de datos final forma de esta manera una matriz de dimensiones 40960 x 8064 ((participantes x canales) x puntos). El WS se aplica a todos los canales de forma independiente, convirtiendo cada una de las señales de 8064 puntos en matrices de 581 x 126 puntos (canales x componentes asociados a frecuencia x componentes asociados al tiempo). Los hyerparámetros que se utilizaron para la transformación de WS correspondieron a $J = 6$, $Q = 100$ y $\text{max order} = 2$. Todos los cálculos se realizaron con la biblioteca kymatio.

En la 4.1 se observan las dimensiones de los datos antes y después de aplicar WS.

Etapas	Dimensiones	Contenido
Datos EEG antes de WS	1280 x 32 x 8064	(video x participante) x canal x data
Datos después de WS	40960 x 581 x 126	(video x participante x canal) x alto x ancho

Tabla 4.2: Dimensiones datos al aplicar WS

4.2.3 Creación de conjuntos de entrenamiento/validación y reducción de dimensionalidad

El conjunto de datos de entrenamiento se dividió para ser evaluado de 2 formas diferentes:

1. **5 k-folds de entrenamiento/validación** dejando 80 % de los datos para el entrenamiento del algoritmo y 20 % para la evaluación de resultados. Para asegurar que los conjuntos de datos

de entrenamiento/validación fueran representativos del conjunto de datos original, se utilizó el método de validación cruzada estratificada K-fold de la biblioteca Python Scikit-learn. Este método garantiza que la proporción entre las categorías que se desean clasificar se mantenga igual que en el conjunto de datos original. Además, se seleccionaron los datos de cada grupo de manera aleatoria, generando así 5 combinaciones diferentes de datos para cada conjunto. Esto asegura que la precisión de los clasificadores no esté influenciada por el orden de los datos.

2. Usando el método **Leave one subject out (LOSO)**, que consiste en dejar únicamente un participante en el conjunto de prueba y el resto en el conjunto de entrenamiento. Esto se realizó con todos los participantes.

Las dimensiones de los conjuntos de entrenamiento/validación fueron las siguientes:

- Conjunto de entrenamiento: 32768 x 581 x 126 (muestra x alto x ancho)
- Conjunto de validación: 8192 X 581 x 126 (muestra x alto x ancho)
- Conjunto de entrenamiento LOSO: 39680 x 581 x 126 (muestra x alto x ancho)
- Conjunto de validación LOSO: 1280 X 581 x 126 (muestra x alto x ancho)

Después de dividir los datos, se aplicó una reducción de dimensionalidad utilizando Análisis de Componentes Principales (PCA) en el que seleccionaron únicamente 973 componentes principales, ya que estas fueron suficientes para conservar el 99.99 % de la variabilidad total de los datos originales. Para esto, en primer lugar, se aplanaron las matrices obtenidas luego de aplicar WS a los datos pre-procesados con la función reshape de la biblioteca numpy. Luego, el algoritmo PCA fue ajustado en el conjunto de entrenamiento. Por último, se aplicó la transformación para la reducción de dimensionalidad tanto en el conjunto de entrenamiento como en el conjunto de validación.

El resultado de aplicar PCA fueron muestras con las dimensiones mostradas en la tabla 4.3.

Etapa	Conjunto	Dimensiones	Contenido
Datos después de WS	Entrenamiento	nt x 581 x 126	muestra x alto x ancho
	Validación	nv x 581 x 126	muestra x alto x ancho
Reshape	Entrenamiento	nt x 73206	muestra x (alto x ancho)
	Validación	nv x 73206	muestra x (alto x ancho)
PCA (componentes=974)	Entrenamiento	nt x 974	muestra x número de componentes
	Validación	nv x 974	muestra x número de componentes

Tabla 4.3: Dimensiones de datos al aplicar PCA. nt denota la cantidad de muestras utilizadas para entrenamiento, dependiendo de si se usa método de 5 k-folds o LOSO. nv denota la cantidad de muestras utilizadas para validación, dependiendo de si se usa método de 5 k-folds o LOSO

4.2.4 Arquitectura del clasificador

Se utilizó un Multilayer Perceptron (MLP) como clasificador, el cual recibe directamente los datos después de aplicarles una reducción de dimensionalidad mediante PCA. La arquitectura del MLP entrenado se detalla en la tabla 4.4.

Layer	Input Size	Output Size
Linear (fc1)	974	475
Dropout (drp=0.5)	475	475
Batch Normalization	475	475
Linear (fc2)	475	1

Tabla 4.4: Arquitectura clasificador de emociones

Los hiperparámetros seleccionados para la red MLP se presentan en la tabla 4.5. Dado que el tamaño del batch es de 128, el tamaño del input en cada iteración del entrenamiento es de 128 x 974. Como se muestra en la tabla 4.4, se incluye una capa de batch normalization para reducir el riesgo de que los valores de los pesos de las neuronas sean demasiado grandes o pequeños, facilitando así el uso de tasas de aprendizaje más altas. Además, se añade una capa de dropout para prevenir el sobreajuste del modelo a los datos de entrenamiento.

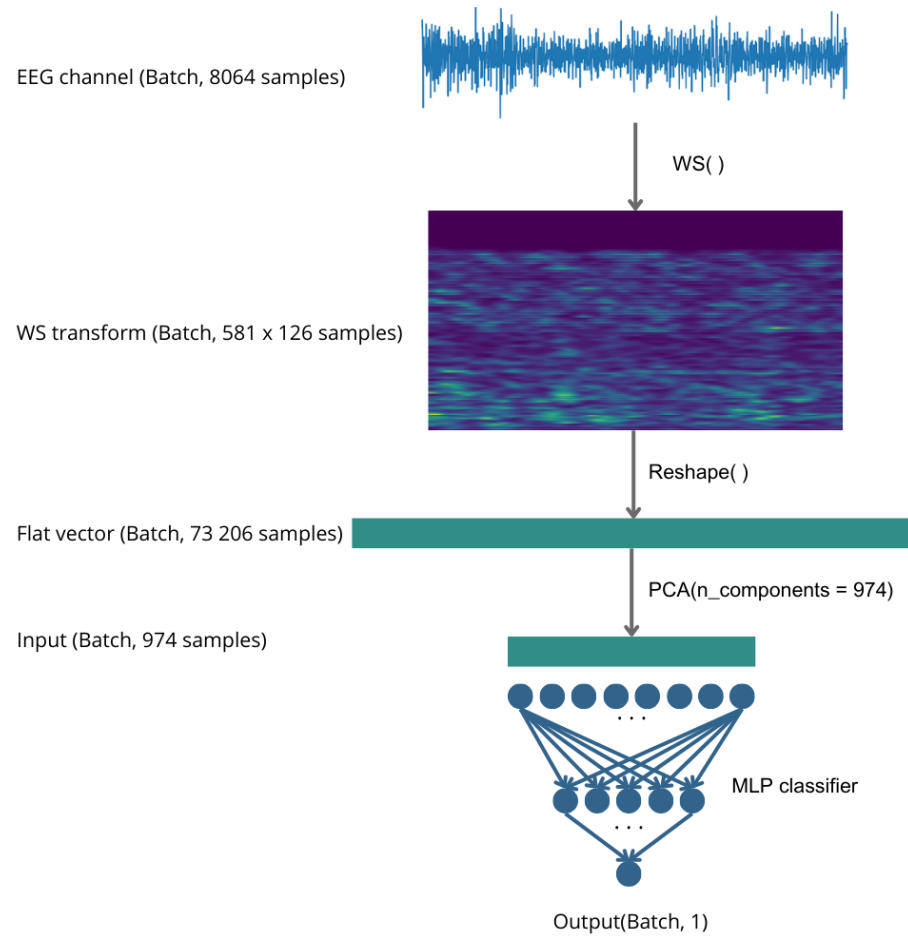


Fig. 4.1: Esquema de entrenamiento

4.2.5 Entrenamiento y evaluación de clasificadores

Para evaluar el rendimiento del modelo, se calcularon las siguientes métricas: **accuracy**, **precision**, **recall** y **F1-score**. Estas métricas son ampliamente utilizadas en la evaluación de modelos de clasificación, permitiendo medir la calidad de las predicciones realizadas por el modelo en diferentes aspectos.

- **Accuracy:** La **accuracy** (exactitud) es la proporción de predicciones correctas entre el total de predicciones realizadas. Se calcula como:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

donde:

- TP (True Positives): Verdaderos positivos

Hyperparameter	Value
Learning Rate	0.001
Weight Decay	0.0001
Batch Size	128
Epochs	30
Criterion	BCE Loss
Optimizer	Adam

Tabla 4.5: Hiperparámetros del MLP

- TN (True Negatives): Verdaderos negativos
- FP (False Positives): Falsos positivos
- FN (False Negatives): Falsos negativos

- **Precision:** La **precision** (precisión) mide la proporción de verdaderos positivos sobre el total de predicciones positivas realizadas por el modelo. Se calcula como:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall:** El **recall** (sensibilidad o tasa de verdaderos positivos) mide la proporción de verdaderos positivos que fueron correctamente identificados entre todos los casos que eran realmente positivos. Se calcula como:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1-score:** El **F1-score** es la media armónica entre la **precision** y el **recall**. Esta métrica es útil cuando se busca un balance entre precisión y sensibilidad. Se calcula como:

$$F\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Estas métricas proporcionarán una visión completa del rendimiento del modelo, permitiendo identificar no solo cuán precisas son las predicciones, sino también la capacidad del modelo para identificar correctamente las clases positivas y negativas.

Además del cálculo de estas métricas, también se graficaron las funciones de exactitud y pérdida en función de las épocas transcurridas durante el entrenamiento.

4.3 Resultados

En las tablas 4.6 y 4.7, se muestran los valores de Accuracy, Precision, Recall y F1-Score para las dimensiones de Valence, Arousal, Dominance y Liking. Estos indicadores reflejan el desempeño del modelo en la clasificación de emociones a partir de señales EEG con método de evaluación generalizado de 5 k-folds aleatorios.

- Valence: El modelo muestra un rendimiento muy alto en la clasificación de valencia, con un Accuracy de 0.956, una Precision de 0.956, un Recall de 0.955 y un F1-Score de 0.955. Estos valores sugieren que el modelo clasifica con precisión y de manera consistente las emociones en términos de valencia. Sin embargo, al evaluar el algoritmo con el método LOSO, se observa un rendimiento cercano al 50 % en todas las métricas.
- Arousal: Para la dimensión de Arousal, el rendimiento es ligeramente inferior al de Valence, con un Accuracy de 0.955 y un F1-Score de 0.953. La Precision y el Recall son también altos (0.955 y 0.952, respectivamente), lo que indica que el modelo mantiene una buena capacidad para identificar correctamente las emociones relacionadas con el nivel de activación. En este caso, al evaluar el algoritmo con el método LOSO se observa lo mismo que con el accuracy, un rendimiento en la clasificación muy bajo, cercano al 50 %.
- Dominance: El modelo también muestra un rendimiento robusto en la clasificación de dominancia al evaluar el algoritmo con 5 k-folds aleatorios, con un Accuracy de 0.955 y un F1-Score de 0.954. La Precision es de 0.955 y el Recall es de 0.954. En el caso de la evaluación con el método LOSO, se obtienen valores en todas las métricas cercanos al 50 %, lo que indica un mal rendimiento del algoritmo.
- Liking: En la clasificación de la dimensión de Liking con 5 k-folds, el modelo presenta un rendimiento similar al de Dominance, con un Accuracy de 0.955 y un F1-Score de 0.954. La Precision y el Recall también se mantienen altos (0.955 y 0.954, respectivamente). En el caso de la evaluación con el método LOSO, se obtienen valores en todas las métricas cercanos al 50 %, lo que indica un mal rendimiento del algoritmo.

El realizar la evaluación con 5 k-folds aleatorios, el modelo muestra un rendimiento altamente consistente y preciso en la clasificación de todas las dimensiones emocionales. Los valores de F1-Score, cercanos o superiores a 0.95 en todas las dimensiones. Sin embargo, al evaluar el modelo con el método LOSO, los valores de Accuracy, precisión, recall y F1-Score están en el rango de 0.53 a

0.57, indicando que el modelo no es eficiente en ninguna de las dimensiones emocionales. Esto puede ser una indicación de que el modelo necesita mejoras adicionales, ya sea en términos de arquitectura, técnicas de preprocesamiento o la cantidad/calidad de los datos utilizados.

	Accuracy	Precision	Recall	F1-Score
Valence	0.9562	0.9557	0.9549	0.9552
Arousal	0.9545	0.9546	0.952	0.9530
Dominance	0.9552	0.9548	0.9537	0.9542
Liking	0.9549	0.9549	0.9539	0.9544

Tabla 4.6: Resultados Promedio de Clasificación en 5 K-Folds para las 4 Dimensiones Emocionales

	Accuracy	Precision	Recall	F1-Score
Valence	0.5717	0.5587	0.5489	0.5390
Arousal	0.5734	0.5629	0.5588	0.5568
Dominance	0.5684	0.5545	0.5444	0.5326
Liking	0.5732	0.5618	0.5564	0.5529

Tabla 4.7: Resultados Promedio de Clasificación con LOSO para las 4 Dimensiones Emocionales

A continuación, se evalúa el comportamiento de las curvas de entrenamiento del algoritmo en las 4 dimensiones originales mostradas en las figuras del anexo 4.2, 4.3, 4.4 y 4.5 para la evaluación con 5 k-folds, y en las figuras 4.6, 4.7, 4.8 y 4.9 en el caso de la evaluación con el método LOSO.



Fig. 4.2: Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Valence.

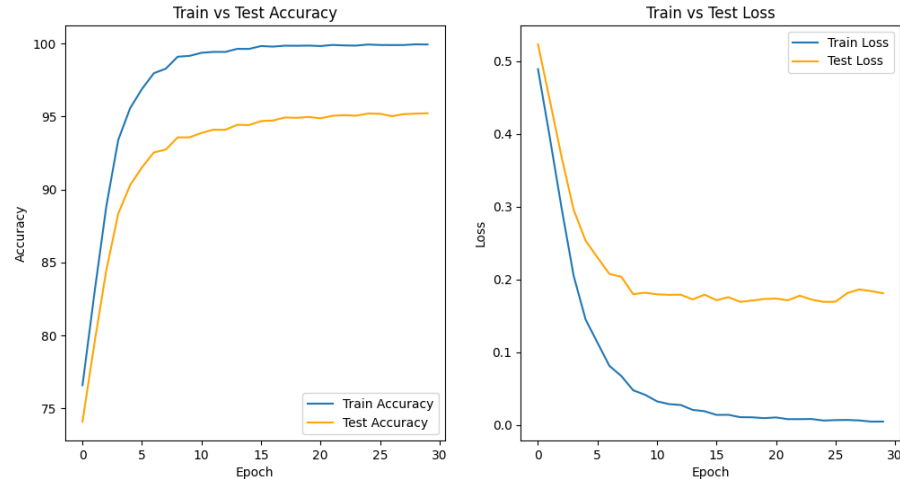


Fig. 4.3: Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Arousal.

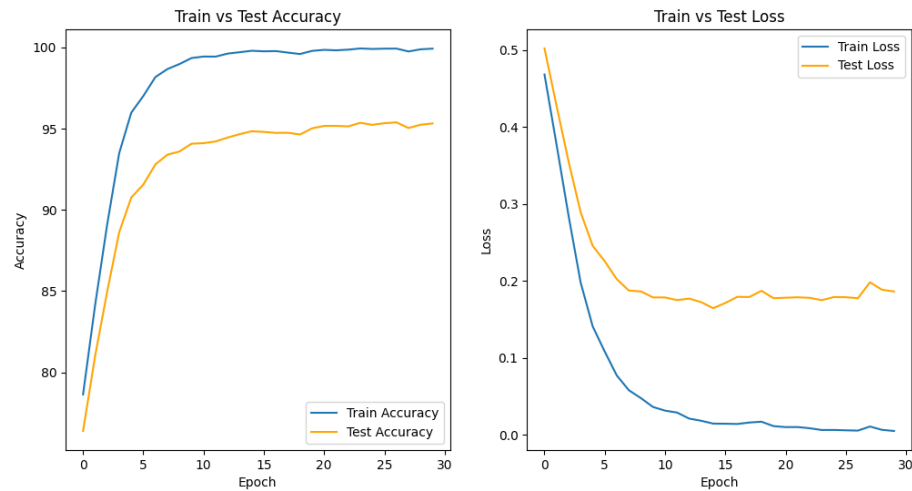


Fig. 4.4: Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Dominance.

4.3.1 Curva de Accuracy

- **Comportamiento del Accuracy de Entrenamiento:** En el caso del método de 5 k-folds, la curva de Accuracy durante el entrenamiento muestra un crecimiento rápido al inicio, alcanzando un valor cercano al 100 % después de unas 10 épocas y manteniéndose estable a partir de ahí. Este comportamiento sugiere que el modelo está aprendiendo bien los patrones de los datos de entrenamiento, ajustándose cada vez más a estos datos a medida que avanza el entrenamiento. Al aplicar el método LOSO, se observa que la Accuracy en el conjunto de entrenamiento mejora de manera constante a lo largo de las épocas, alcanzando valores cercanos al 90 %. Esto indica que el modelo está aprendiendo bien en el conjunto de entrenamiento.
- **Comportamiento del Accuracy de Prueba para las 4 dimensiones emocionales:** En el caso del método de 5 k-folds la precisión en los datos de prueba también aumenta rápidamente al princi-

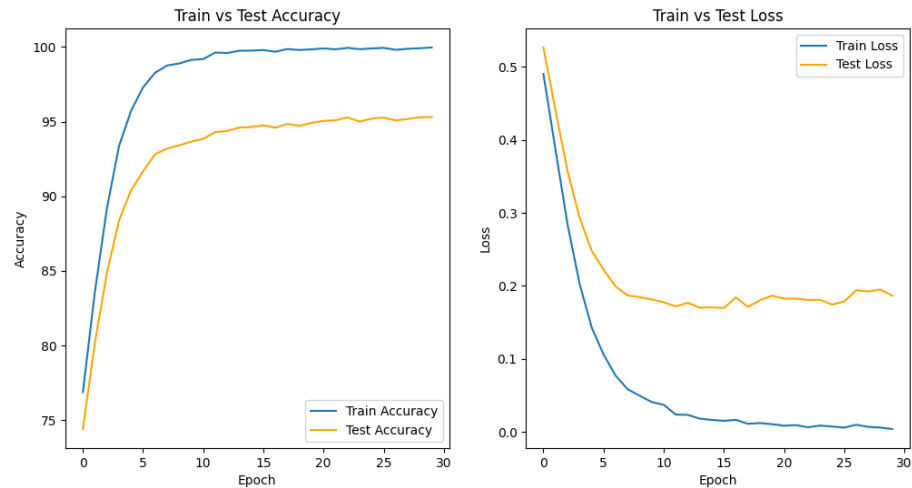


Fig. 4.5: Accuracy y pérdida media en 5 k-folds para la clasificación del nivel de Liking.

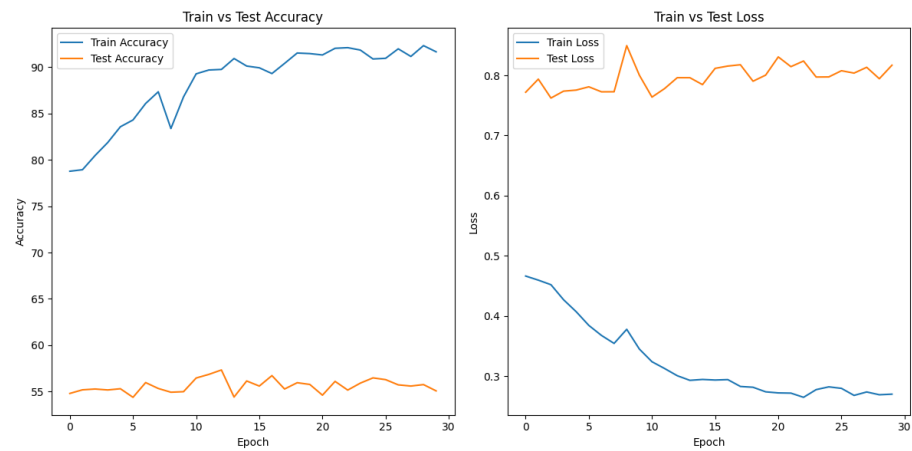


Fig. 4.6: Accuracy y pérdida media al aplicar método LOSO en Valence.

pio, pero se estabiliza alrededor del 95 % después de aproximadamente 10 épocas. La diferencia entre las curvas de Accuracy de entrenamiento y prueba, con la precisión de prueba estabilizándose por debajo de la de entrenamiento, sugiere que el modelo podría estarse sobreajustando a los datos de entrenamiento. Aunque el modelo generaliza bien, dado que la precisión de prueba es alta, el margen entre las dos curvas indica una posible sobreadaptación a las características específicas de los datos de entrenamiento. El realizar la evaluación con el método LOSO, La precisión en el conjunto de prueba se mantiene relativamente baja y fluctuante, alrededor del 55 %. Esto sugiere que el modelo no está generalizando bien a los datos de prueba.



Fig. 4.7: Accuracy y pérdida media al aplicar método LOSO en Arousal.

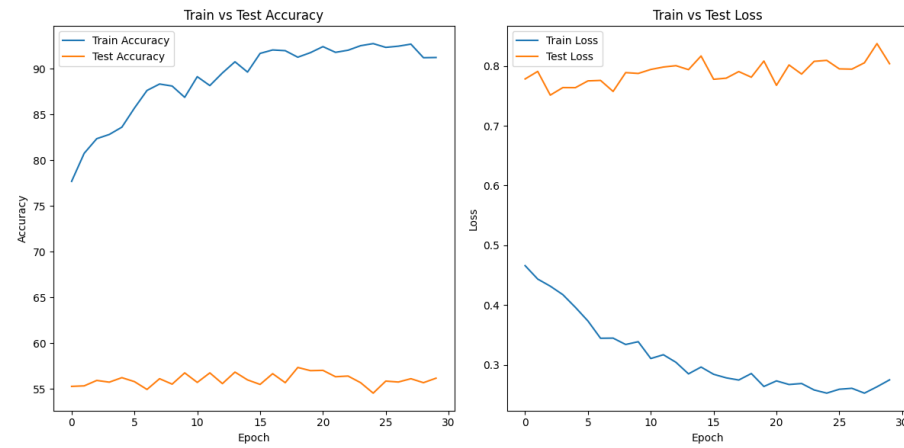


Fig. 4.8: Accuracy y pérdida media al aplicar método LOSO en Dominance.

4.3.2 Comportamiento de la Pérdida en Entrenamiento en las 4 dimensiones emocionales:

En el caso del entrenamiento con 5 k-folds, la curva de pérdida en el conjunto de entrenamiento desciende rápidamente en las primeras épocas, lo que es un indicador de que el modelo está aprendiendo eficazmente a minimizar el error. Sin embargo, la curva continúa descendiendo incluso cuando la pérdida en los datos de validación se ha estabilizado, lo que refuerza la idea de un posible sobreajuste.

En el caso del uso del método LOSO, la pérdida en el conjunto de entrenamiento disminuye de manera constante, lo que es un comportamiento esperado y consistente con el aumento en la precisión de entrenamiento. Sin embargo, la pérdida en el conjunto de prueba se mantiene alta y con fluctuaciones, lo que indica que el modelo está teniendo dificultades para predecir correctamente en el conjunto de prueba.

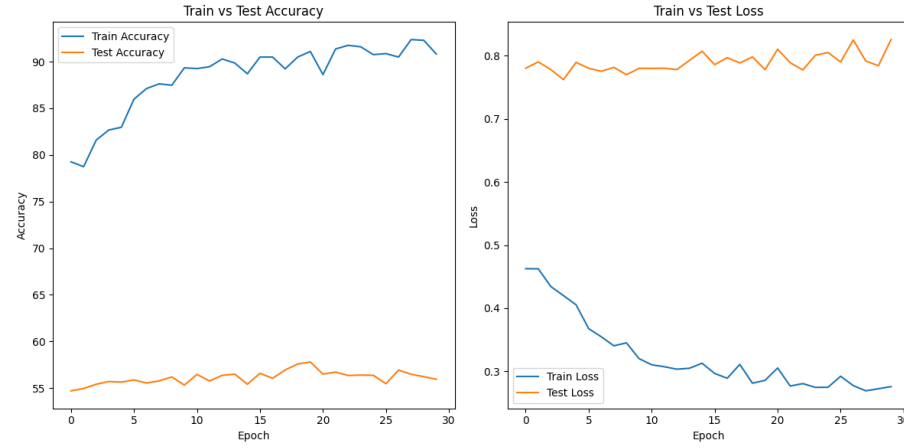


Fig. 4.9: Accuracy y pérdida media al aplicar método LOSO en Liking.

4.4 Discusión

4.4.1 Comparación con estado del arte:

El análisis de los resultados presentados en la tabla 4.8 revela que el método propuesto, que emplea una MLP y es evaluado en 5 k-folds aleatorios, logra una precisión general del 95.4 % en la clasificación de emociones a partir de datos EEG de la base de datos DEAP. Aunque la 2D CNN utilizada por [20] reporta precisiones superiores (99.99 % para arousal y 99.98 % para valencia), nuestro método muestra un rendimiento notablemente superior al de [23] y comparable al de [13], [8], y [19], alcanzando todas precisiones superiores al 95 %. La comparación con [8] es especialmente relevante, ya que comparte una metodología de extracción de descriptores similar a la de este trabajo. Aunque nuestros resultados son un 3 % inferiores a los reportados por [8], este estudio demuestra que es posible aplicar esta metodología de extracción de características manteniendo alta precisión en la clasificación binaria de emociones sin necesidad de entrenar un clasificador independiente para cada sujeto.

No obstante, la tabla 4.9 muestra que el algoritmo de clasificación propuesto en esta tesis obtiene el resultado más bajo, con una precisión del 55.7 %. Esto se debe a que, en un principio, no se planteó utilizar este método como técnica de evaluación, aunque se decidió incluirlo posteriormente, considerando su importancia para evaluar un algoritmo que tenga en cuenta las diferencias individuales entre sujetos. Por lo tanto, en futuros trabajos será necesario implementar esta metodología desde el inicio del entrenamiento.

References	Database	Classification methods	Acc . (%)
[13] (2020)	DEAP	CNN , K - NN , NB , DT	Overall accuracy : 95.20 (using CNN)
[19] (2020)	DEAP and DREAMR	Regional- Asymmetric Convolutional Neural Network (RACNN)	Overall accuracy : 96.65 (Valence) , 97.11 (Arousal)
[20] (2022)	DEAP	2D CNN	Average accuracy : Arousal : 99.99 , valence : 99.98
[23] (2024)	DEAP	TSM-CNN	Overall accuracy: Arousal: 93.43 % Valence: 93.26 % Overall accuracy
[8] (2024)	DEAP	SVM	Overall accuracy: 2 categorías: 98 % 3 categorías: 97 %
Método propuesto (2024)	DEAP	MLP	Overall accuracy: 95.4 %

Tabla 4.8: Algoritmo propuesto en comparación con estado del arte

4.4.2 Fortalezas y debilidades del clasificador propuesto

El algoritmo probado demostró ser efectivo en la clasificación de emociones al utilizar el método generalizado, en el cual cada uno de los canales de EEG se evalúa como un dato independiente y se forman 5 k-folds aleatorios para la validación. Sin embargo, esta alta precisión en la clasificación no se mantuvo al aplicar el método LOSO. Este comportamiento podría deberse a dos problemas presentes en la metodología de evaluación basada en canales independientes con 5 k-folds:

- Diferencias individuales entre los participantes.
- Correlación entre los canales de EEG.

Las señales de EEG suelen variar considerablemente entre diferentes participantes, y el método LOSO destaca esta característica al evaluar el rendimiento del algoritmo de manera independiente para cada sujeto. Además, es probable que el aparente buen desempeño del algoritmo bajo el enfoque generalizado se deba a una contaminación en los datos, ya que al considerar los canales como datos independientes, es posible que canales de una misma muestra se encuentren tanto en el conjunto de

References	Database	Classification methods	Acc . (%)
[21] (2022)	DEAP	Resnet-18	Overall accuracy LOSO: 95.05 %
[14] (2023)	DEAP, SEED	SVM	Overall accuracy LOSO : DEAP: 86.11 % , SEED: 87.92 %
[22] (2023)	DEAP, MANHOB-HCI	Resnet-18 + SVM	Overall accuracy LOSO : DEAP: 87.45 % , MANHOB-HCI : 77.47 %
[23] (2024)	DEAP	TSM-CNN	Overall accuracy LOSO: Arousal: 64.91 % Valence: 62.52 %
Método propuesto			
(2024)	DEAP	MLP	Overall accuracy LOSO: 55.7 %

Tabla 4.9: Algoritmo propuesto en comparación con estado del arte método LOSO

entrenamiento como en el de prueba. De esta forma, dado que los canales están correlacionados entre sí, esto podría mejorar artificialmente el rendimiento del algoritmo.

Por lo tanto, para comprobar la efectividad de la metodología WS en la clasificación de emociones, es necesario ajustar la metodología de evaluación de los algoritmos de EEG. Esto implica garantizar que no haya contaminación de datos entre los conjuntos de entrenamiento y validación, y asegurar que el algoritmo funcione de manera efectiva en diferentes participantes utilizando el método LOSO.

Además, se debe probar el uso de otros tipos de clasificadores como CNN o RNN, que podrían ser más adecuados para capturar las complejas relaciones espaciales y temporales presentes en las señales EEG. La implementación de CNNs podría ayudar a explotar la estructura espacial inherente de las señales, mientras que las RNNs podrían ser beneficiosas para modelar las dependencias temporales en las secuencias de EEG. Estas alternativas podrían ofrecer mejoras significativas en la clasificación de emociones, especialmente bajo metodologías de evaluación más estrictas como el método LOSO.

Finalmente, es importante considerar que la combinación de diferentes modelos y la utilización de técnicas de ensamble podrían potenciar el rendimiento general del sistema, aprovechando las fortalezas individuales de cada clasificador. Esto refuerza la necesidad de explorar no solo diferentes modelos, sino también distintas estrategias de fusión y preprocesamiento que optimicen la clasificación de emociones en entornos con alta variabilidad entre sujetos.

5 Conclusión

El presente trabajo investigó la efectividad de un algoritmo de clasificación de emociones que integra la Transformada Wavelet Scattering (WS) para la extracción de descriptores de señales EEG, junto con PCA para la reducción de dimensionalidad y un MLP para la clasificación binaria de la intensidad emocional en las dimensiones de Valence, Arousal, Dominance y Liking. Los resultados obtenidos bajo el método de validación cruzada de 5 k-folds mostraron un rendimiento consistente y elevado, con métricas de precisión, recall y F1-Score superiores al 95 % en todas las dimensiones emocionales.

Sin embargo, al aplicar el método LOSO, el rendimiento del modelo disminuyó notablemente a un 55 %, lo que sugiere que la metodología de evaluación basada en canales independientes con 5 k-folds podría estar generando una contaminación de datos entre los conjuntos de entrenamiento y prueba. Esta contaminación, probablemente originada por la correlación entre canales de EEG y el uso de conjuntos que consideran a todos los participantes, podría haber aumentado artificialmente los resultados obtenidos con k-folds.

Estos hallazgos subrayan la importancia de utilizar métodos de evaluación robustos, como LOSO, que permitan un análisis más fiel del desempeño del modelo en contextos reales, donde la variabilidad entre sujetos es un factor crucial. A futuro, será esencial ajustar la metodología de evaluación para evitar la contaminación de datos y asegurar que el algoritmo pueda generalizar adecuadamente en diferentes sujetos. Esto no solo contribuirá a una mejor comprensión del potencial de WS en la clasificación de emociones, sino que también garantizará la validez y confiabilidad de los modelos desarrollados para aplicaciones prácticas.

6 Trabajo futuro

Los resultados obtenidos en este estudio destacan la necesidad de ajustar la metodología de evaluación para evitar la contaminación de datos entre los conjuntos de entrenamiento y prueba. Una posible dirección a futuro sería la implementación de estrategias que aseguren la independencia de los datos utilizados en ambos conjuntos. Esto podría lograrse mediante técnicas que tengan en cuenta la correlación entre canales EEG o que separen más claramente los datos de los distintos sujetos.

Otra dirección importante para futuros trabajos es la mejora del modelo mediante la incorporación de técnicas de regularización avanzadas, como la regularización L1/L2, que podrían mejorar la capacidad de generalización del modelo. También sería valioso explorar la optimización de los hiperparámetros del MLP y del PCA, así como experimentar con otras técnicas de reducción de dimensionalidad que se adapten mejor a los datos EEG. El uso de arquitecturas más complejas, como redes neuronales recurrentes (RNNs) o redes neuronales convolucionales (CNNs), podría permitir capturar de manera más efectiva la temporalidad y la relación espacial entre los canales de EEG, mejorando así la precisión de la clasificación emocional.

La obtención de un conjunto de datos más extenso y diverso es otra prioridad para futuros estudios. La recopilación de datos EEG de una población más amplia, que incluya variabilidad en términos de edad, género y condiciones experimentales, ayudaría a reducir el riesgo de sobreajuste y a mejorar la capacidad del modelo para generalizar a nuevos sujetos. Además, la integración de señales multimodales, como Electrooculograma, Electromiografía o datos de imagen, podría proporcionar información complementaria valiosa, mejorando la precisión de la clasificación emocional al combinar diferentes tipos de datos fisiológicos.

Finalmente, un objetivo clave a largo plazo es la implementación del modelo en aplicaciones de tiempo real, especialmente en dispositivos con recursos computacionales limitados, como dispositivos móviles o sistemas embebidos. Esto requeriría optimizar el procesamiento de señales y la inferencia del modelo para funcionar eficientemente en estos entornos. Asimismo, la validación del modelo en entornos naturales y no controlados proporcionaría una comprensión más profunda de su robustez y aplicabilidad práctica, facilitando su uso en áreas como la educación, la salud mental o la interacción humano-computadora.

Glosario

CNN : Convolutional Neural Network

DBN : Deep beliefs networks

DE : Differential entropy

EEG: Electroencefalograma

EMD : Empirical mode decomposition

FC: Functional connectivity maps

LOSO : Leave One subject out

MLP : Multilayer perceptron

PCA : Principal component analysis

PSD : Power spectral density

RFR : Random Forest Regression

RNN : Recurrent neural network

SVM : Support vector Machine

WS : Wavelet scatttering

WT : Wavelet Transform

Bibliografía

- [1] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, “DEAP: A database for Emotion Analysis; using physiological signals,” *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.
- [2] P. C. Broderick and P. A. Jennings, “Mindfulness for adolescents: A promising approach to supporting emotion regulation and preventing risky behavior,” *New directions for youth development*, no. 136, pp. 111–126, 2012.
- [3] Y. Koush, D.-E. Meskaldji, S. Pichon, G. Rey, S. W. Rieger, D. E. Linden, D. Van De Ville, P. Vuilleumier, and F. Scharnowski, “Learning control over emotion networks through connectivity-based neurofeedback,” *Cerebral cortex*, vol. 27, no. 2, pp. 1193–1202, 2017.
- [4] M. Y. Melnikov *et al.*, “The current evidence levels for biofeedback and neurofeedback interventions in treating depression: a narrative review,” *Neural Plasticity*, 2021.
- [5] E. Kuznetsova, A. V. P. Veilahti, R. Akhundzadeh, S. Radev, L. Konicar, and B. U. Cowley, “Evaluation of neurofeedback learning in patients with ADHD: A systematic review,” *Applied Psychophysiology and Biofeedback*, vol. 48, no. 1, pp. 11–25, 2023.
- [6] F. Wang, S. Wu, W. Zhang, Z. Xu, Y. Zhang, C. Wu, and S. Coleman, “Emotion recognition with convolutional neural network and EEG-based EFDMS,” *Neuropsychologia*, vol. 146, 2020.
- [7] X. Wang, Y. Ren, Z. Luo, W. He, J. Hong, and Y. Huang, “Deep learning-based EEG emotion recognition: Current trends and future perspectives,” *Frontiers in Psychology*, vol. 14, 2023.
- [8] A. Elrefaiy, N. Tawfik, N. Zayed, and I. Elhenawy, “EEG emotion recognition framework based on invariant wavelet scattering convolution network,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–19, 2024.
- [9] J. Wang and M. Wang, “Review of the emotional feature extraction and classification using EEG signals,” *Cognitive robotics*, vol. 1, pp. 29–40, 2021.
- [10] M. Seeck, L. Koessler, T. Bast, F. Leijten, C. Michel, C. Baumgartner, B. He, and S. Beniczky, “The standardized EEG electrode array of the IFCN,” *Clinical neurophysiology*, vol. 128, no. 10, pp. 2070–2077, 2017.

- [11] P. Nandwani and R. Verma, “A review on sentiment analysis and emotion detection from text,” *Social Network Analysis and Mining*, vol. 11, no. 1, p. 81, 2021.
- [12] D. Ouyang, Y. Yuan, G. Li, and Z. Guo, “The effect of time window length on EEG-based emotion recognition,” *Sensors*, vol. 22, no. 13, p. 4939, 2022.
- [13] R. Alhalaseh and S. Alasasfeh, “Machine-learning-based emotion recognition system using EEG signals,” *Computers*, vol. 9, no. 4, p. 95, 2020.
- [14] Z.-M. Wang, Z.-Y. Chen, and J. Zhang, “EEG emotion recognition based on PLV-rich-club dynamic brain function network,” *Applied Intelligence*, vol. 53, no. 14, pp. 17 327–17 345, 2023.
- [15] S. Sanei and J. A. Chambers, *EEG Signal Processing and Machine Learning*. John Wiley & Sons Ltd, 2021.
- [16] R. Rojas, *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.
- [17] J. Bruna and S. Mallat, “Invariant scattering convolution networks,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1872–1886, 2013.
- [18] “Kymat: Wavelet Scattering and Signal Processing Solutions,” <https://www.kymat.io/>, 2024, accessed: 2024-07-25.
- [19] H. Cui, A. Liu, X. Zhang, X. Chen, K. Wang, and X. Chen, “EEG emotion recognition based on PLV-rich-club dynamic brain function network,” *Knowledge-Based Systems*, vol. 205, 2020.
- [20] Y. Wang, L. Zhang, P. Xia, P. Wang, X. Chen, L. Du, Z. Fang, and M. Du, “EEG-based emotion recognition using a 2D CNN with different kernels,” *Bioengineering*, vol. 9, no. 6, p. 231, 2022.
- [21] S. Bagherzadeh, K. Maghooli, A. Shalbaf, and A. Maghsoudi, “Recognition of emotional states using frequency effective connectivity maps through transfer learning approach from electroencephalogram signals,” *Biomedical Signal Processing and Control*, vol. 75, p. 103544, 2022.
- [22] S. Bagherzadeh, K. Maghooli, M. Shalbaf, Ahmad, and Arash, “A Hybrid EEG-based emotion recognition approach using wavelet convolutional neural networks and support vector machine,” *Basic and Clinical Neuroscience*, vol. 14, no. 1, p. 87, 2023.
- [23] Y. Chen, H. Zhang, J. Long, and Y. Xie, “Temporal shift residual network for EEG-based emotion recognition: a 3D feature image sequence approach,” *Multimedia Tools and Applications*, vol. 83, no. 15, pp. 45 739–45 759, 2024.

Anexo A

Número canal	Canal EEG	Número canal	Canal EEG
1	Fp1	17	Fp2
2	AF3	18	AF4
3	F3	19	Fz
4	F7	20	F4
5	FC5	21	F8
6	FC1	22	FC6
7	C3	23	FC2
8	T7	24	Cz
9	CP5	25	C4
10	CP1	26	T8
11	P3	27	CP6
12	P7	28	CP2
13	PO3	29	P4
14	O1	30	P8
15	Oz	31	PO4
16	Pz	32	O2

Tabla 6.1: Orden de Geneva para canales de EEG [1].

Resumen FI

UNIVERSIDAD DE CONCEPCIÓN – FACULTAD DE INGENIERÍA RESUMEN DE MEMORIA DE TÍTULO

Departamento : Departamento de Ingeniería
Carrera : Ingeniería Civil Biomédica
Nombre del memorista : Alonso Valderrama Hickmann
Título de la memoria : Clasificación de emociones en Electroencefalograma con Wavelet Scattering y algoritmos de Machine Learning.
Fecha de la presentación oral :

Profesor(es) Guía : Esteban Pino y Mabel Urrutia
Profesor(es) Revisor(es) : Pamela Guevara
Concepto :
Calificación :

Resumen (máximo 200 palabras)

La clasificación de emociones a partir de señales EEG es crucial para avances en neurociencia afectiva e interfaces cerebro-computadora. Esta memoria, vinculada al proyecto FONDECYT 13220040, evalúa la Transformada Wavelet Scattering (WS) como técnica para la extracción de características EEG, buscando mejorar el rendimiento de los clasificadores de emociones. Utilizando la base de datos DEAP como método de validación, se combinó WS con un Perceptrón Multicapa (MLP) y Principal component analysis (PCA) para la reducción de dimensionalidad. Los resultados muestran precisiones superiores al 95 % en la clasificación de Valence, Arousal, Dominance y Liking con validación cruzada de 5 k-folds. Sin embargo, al aplicar el método Leave-One-Subject-Out (LOSO), la precisión disminuyó, lo que indica la necesidad de ajustar la metodología para abordar la correlación que existe entre canales de EEG y la alta variabilidad entre sujetos. En conclusión, aunque WS muestra un potencial significativo para la clasificación de emociones, es esencial mejorar las estrategias de evaluación de clasificadores para asegurar la robustez y generalización del modelo en aplicaciones reales.