



UNIVERSIDAD DE VALLADOLID

FACULTAD DE MEDICINA

ESCUELA DE INGENIERÍAS INDUSTRIALES

TRABAJO DE FIN DE GRADO

GRADO DE INGENIERÍA BIOMÉDICA

Reconocimiento de Emociones a partir de la Actividad Eléctrica Cerebral con Técnicas de Deep Learning

Autor/a:

D.^a Ruth Zamora García

Tutor/a:

D. Sergio Pérez Velasco

D. Roberto Hornero Sánchez

TÍTULO: **Reconocimiento de Emociones a partir de la Actividad Eléctrica Cerebral con Técnicas de Deep Learning**

AUTOR: **D^a. Ruth Zamora García**

TUTOR/ES: **D. Sergio Pérez Velasco**
Dr. D. Roberto Hornero

DEPARTAMENTO: **Teoría de la Señal y las Comunicaciones e Ingeniería Telemática**

Tribunal

PRESIDENTE: **Dr. D. Roberto Hornero Sánchez**

SECRETARIA: **Dr. D. Carlos Gómez Peña**

VOCAL: **Dr. D. Jesús Poza Crespo**

P. SUPLENTE: **Dra. D^a. María García Gardañón**

S. SUPLENTE: **Dr. D. Daniel Álvarez González**

FECHA:

CALIFICACIÓN:

Resumen

Este Trabajo de Fin de Grado (TFG) se centra en el desarrollo y aplicación de técnicas de *deep learning* para el reconocimiento de emociones a partir de la actividad eléctrica cerebral (EEG). Se parte de la premisa de que el EEG es una valiosa fuente de información capaz de revelar los estados emocionales de un individuo. Esta investigación se enfoca en abordar el problema de detectar emociones a partir de datos EEG, lo cual tiene una relevancia significativa en áreas como *Brain Computer Interface* (BCI), la salud mental y la neurociencia afectiva.

A lo largo de este estudio, se ha utilizado la red neuronal profunda EEG-Inception, para analizar y clasificar patrones específicos en las señales EEG asociados con diferentes estados emocionales. Estas señales EEG fueron extraídas de la base de datos SEED, y su contenido permitió llevar a cabo una clasificación multiclase que incluye la identificación de las siguientes emociones triste, neutra y feliz, así como una clasificación binaria destinada a distinguir entre los estados emocionales de felicidad y tristeza.

Los resultados obtenidos han sido notables. En la clasificación multiclase se ha alcanzado una precisión del 51,58%, mientras que, en la clasificación binaria se ha obtenido una mejora significativa alcanzando un 77,49% de precisión. Este tipo de investigaciones se enfrenta a desafíos debido a las diferencias entre sujetos y a la dificultad del modelo para generalizar patrones entre ellos. Sin embargo, los resultados obtenidos indican un enfoque prometedor y sugieren oportunidades de mejora en este campo de estudio.

En conclusión, este TFG representa un avance significativo en el campo del reconocimiento de emociones a partir de señales EEG utilizando técnicas de DL. Apesar de los desafíos presentes, los resultados obtenidos proporcionan una base sólida para futuras investigaciones y mejoras en la detección de emociones a partir de señales biomédicas. Este trabajo no solo amplía nuestro conocimiento sobre la relación entre la actividad cerebral y las emociones, sino que también abre la puerta a una serie de aplicaciones prácticas que pueden tener un impacto positivo en la sociedad en general.

Palabras clave

Electroencefalograma, Deep Learning, Emociones, EEG-Inception.

Abstract

This TFG focuses on the development and application of deep learning techniques for the recognition of emotions based on electroencephalogram (EEG) data. It starts from the premise that EEG is a valuable source of information capable of revealing an individual's emotional states. This research aims to address the challenge of detecting emotions from EEG data, which holds significant relevance in areas such as Brain Computer Interfaces (BCI), mental health, and affective neuroscience.

Throughout this study, the deep neural network EEG-Inception has been employed to analyze and classify specific patterns in EEG signals associated with various emotional states. These EEG signals were extracted from the SEED database, enabling both multiclass classification, encompassing emotions such as sadness, neutrality, and happiness, and binary classification, aimed at distinguishing between the emotional states of happiness and sadness.

The results obtained have been noteworthy. In the multiclass classification, an accuracy of 51.58% was achieved, while in the binary classification, there was a substantial improvement, reaching an accuracy of 77.49%. Such research encounters challenges due to inter-subject differences and the model's difficulty in generalizing patterns among them. Nevertheless, the results obtained indicate a promising approach and suggest opportunities for enhancement in this field of study.

In conclusion, this BT represents a significant advancement in the field of emotion recognition from EEG signals using deep learning techniques. Despite the inherent challenges, the results obtained provide a solid foundation for future research and improvements in the detection of emotions from biomedical signals. This work not only expands our understanding of the relationship between brain activity and emotions but also opens the door to various practical applications that can have a positive impact on society.

Keywords

Electroencephalogram, Deep Learning, Emotions, EEG-Inception.

Glosario de siglas y acrónimos

| | |
|-------------|---------------------------------------|
| TFG | Trabajo Fin de Grado |
| BCI | Brain Computer Interface |
| ML | Machine Learning |
| DL | Deep Learning |
| EEG | Electroencefalograma |
| SVM | Support Vector Machines |
| CNN | Redes Neuronales Convolucionales |
| RNN | Redes Neuronales Recurrentes |
| LOSO | Leave-One-Subject-Out |
| VP | Verdaderos positivos |
| FP | Falsos positivos |
| VN | Verdaderos negativos |
| FN | Falsos negativos |
| EO | Exactitud Observada |
| EE | Exactitud Esperada |
| ERPs | Potenciales relacionados con eventos |
| CV | Computer visión |
| NLP | Natural language preprocessing |
| ICA | Análisis de Componente Independientes |

Índice general

| | |
|--|-----------|
| Resumen..... | 4 |
| Palabras clave..... | 4 |
| Abstract | 5 |
| Keywords | 5 |
| Glosario de siglas y acrónimos..... | 6 |
| Índice de figuras | 9 |
| Índice de tablas..... | 10 |
| 1. Capítulo I..... | 11 |
| 1.1. Ingeniería Biomédica y procesado de señales neuronales..... | 12 |
| 1.2. Introducción a la clasificación de emociones a través de EEG..... | 12 |
| 1.3. Inteligencia Artificial..... | 13 |
| 1.4. Motivación..... | 15 |
| 1.5. Hipótesis | 15 |
| 1.6. Objetivos..... | 16 |
| 1.7. Estructura del documento..... | 16 |
| 2. Capítulo II..... | 18 |
| 2.1. Historia del EEG..... | 19 |
| 2.2. Introducción al EEG..... | 20 |
| 2.3. Posicionamiento de los electrodos..... | 21 |
| 2.4. Adquisición y procesamiento de señales de EEG..... | 21 |
| 3. Capítulo III | 24 |
| 3.1. Introducción..... | 25 |
| 3.2. Clasificación de las emociones..... | 25 |
| 3.3. Emociones vs sentimientos..... | 28 |
| 3.4. ¿Cómo se generan las emociones? | 28 |
| 3.5. ¿Cómo se detectan las emociones?..... | 29 |
| 4. Capítulo IV..... | 32 |
| 4.1. Introducción..... | 33 |
| 4.2. Base de datos SEED..... | 33 |
| 4.3. Clasificador de emociones..... | 34 |
| 4.3.1. Extracción de datos..... | 34 |
| 4.3.2. EEG-Inception | 36 |

| | |
|--|-----------|
| 4.3.3. Modificaciones de la red..... | 38 |
| 4.4. Tipos de clasificación | 38 |
| 4.4.1. Clasificación multiclase..... | 38 |
| 4.4.2. Clasificación binaria..... | 40 |
| 5. Capítulo V..... | 43 |
| 5.1. Introducción..... | 44 |
| 5.2. Resultados de clasificación multiclase..... | 44 |
| 5.3. Resultados de clasificación binaria..... | 51 |
| 6. Capítulo VI..... | 59 |
| 6.1. Introducción..... | 60 |
| 6.2. Discusión..... | 60 |
| 6.3. Comparación con otros estudios similares..... | 61 |
| 6.4. Limitaciones | 63 |
| 7. Capítulo VII | 64 |
| 7.1. Introducción..... | 65 |
| 7.2. Objetivos propuestos | 65 |
| 7.3. Aportaciones..... | 66 |
| 7.4. Conclusiones..... | 66 |
| 7.5. Líneas futuras..... | 67 |
| Bibliografía..... | 68 |

Índice de figuras

| | |
|--|----|
| FIGURA 2.1: DIFERENTES RITMOS ELECTROENCEFALOGRÁFICOS SEGÚN EL ESTADO DEL SUJETO [19]..... | 20 |
| FIGURA 2.2: LOCALIZACIÓN DE LOS ELECTRODOS EN EL SISTEMA INTERNACIONAL 10-20 [22]. | 21 |
| FIGURA 3.1: EJEMPLO DEL CUBO DE LAS EMOCIONES LÖVHEIM [26]. | 26 |
| FIGURA 3.2: RUEDA DE LAS EMOCIONES DE PLUTCHIK [27]. | 26 |
| FIGURA 3.3: DIAGRAMA DEL MODELO CIRCUMPLEJO DE RUSSEL [28]. | 27 |
| FIGURA 4.1: LOCALIZACIÓN DE LOS ELECTRODOS EN NUESTRO ESTUDIO [23]. | 34 |
| FIGURA 4.2: RESUMEN DE LA ARQUITECTURA DE EEG-INCEPTION. LOS BLOQUES DE CONVOLUCIÓN 2D Y LOS BLOQUES DE CONVOLUCIÓN 2D DE PROFUNDIDAD INCLUYEN NORMALIZACIÓN POR LOTES, ACTIVACIÓN Y REGULARIZACIÓN MEDIANTE ELIMINACIÓN. EL TAMAÑO DEL NÚCLEO SE MUESTRA PARA LAS CAPAS DE CONVOLUCIÓN Y DE PROMEDIADO. | 36 |
| FIGURA 4.3: ESQUEMA DE LA CONFIGURACIÓN DE LA RED. | 36 |
| FIGURA 4.4: ESQUEMA DE MATRIZ DE CONFUSIÓN PARA NUESTRA CLASIFICACIÓN MULTICLASE. | 40 |
| FIGURA 4.5: ESQUEMA MATRIZ DE CONFUSIÓN PARA CLASIFICACIÓN BINARIA. | 42 |
| FIGURA 5.1: MATRIZ DE CONFUSIÓN CON LA SUMA DE PREDICCIONES DEL MODELO 1. | 46 |
| FIGURA 5.2: MATRIZ DE CONFUSIÓN CON LA SUMA DE PREDICCIONES DEL MODELO 2. | 46 |
| FIGURA 5.3: MATRIZ DE CONFUSIÓN CON LA SUMA DE PREDICCIONES DEL MODELO 3. | 47 |
| FIGURA 5.4: MATRIZ DE CONFUSIÓN DEL SUJETO 14 EN EL MODELO 2. | 48 |
| FIGURA 5.5: MATRIZ DE CONFUSIÓN DEL SUJETO 7 EN EL MODELO 2. | 48 |
| FIGURA 5.6: MATRIZ DE CONFUSIÓN DEL MODELO DE CLASIFICACIÓN DONDE LA LONGITUD DE LOS SEGMENTOS DE ENTRADA ES DE 3 SEGUNDOS. | 52 |
| FIGURA 5.7: MATRIZ DE CONFUSIÓN DEL MODELO DE CLASIFICACIÓN DONDE LA LONGITUD DE LOS SEGMENTOS DE ENTRADA ES DE 1 SEGUNDO. | 53 |
| FIGURA 5.8: MATRIZ DE CONFUSIÓN DEL MODELO DE CLASIFICACIÓN DONDE LA LONGITUD DE LOS SEGMENTOS DE ENTRADA ES LA MEDIA DE LOS SEGMENTOS DE 1 SEGUNDO EN INTERVALOS DE 10 SEGUNDOS. | 53 |
| FIGURA 5.9: MATRIZ DE CONFUSIÓN DEL SUJETO 12 EN EL MODELO ELEGIDO. SUJETO CON PEOR RENDIMIENTO EN ESTE MODELO. | 54 |
| FIGURA 5.10: MATRIZ DE CONFUSIÓN DEL SUJETO 15 EN EL MODELO ELEGIDO. SUJETO CON MEJOR RENDIMIENTO EN ESTE MODELO. | 55 |
| FIGURA 5.11: GRÁFICO DE ACIERTO DE EMOCIÓN FELIZ EN VIDEO 3 DEL SUJETO 15. | 56 |
| FIGURA 5.12: GRÁFICO DE ACIERTO DE LA EMOCIÓN TRISTE EN EL VIDEO 6 DEL SUJETO 15. | 57 |
| FIGURA 5.13: GRÁFICO DE ACIERTO DE LA EMOCIÓN FELIZ EN EL VIDEO 3 DEL SUJETO 12. | 57 |
| FIGURA 5.14: GRÁFICO DE ACIERTO DE LA EMOCIÓN TRISTE EN EL VIDEO 6 DEL SUJETO 12. | 58 |

Índice de tablas

| | |
|---|----|
| TABLA 4.1: HIPERPARÁMETROS DE LA RED | 38 |
| TABLA 5.1: VALORES OBTENIDOS DE ACCURACY EN LOS 3 MODELOS DE CLASIFICACIÓN MULTICLASE. ACC 1 (MODELO 1), ACC 2 (MODELO 2) Y ACC 3 (MODELO 3)..... | 44 |
| TABLA 5.2: VALORES OBTENIDOS DEL COEFICIENTE KAPPA EN LOS 3 MODELOS DE CLASIFICACIÓN MULTICLASE. KAPPA 1 (MODELO 1), KAPPA 2 (MODELO 2) Y KAPPA 3 (MODELO 3)..... | 45 |
| TABLA 5.3: TABLA CON LOS VALORES DE ACCURACY Y COEFICIENTE KAPPA DE LOS SUJETOS 7 Y 14, SIENDO ESTOS EL SUJETO CON PEOR Y MEJOR RENDIMIENTO EN NUESTRO MODELO RESPECTIVAMENTE..... | 47 |
| TABLA 5.4: NÚMERO DE VIDEOS ACERTADOS POR CADA SUJETO EN EL MODELO 2. SE MUESTRA CON UNA X LOS VIDEOS QUE CADA SUJETO ACIERTA. LOS VIDEOS EN VERDE CORRESPONDEN A FELIZ, LOS BLANCOS A NEUTRO Y LOS ROJOS A TRISTE. | 49 |
| TABLA 5.5 PORCENTAJE DE ACIERTOS DE VIDEOS TOTALES. | 49 |
| TABLA 5.6: TABLA DE LOS PORCENTAJES DE ACIERTOS QUE CADA SUJETO TUVO EN CADA UNA DE LAS TRES SESIONES | 50 |
| TABLA 5.7: VALORES OBTENIDOS DE ACCURACY EN LAS DISTINTAS CONFIGURACIONES DE LONGITUD DE MUESTRAS DE ENTRADA. 3 SEGUNDOS (3s), 1 SEGUNDO (1s) Y MEDIA DE 10 SEGUNDOS (10s)..... | 51 |
| TABLA 5.8: VALORES OBTENIDOS DE COEFICIENTE KAPPA EN LAS DISTINTAS CONFIGURACIONES DE LONGITUD DE MUESTRAS DE ENTRADA. 3 SEGUNDOS (3s), 1 SEGUNDO (1s) Y MEDIA DE 10 SEGUNDOS (10s)..... | 52 |
| TABLA 5.9: TABLA DE PORCENTAJE DE ACIERTOS DE VIDEOS EN TOTAL | 55 |
| TABLA 5.10: NÚMERO DE VIDEOS ACERTADOS POR CADA SUJETO EN EL MODELO 2 CON SEGMENTOS DE MEDIA DE 10 SEGUNDOS. SE MUESTRA CON UNA X LOS VIDEOS QUE CADA SUJETO ACIERTA. LOS VIDEOS EN VERDE CORRESPONDEN A FELIZ Y LOS ROJOS A TRISTE..... | 56 |
| TABLA 6.1: TABLA DE COMPARACIÓN F1-SCORE DE ESTUDIO DE REFERENCIA CON MODELO DE CLASIFICACIÓN BINARIA RBM+LSTM Y NUESTRO ESTUDIO CON CONFIGURACIÓN DE LONGITUD DE ENTRADA DE LAS MUESTRAS DE 3 SEGUNDOS (3s)..... | 62 |
| TABLA 6.2: TABLA DE COMPARACIÓN F1-SCORE DE ESTUDIO DE REFERENCIA CON MODELO DE CLASIFICACIÓN BINARIA RBM+LSTM Y NUESTRO ESTUDIO CON CONFIGURACIÓN DE LONGITUD DE ENTRADA DE LAS MUESTRAS DE MEDIA DE SEGMENTOS DE 1 SEGUNDO EN INTERVALOS DE 10 SEGUNDOS | 62 |

1. Capítulo I

Introducción

Índice

| | |
|---|----|
| 1.1. Ingeniería Biomédica y procesamiento de señales neuronales | 12 |
| 1.2. Introducción a la clasificación de emociones a través de EEG | 12 |
| 1.3. Inteligencia Artificial..... | 13 |
| 1.4. Motivación..... | 15 |
| 1.5. Hipótesis..... | 15 |
| 1.6. Objetivos..... | 16 |
| 1.7. Estructura del documento..... | 16 |

1.1. Ingeniería Biomédica y procesamiento de señales neuronales

La Ingeniería Biomédica constituye un campo de conocimiento interdisciplinario que emplea principios matemáticos, físicos e ingenieriles en el ámbito de la medicina. Uno de sus objetivos es diseñar y fabricar dispositivos destinados a interactuar con el cuerpo humano, con el propósito de mejorar la calidad de vida de las personas [1]. Una de las áreas clave de la IB se centra en el procesamiento de señales biomédicas, que abarca la adquisición, análisis y aplicación de datos procedentes del cuerpo humano para comprender mejor su funcionamiento y diagnosticar enfermedades [2]. Siendo las señales biomédicas de origen neural una de las más relevantes, que además van a ser objeto de estudio en el presente TFG.

Durante el análisis de señales neuronales, se utilizan técnicas para examinar señales electromagnéticas y obtener información del sistema biológico en cuestión, en este caso, el cerebro, con fines de diagnóstico y monitorización posterior [2]. La señal registrada tiene su origen en el cerebro, donde diferentes tipos de células y capas en el córtex cerebral interactúan en la generación de señales electromagnéticas con características particulares en términos de su distribución espacial, temporal y frecuencial, las cuales están asociadas a procesos cognitivos específicos [3]. La sinapsis desempeña un papel fundamental en este proceso, ya que es un fenómeno fisiológico que ocurre cuando las células nerviosas se estimulan mediante la liberación de neurotransmisores por parte de una neurona adyacente [4]. Uno de los principales contribuyentes a la generación de señales neuronales es el intercambio electroquímico que ocurre durante la sinapsis [4].

El procesamiento de señales neuronales implica la aplicación de técnicas y algoritmos para analizar y extraer información relevante de estas señales. Esto puede incluir la detección de patrones, la identificación de eventos, la cuantificación de características específicas y la interpretación de la actividad cerebral [3]. Información que utilizaremos para el desarrollo de este TFG. Ya que, se va a centrar en la clasificación de emociones a través de registros de electroencefalogramas (EEG).

1.2. Introducción a la clasificación de emociones a través de EEG

El reconocimiento y la comprensión de las emociones humanas son elementos fundamentales en la interacción social y en diversos campos médicos. En el campo de la psicología, el reconocimiento de emociones es fundamental para la comprensión de los procesos emocionales en los pacientes, lo que puede ayudar a mejorar la eficacia de las terapias psicológicas y la calidad de vida de las personas. Por ejemplo, la detección temprana de la depresión y la ansiedad en pacientes puede ser clave para prevenir problemas más graves y mejorar la efectividad del tratamiento [5]. Por otro lado, en la medicina, el reconocimiento de emociones puede ser útil en la evaluación del dolor y el estrés en pacientes, lo que puede ayudar a mejorar la atención y el tratamiento médico [6]. Por ejemplo, el análisis de la actividad eléctrica cerebral y la detección de patrones relacionados con el dolor pueden ser utilizados para mejorar la precisión de los diagnósticos y reducir la necesidad de pruebas invasivas. En resumen, el reconocimiento de emociones es un campo de estudio de gran importancia en diversos campos de aplicación, ya que permite una mejor comprensión de los procesos emocionales humanos y la creación de sistemas más inteligentes y adaptativos.

Una de las áreas que ha cobrado gran relevancia en la identificación de emociones es el uso de los registros de encefalografía para clasificar las emociones mediante diferentes métodos. Debido a que, el EEG ofrece una ventana única para explorar las relaciones neurales de las emociones, ya que captura la actividad cerebral en tiempo real. Este enfoque innovador se centra en la idea de que diferentes emociones generan patrones específicos de actividad eléctrica en el cerebro, y estos patrones pueden ser detectados y utilizados para clasificar las emociones experimentadas por un individuo.

Las técnicas tradicionales de análisis de EEG para el reconocimiento de emociones se han centrado en el análisis de bandas de frecuencia y el análisis de potenciales evocados.

El análisis de bandas de frecuencia consiste en descomponer la señal de EEG en diferentes bandas de frecuencia, como delta, theta, alfa, beta y gamma [7]. Cada banda de frecuencia se asocia con diferentes estados de la actividad cerebral y puede ser utilizada para inferir el estado emocional del sujeto. Por ejemplo, según diversos estudios la banda de frecuencia alfa se relaciona con estados de relajación y atención sostenida, mientras que la banda de frecuencia beta se relaciona con la actividad mental y la concentración. De esta manera, se puede inferir el estado emocional del sujeto a partir de la actividad eléctrica cerebral en diferentes bandas de frecuencia [7], [8]. Por otra parte, el análisis de potenciales evocados consiste en medir la actividad eléctrica cerebral en respuesta a estímulos específicos, como sonidos, imágenes o palabras. Se registran los potenciales evocados en diferentes puntos del cerebro y se comparan las diferencias en la actividad eléctrica entre diferentes estímulos. De esta manera, se puede inferir la respuesta emocional del sujeto a diferentes estímulos emocionales [9].

Aunque estas técnicas han sido útiles para el reconocimiento de emociones a partir de la actividad eléctrica cerebral, tienen algunas limitaciones. Por ejemplo, el análisis de bandas de frecuencia no tiene en cuenta la interacción entre diferentes bandas de frecuencia, y el análisis de potenciales evocados no tiene en cuenta la actividad espontánea del cerebro y solo se enfoca en la respuesta a estímulos específicos [7], [9].

Para abordar estas limitaciones, en los últimos años se han desarrollado técnicas de análisis de EEG basadas en el aprendizaje automático y el *deep learning* (DL), que permiten una mayor precisión en el reconocimiento de emociones a partir de la actividad eléctrica cerebral [8]. Estas técnicas tienen en cuenta la interacción entre diferentes bandas de frecuencia y la actividad espontánea del cerebro, lo que permite una mejor comprensión de los procesos emocionales en el cerebro humano [8].

1.3. Inteligencia Artificial

La inteligencia artificial es un campo en auge, centrado en el desarrollo de algoritmos y modelos capaces de aprender y realizar tareas sin una programación explícita. Las aplicaciones reales de este campo se centran en la visión por ordenador o *computer vision* (CV) y el procesamiento del lenguaje natural o *natural language preprocessing* (NLP), técnicas que con amplio desarrollo están empezando a ayudar en decisiones médicas, aunque con alguna que otra barrera [10].

El *machine learning* (ML), es un concepto que apareció en 1959 y fue desarrollado por Arthur Samuel programando un ordenador para jugar a las damas [11]. Este era capaz de comparar las mejores opciones de movimiento y recordar las posiciones de juegos previos, lo que permitían a la máquina aprender y mejorar con cada partida. Samuel en su artículo definió el ML como “una tecnología que utiliza técnicas estadísticas y algoritmos computacionales para proporcionar a los

ordenadores la capacidad de aprender, es decir, mejorar sus resultados en una tarea específica tras procesar datos en suficiente cantidad y sin instrucciones explícitas externas proporcionadas por el programador [11]”.

Para entender mejor este concepto utilizamos la siguiente cita de Mitchell *et al* [12]: “Se dice que un programa de computadora aprende de la experiencia E con respecto a alguna clase de tareas T y mide el desempeño P si su desempeño en las tareas en T, medido por P, mejora con la experiencia E”, que podemos traducir en diseñar algoritmos con 3 sencillos pasos [12]:

1. Consumir datos (‘E’)
2. Aplicar técnicas estadísticas para detectar patrones
3. Y de ese modo ‘aprender’ cómo realizar alguna tarea (‘T’)

Este es un proceso iterativo donde el intento de la cada iteración de la tarea particular se mide contra el resultado correcto de esa tarea (‘P’). Si el algoritmo mejora se ajusta más en la dirección de mejora del rendimiento, por el contrario, si empeora el algoritmo tenderá a la acción contraria [12].

Pero en la actualidad existen tres variantes del ML:

- Aprendizaje supervisado: en este caso se entrena a la máquina proporcionando entradas (*inputs*) asociadas a etiquetas (*outputs*) que han sido determinadas por humanos. El objetivo de estos algoritmos es aprender las reglas generales categorizan los inputs con los outputs correctos [13].
- Aprendizaje no supervisado: a diferencia del caso anterior, aquí no se proporciona información previa categorizada por humanos, sino que se introducen grandes cantidades de datos no etiquetados y es el sistema quien se encarga de encontrar tendencias o patrones y separar esta información por grupos de forma automática [14]
- Aprendizaje por refuerzo: en esta última variante se proporcionan tanto datos etiquetados como sin etiquetar, el sistema interactúa con el entorno dinámico y recibe respuestas negativas o positivas dependiendo de sus acciones, lo que le permite desarrollar mejores clasificaciones y caracterizaciones de los datos [13]

Por otro lado, el aprendizaje profundo o DL es un subconjunto del ML que puede procesar una amplia gama de datos, requiere menor preprocesamiento por parte de los humanos y a veces, puede producir resultados más precisos comparándolos los enfoques tradicionales del ML [15]. Sin embargo, el DL también tiene sus desafíos debido a que, es computacionalmente más costoso en términos de tiempo de ejecución, requisitos de hardware y cantidad de datos necesaria [16]

En concreto, el DL utiliza múltiples capas de unidades de procesamiento no lineal, denominadas neuronas, para la extracción y transformación de características [15]. Un conjunto de estas capas interconectadas de neuronas conforma las llamadas redes neuronales artificiales, que procesan los datos aprendiendo características cada vez más complejas en cada capa. La idea es simular una comprensión abstracta de cómo el cerebro humano procesa información y aprende de su entorno y de la información sensorial.

Esta capacidad de aprender representaciones jerárquicas de los datos es especialmente valiosa cuando se trata de la detección de patrones sutiles en las señales cerebrales asociadas a las emociones. Además, las redes neuronales profundas pueden adaptarse y mejorar su rendimiento a medida que se les proporciona más datos, lo que las hace altamente escalables y flexibles en la tarea de clasificación emocional [15].

Construir una red neuronal desde cero puede ser un proceso complejo, pero también es posible utilizar bibliotecas de código preconstruidas que simplifican su implementación. Esto reduce la

barrera de entrada para los investigadores y profesionales que desean utilizar DL en la clasificación de emociones a través de EEG. Aunque los conceptos matemáticos subyacentes pueden requerir tiempo para comprender completamente, las redes neuronales se basan en matrices, álgebra lineal y otras matemáticas avanzadas, lo que proporciona una base sólida para la modelización de las señales neuronales y la interpretación de los resultados. En última instancia, el DL no solo ofrece la capacidad de mejorar la precisión en la clasificación de emociones, sino que también amplía el potencial para descubrir patrones y correlaciones previamente desconocidos en los datos EEG, enriqueciendo así nuestra comprensión de las bases neurales de las emociones humanas [12].

1.4. Motivación

En el campo de la clasificación de emociones se ha desarrollado sobre todo algoritmos con métodos de ML mediante una previa extracción de características de los EEG, que podían estar relacionadas con determinadas emociones. Entre las características utilizadas para clasificar emociones podemos encontrar [17]:

- Características en el dominio del tiempo (potenciales relacionados con eventos),
- Características en el dominio de la frecuencia (densidad espectral de potencia, transformada rápida de Fourier)
- Características tiempo-frecuencia (Transformada de Fourier de tiempo corto, transformada wavelet discreta)
- Características del sistema dinámico no lineal (entropía diferencial)
- Características de asimetría (asimetría de potencia espectral)
- Características de red cerebral (construcción de redes cerebrales mediante estudio de correlación temporal o coherencia espectral).

Pero con estos métodos de ML, en los que las entradas son características extraídas de la señal de EEG dependientes de la interpretación humana, se puede perder mucha información valiosa que está implícita en el EEG sin procesar. Algunas de las primeras aplicaciones de DL en la clasificación de emociones consistían en solo sustituir la fase de clasificación de las características extraídas previamente [17]. Pero en estas aplicaciones estaríamos limitando la capacidad que tienen las redes de DL para extraer sus propias características del EEG sin procesar y tener un mejor desempeño en la clasificación de emociones.

Existen algunos estudios en los que se extraen estas características para ser clasificadas mediante alguna red de DL, lo que puede seguir dando pérdida de cierta información [17]. Es por ello, que vamos a estudiar las distintas formas que tenemos para clasificar emociones estableciendo como entrada de una red de DL un EEG en crudo (“*raw data*”) que solo se le ha practicado un preprocesado simple.

1.5. Hipótesis

Basándonos en la capacidad del DL para procesar directamente datos de EEG sin una extracción previa de características, y en su habilidad para transferir conocimientos entre tareas y entornos de aplicación, nuestra hipótesis principal plantea que el DL puede abordar con éxito los desafíos inherentes a la clasificación de emociones basada en señales EEG. En particular, se espera que el DL ofrezca ventajas significativas sobre las aproximaciones tradicionales basadas

en ML, superando problemas como la variabilidad intersujeto, que ha sido un obstáculo común en los enfoques convencionales de clasificación emocional. Esta variabilidad intersujeto se refiere a las diferencias individuales en las respuestas cerebrales a las mismas emociones, lo que ha requerido enfoques de entrenamiento personalizados en ML.

La hipótesis se fundamenta en la idea de que las redes neuronales profundas, al aprender representaciones jerárquicas y abstractas de los datos, tienen el potencial de capturar patrones complejos en las señales EEG que son difíciles de discernir mediante métodos tradicionales. Además, la capacidad del DL para transferir conocimientos entre sujetos puede permitir un enfoque más generalizable y menos dependiente de datos específicos del individuo. En este contexto, se espera que los resultados de este estudio demuestren que el DL es una herramienta prometedora para la clasificación de emociones a partir de señales EEG y que puede contribuir significativamente a mejorar la precisión en la detección de estados emocionales, abordando así algunas de las limitaciones previamente identificadas en enfoques basados en ML.

1.6. Objetivos

De acuerdo con estas hipótesis, el objetivo principal de este Trabajo Fin de Grado es desarrollar un clasificador de emociones basado en DL a partir de EEG, utilizando bases de datos públicas que permitirán la comparación de la herramienta desarrollada con las soluciones ya existentes. Para lograr este objetivo general, se establecen los siguientes objetivos específicos:

1. Realizar un revisión bibliográfica exhaustiva de los trabajos previos relacionados con el reconocimiento de emociones utilizando la actividad eléctrica cerebral y técnicas de DL.
2. Recopilar las bases de datos públicas y estudio profundizado de las mismas.
3. Procesar los datos para garantizar la calidad de estos. Esto puede implicar limpieza de artefactos, normalización de los datos, división en épocas, etiquetado, etc.
4. Proponer y desarrollar un modelo de DL adecuado para el reconocimiento de emociones a partir de la actividad eléctrica cerebral. Esto supone selección de arquitectura de la red, definición de capas y configuración de otros hiperparámetros.
5. Evaluar el rendimiento del modelo de DL propuesto utilizando métricas apropiadas, como la precisión, la sensibilidad y la especificidad.
6. Comparar los resultados obtenidos con los trabajos previos y discutir las ventajas y limitaciones del enfoque propuesto.

1.7. Estructura del documento

La memoria del presente Trabajo Fin de Grado se ha estructurado siguiendo los pasos habituales establecidos en una investigación que conlleve relación con las señales biomédicas.

En este primer capítulo se presenta el problema bajo estudio, introduciendo la clasificación de emociones a través de EEG y se proporciona una introducción a la inteligencia artificial, explicando brevemente los conceptos de DL y ML. Además, se plantea la hipótesis, objetivos y motivación para el desarrollo de este TFG.

En el segundo capítulo se aborda el EEG como técnica utilizada para el estudio de la actividad eléctrica cerebral. Se describe qué es, cómo se obtiene y se realiza una breve revisión histórica y evolución de la técnica. Se explica el posicionamiento de los electrodos, la adquisición y procesamiento de las señales EEG.

En el tercer capítulo se estudia la clasificación de las emociones. Se presenta la diferencia entre emociones y sentimientos y los diferentes modelos de clasificación posibles. Se explica cómo se generan las emociones desde una perspectiva fisiológica y se discuten las distintas formas de detectar emociones.

En el cuarto capítulo se describe la metodología utilizada en el estudio. Se detalla el clasificador de emociones utilizado, que se basa en la red neuronal EEG-Inception original y modificada en distintas pruebas. Se explica también la clasificación binaria y multiclase y se presentan las métricas de evaluación empleadas.

En el quinto capítulo se muestran los resultados obtenidos en el estudio. Se presentan los resultados de la clasificación multiclase, incluyendo tablas y gráficos de aciertos en los videos. Se explican las modificaciones realizadas en la red y se presentan los diferentes resultados obtenidos. Además, se exponen los resultados de la clasificación binaria con las mismas métricas que lo hemos hecho en clasificación multiclase.

En el sexto capítulo se lleva a cabo una discusión de los resultados obtenidos. Se ofrecen hipótesis sobre los datos obtenidos y se comparan los resultados con otros estudios similares. Se analizan las limitaciones encontradas durante el desarrollo del TFG, tanto en la base de datos o en el procesamiento de los datos.

En este último capítulo se presentan las conclusiones a las que se ha llegado con este estudio. Se destacan los resultados relevantes y se plantean líneas futuras de investigación derivadas de este estudio.

2. Capítulo II

EEG

Índice

| | |
|--|----|
| 2.1. Historia del EEG..... | 19 |
| 2.2. Introducción al EEG..... | 20 |
| 2.3. Posicionamiento de los electrodos | 21 |
| 2.4. Adquisición y procesamiento de señales de EEG | 21 |

2.1. Historia del EEG

Fue una guerra lo que dio la oportunidad de estudiar el cerebro humano por primera vez. En 1870, los médicos militares del ejército prusiano, Fritsch y Hitzig, descubrieron que, al estimular ciertas áreas del cerebro humano con corriente galvánica, se producían movimientos en el lado opuesto del cuerpo. Esto se confirmó en 1874, cuando el pediatra Richard Caton observó distintas variaciones en las corrientes cerebrales, las cuales se incrementaban durante el sueño y desaparecían después de la muerte. Este pediatra utilizó un galvanómetro reflector con el que consiguió proyectar las ondas cerebrales en las paredes. Publicó su estudio en 1875, sentando con él, las bases para la investigación de las ondas electromagnéticas cerebrales[18].

No fue hasta 1890 que tres científicos, Vasili Yakovlevich Danilevsky, Ernest Fleischl von Marxow y Adolf Beck, comenzaron a investigar las ondas electromagnéticas cerebrales en Cracovia. Beck merece una mención especial, ya que intentó localizar las modalidades sensoriales de la corteza cerebral mediante estimulación eléctrica y sensorial. Encontró respuestas evocadas, fluctuaciones espontáneas de los potenciales cerebrales y observó una disminución en la amplitud de estos potenciales durante la estimulación sensorial [19]. Beck fue el primero en describir la desincronización de las ondas eléctricas después de la estimulación y los potenciales evocados posteriores a la estimulación sensorial [18].

En 1908, Vladimir Pravdich-Neminsky, un joven científico en Kiev, utilizó un galvanómetro de hilo en sus experimentos con animales. Descubrió corrientes de acción en el sistema nervioso central de una rana y fluctuaciones eléctricas espontáneas en el cerebro de un perro mediante medidas eléctricas no invasivas. Sus hallazgos fueron publicados en revistas científicas en 1913 [18], [20].

En 1924, el psiquiatra Hans Berger realizó experimentos en sí mismo, en su hijo Klaus y en otras personas. Registró los primeros EEGs en humanos y descubrió los ritmos alfa. Sin embargo, debido a sus dudas, pospuso la publicación de sus observaciones durante cinco años. Cuando finalmente publicó sus resultados basados en el EEG en 1929, fue ridiculizado y su descubrimiento fue recibido con incredulidad y desprecio. Tomó otros ocho años para que sus hallazgos fueran confirmados por los científicos británicos Edgar Douglas Adrián y Bryan Matthews [20].

Edgar Douglas Adrián y Bryan Matthews no solo confirmaron los hallazgos de Berger, sino que también realizaron otros descubrimientos importantes en la captura de la actividad nerviosa. Adrián ganó el Premio Nobel en 1932 y Matthews fue un ingeniero brillante que inventó el oscilógrafo de Matthews para capturar la actividad nerviosa y el amplificador diferencial, utilizado para el registro de la actividad eléctrica en sistemas biológicos. Mediante el uso de este equipo utilizaron un parpadeo visual en varias frecuencias para inducir actividad alfa en la misma frecuencia. Describieron también la frecuencia que seguía a la respuesta cerebral en el EEG, sentando las bases para muchas aplicaciones de entretenimiento audiovisual [18].

Mientras tanto, en Estados Unidos, la primera publicación sobre el EEG vino de Herbert Jasper y Leonard Carmichael en 1935. Luego, Alfred Loomis, un hombre rico e inteligente que construyó su propio laboratorio en Tuxedo Park, atrajo a grandes científicos como Einstein y Heisenberg. Loomis experimentó con su propio hijo y estaba interesado en el EEG durante el sueño, describiendo por primera vez los usos del sueño. Esto resultaría fundamental para la comprensión de los mecanismos de funcionamiento de la *neurofeedback* basados en ritmos sensoriomotores [18].

La aplicación clínica regular del EEG en hospitales comenzó en 1937. En España, se instalaron equipos de EEG en Barcelona y Madrid entre la década de 1940 y mediados de 1950 y

aún hoy, viven personas, de unos 80 años, cuyos registros EEG en papel se conservan en los Archivos de Historias Clínicas de hospitales de todas partes del mundo [19], [21].

En 1965, se aplicó el algoritmo de Fourier al estudio del EEG, lo que permitió el análisis espectral de las señales cerebrales. Desde 1980, los avances en este campo han ocurrido a una velocidad vertiginosa, siendo uno de los más importantes la conversión digital del EEG. Esta conversión ha permitido la transmisión del EEG por diversos medios sin pérdida alguna de información [21].

Estos avances han revolucionado la forma en que estudiamos y comprendemos el cerebro, abriendo nuevas posibilidades en la investigación y el diagnóstico de trastornos neurológicos [20].

2.2. Introducción al EEG

El EEG es una señal biomédica que registra la actividad eléctrica del cerebro de manera no invasiva, mediante la colocación de electrodos en el cuero cabelludo. Esta actividad es la combinación de las señales eléctricas generadas por millones de neuronas simultáneamente y se presenta de forma oscilatoria y repetitiva, a lo que comúnmente llamamos "ritmo". Los ritmos se clasifican según su rango de frecuencia y su amplitud relativa.

La amplitud del EEG está relacionada con la sincronización de la actividad neuronal. Cuando un grupo de neuronas se excita de forma sincronizada, produce una amplitud alta en la superficie del cráneo debido a la interferencia constructiva generada por esa sincronización. En cambio, si las neuronas se excitan de manera asincrónica, el EEG resultante será irregular y de baja amplitud debido a la interferencia destructiva entre las actividades neuronales individuales. El rango de amplitudes del EEG abarca desde unos pocos microvoltios hasta 200 μV [22].

El rango de frecuencia del EEG depende de la actividad que se esté realizando y está relacionado con su amplitud. En estados de somnolencia, se producen bajas frecuencias y una sincronización entre las neuronas, lo que genera una amplitud elevada. Por el contrario, cuando el cerebro está concentrado en una tarea específica, las frecuencias se desplazan hacia valores más altos y se produce una actividad neuronal asincrónica, lo que resulta en una amplitud más baja. En la Figura 2.1 se pueden observar varios ritmos en función del estado del sujeto [22] [23].

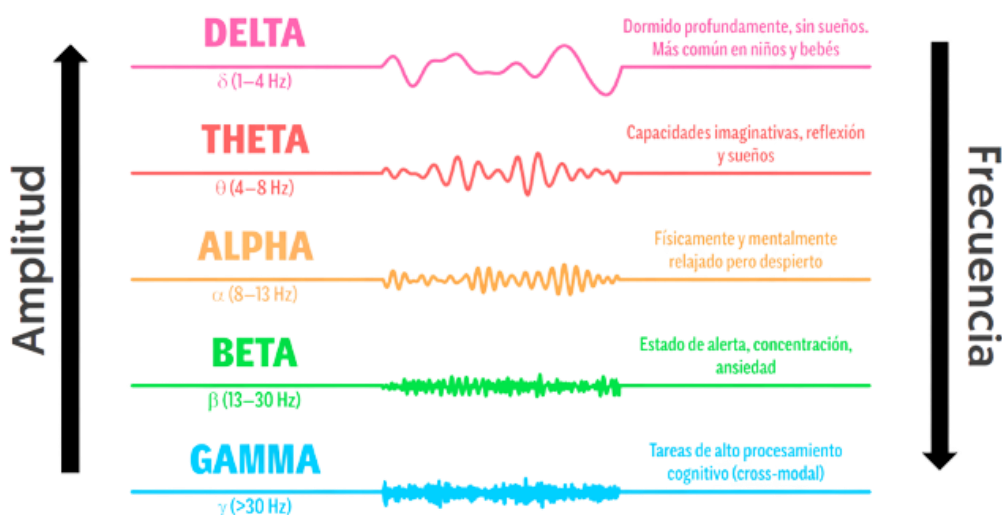


Figura 2.1: Diferentes ritmos electroencefalográficos según el estado del sujeto [19]

2.3. Posicionamiento de los electrodos

La ubicación precisa de los electrodos en el cuero cabelludo es de suma importancia en los registros de EEG, ya sea para garantizar la reproducibilidad de los experimentos o para comparar registros con datos procedentes de distintos individuos [24]. Existen diversos sistemas de posicionamiento de electrodos en EEG, por ejemplo 10-5, 10-10 o 10-20, siendo este último el sistema internacional de colocación de EEG [24].

En el sistema internacional 10-20, mostrado en la *Figura 2.2*, se basa en la colocación de los electrodos separándolos con un margen del 10% o el 20% de la distancia entre el nasión y el inión. Los nombres de los electrodos son letras y números, donde las letras “Fp, F, T, P, O” indican el lóbulo frontopolar, frontal, temporal, parietal y occipital respectivamente, y la letra “C” indica la región central [25]. En cuanto a la ubicación lateralizada, los números pares se refieren a los electrodos colocados en el hemisferio derecho, mientras que los impares se refieren a los del hemisferio izquierdo y los electrodos que se sitúan sobre la línea media están etiquetados con la letra “z” [24].

Los sistemas de posicionamiento 10-5 y 10-10 comparten la misma base que el sistema internacional. La única diferencia entre ellos se encuentra en la separación de los electrodos, que varía de la siguiente manera: en lugar de tener una separación del 10% y 20% de la distancia entre el nasión y el inión, en el sistema 10-5 la separación es del 10% y 5%, mientras que en el sistema 10-10, la separación es del 10% y 10% [24]. Además, la nomenclatura de los electrodos sigue las mismas reglas que el sistema internacional.

2.4. Adquisición y procesamiento de señales de EEG

El EEG es una técnica utilizada para medir señales cerebrales, las cuales tienen una variación eléctrica entre 5 y 200 μV y requieren amplificación posterior para su registro [26]. Estas señales suelen medirse de forma no invasiva situando los electrodos sobre el cuero cabelludo. Los experimentos de EEG pueden categorizarse como bipolares o unipolares. Los registros bipolares muestran fluctuaciones de diferencia de potencial entre dos electrodos corticales, mientras que los registros unipolares muestran la diferencia de potencial entre un electrodo cortical y un electrodo de referencia, situado generalmente en el lóbulo la oreja [25].

La elección de la disposición de los electrodos y su número depende de varios factores, incluida la resolución espacial deseada y el área cerebral de interés. Los sistemas EEG utilizan canales que representan diferentes regiones cerebrales. Se pueden utilizar sistemas con pocos

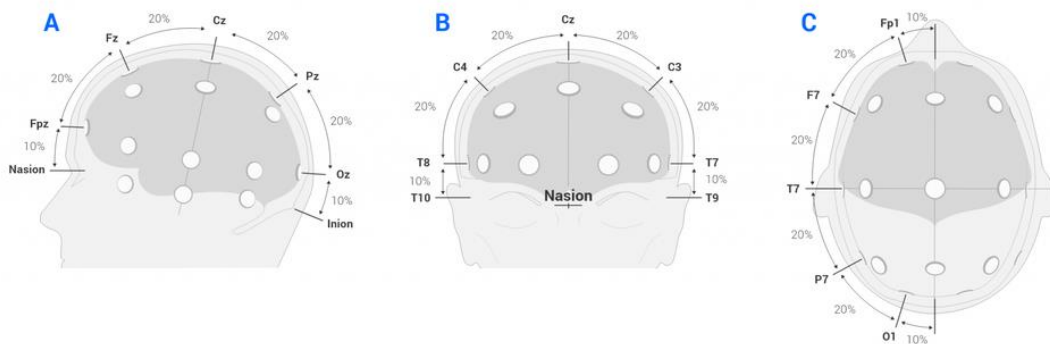


Figura 2.2: Localización de los electrodos en el sistema internacional 10-20 [22].

canales para aplicaciones básicas o sistemas de alta densidad con cientos de electrodos para una mayor precisión espacial.

En términos de tipos de electrodos, existen dos categorías principales: electrodos activos y pasivos [24]:

- **Electrodos Activos:** Estos electrodos requieren un preamplificador cerca del cuero cabelludo para amplificar las señales débiles antes de que sean transmitidas a un amplificador más lejano. Los electrodos activos son beneficiosos porque reducen la cantidad de ruido capturado por el cable largo entre el cuero cabelludo y el amplificador. Sin embargo, su diseño y colocación requieren más cuidado debido a la sensibilidad del preamplificador a las interferencias electromagnéticas.
- **Electrodos Pasivos:** Son electrodos más simples y no requieren un preamplificador cercano al sitio de registro. En cambio, se conectan directamente a un amplificador más alejado. Aunque son más fáciles de usar y colocar, la longitud del cable entre el electrodo y el amplificador puede aumentar la susceptibilidad al ruido eléctrico y de movimiento.

Por tanto, el preprocesamiento de las señales de EEG es un proceso esencial para mejorar la calidad de las señales registradas y prepararlas para su análisis y clasificación. Este proceso implica varias etapas que podemos distinguir según vamos a hacer una clasificación de ML ó DL.

En primer lugar, en el caso de utilizar métodos de ML para la clasificación de emociones deberíamos de seguir los siguientes pasos de procesamiento:

Preprocesamiento: En esta etapa, el objetivo es reducir el ruido y las interferencias en las señales EEG. Una vez adquiridos los registros a estudiar se necesita un procesamiento de estas señales para eliminar los posibles ruidos y artefactos que puedan tener. Podemos distinguir dos tipos de artefactos dependiendo de su origen. Por un lado, existen los artefactos provocados por señales exógenas, es decir, externas al sujeto al que se le está realizando el EEG, como, por ejemplo: interferencia de los electrodos por posibles cambios de impedancia, la componente de frecuencia de una corriente alterna, señales de radio o televisión, entre otros. Por otro lado, también aparecen artefactos endógenos, que son aquellos que tienen su origen en el propio organismo del sujeto, como pueden ser: los provocados por movimientos involuntarios como son la respiración, los latidos cardiacos o el parpadeo de los ojos y también puede haber movimientos musculares por espasmos o tirones. Para detectar y eliminar dichos artefactos se utilizan diferentes técnicas y procedimientos.

- 1) Las principales técnicas de **preprocesamiento** incluyen:
 - Filtrado: Aplicación de filtros para eliminar artefactos y ruido de la señal. Esto puede incluir filtros paso alto para eliminar componentes de baja frecuencia (como ruido de línea) y filtros paso bajo para eliminar componentes de alta frecuencia (como ruido muscular).
 - Rechazo de Artefactos: Identificación y eliminación de artefactos específicos, como el parpadeo de los ojos y los movimientos corporales, utilizando técnicas como el promedio de rechazo o el Análisis de Componentes Independientes (ICA).
- 2) **Extracción de Características:** En los enfoques tradicionales de ML, se seleccionan y extraen características relevantes de la señal EEG preprocesada. Estas características pueden incluir:
 - Frecuencias Dominantes: Identificación de las frecuencias dominantes en diferentes bandas de frecuencia, como delta, theta, alfa, beta y gamma. Ciertas emociones pueden estar asociadas con patrones específicos de actividad en estas bandas.

- Amplitudes y Potencia: Medición de las amplitudes y la potencia de las señales en diferentes regiones cerebrales y bandas de frecuencia. Cambios significativos en estas métricas pueden indicar respuestas emocionales.
 - Asimetría Hemisférica: Exploración de las diferencias en la actividad eléctrica entre los hemisferios izquierdo y derecho del cerebro. Esto puede estar relacionado con la experiencia emocional.
- 3) **Normalización y Preparación de Datos:** Las señales EEG pueden variar en amplitud y rango. La normalización se realiza para hacer que las señales sean comparables y para evitar la dominación de características en el análisis posterior.
 - 4) **Selección de Características:** En algunos casos, se puede aplicar un proceso de selección de características para reducir la dimensionalidad y eliminar características redundantes o irrelevantes. Esto puede mejorar la eficiencia y el rendimiento del modelo de clasificación.
 - 5) **Clasificación:** Finalmente, las características preprocesadas se utilizan para entrenar modelos de clasificación, como *support vector machines* (SVM), *random forest*, o, en el caso de DL, redes neuronales convolucionales (CNN) o recurrentes (RNN). Estos modelos aprenderán a relacionar los patrones en las señales EEG con las emociones correspondientes, lo que permitirá la clasificación precisa de las respuestas emocionales.

Por otro lado, en el caso de nuestro estudio hemos utilizado una clasificación basada en DL y el procesamiento seguido ha sido bastante más simple que en el caso del ML tradicional. Ya que el DL permite clasificar señales de EEG en bruto, sin necesidad de extraer previamente características relevantes ya que la propia red se encarga de extraerlas en capas intermedias. Por tanto, los pasos seguidos fueron:

- 1) Preprocesamiento (datos utilizados ya preprocesados)
- 2) Normalización y preparación de los datos
- 3) Extracción de características y clasificación de las señales de EEG mediante nuestra red de DL

3. Capítulo III

Emociones

Índice

| | |
|---|----|
| 3.1. Introducción | 25 |
| 3.2. Clasificación de las emociones | 25 |
| 3.3. Emociones vs sentimientos | 28 |
| 3.4. ¿Cómo se generan las emociones? | 28 |
| 3.5. ¿Cómo se detectan las emociones? | 29 |

3.1. Introducción

En primer lugar, introducimos el concepto de emoción con una definición que hacen Francisco Mora *et al* [27]: “La emoción es ese motor que todos llevamos dentro. Una energía codificada en ciertos circuitos neuronales localizados en zonas profundas de nuestro cerebro (en el sistema límbico) que nos mueve y nos empuja a ‘vivir’, a querer estar vivos en interacción constante con el mundo y con nosotros mismos. Circuitos que, mientras estamos despiertos, se encuentran siempre activos, en alerta, y nos ayudan a distinguir estímulos importantes para nuestra supervivencia” [27].

Si reflexionamos sobre la definición anterior, podemos entender que las emociones son una fuerza impulsora que nos empuja a vivir. ¿Qué sería de los humanos sin las emociones? Sin embargo, las emociones no son totalmente subjetivas, ya que tienen un fundamento científico que es desarrollado por una parte del cerebro. El problema principal está en nuestra capacidad para ocultar o disfrazar nuestras emociones, lo que se convierte en un problema cuando las emociones están relacionadas con alguna enfermedad. De ahí la importancia de desarrollar métodos y herramientas que puedan identificar de manera objetiva la emoción que experimenta una persona en un momento dado.

3.2. Clasificación de las emociones

El estudio de las emociones es un tema complejo debido a su naturaleza dinámica, ya que experimentan constantes cambios y transformaciones tanto a nivel consciente como inconsciente en los seres humanos. Por ejemplo, la vergüenza puede convertirse en rabia. Pero para facilitar su estudio y clasificación se han propuesto modelos teóricos que actúan como representaciones o abstracciones de la realidad. Estos modelos se basan en los cambios que experimentan las emociones en diversos componentes, como la valencia-activación y los cambios hormonales, lo que permite medirlas y clasificarlas de manera más sencilla [27]. A continuación, se presentan algunos modelos emocionales que resultan relevantes para el reconocimiento de emociones en el campo de la Computación Afectiva [27]:

1) Cubo de las emociones de Lövheim

Lövheim, propuso un modelo en el que se sitúan ocho emociones básicas. Estas emociones se representan en un cubo tridimensional, donde los tres ejes del cubo representan la Serotonina, la Noradrenalina y la Dopamina [28]. En la Figura 3.1, vemos como en los ejes se indican los niveles de mayor o menor cantidad de cada una de estas hormonas.

2) Rueda de emociones de Plutchik

El psicólogo estadounidense Robert Plutchik, desarrolló un sistema de emociones que se asemeja a la forma en que se crean los colores [29]. En este modelo, existen emociones básicas que pueden combinarse para generar emociones más complejas. Además, cada emoción consta de 3 niveles: bajo, medio y alto. Que, por ejemplo, se corresponden con serenidad, alegría y éxtasis respectivamente. Es importante destacar que esta clasificación solo es aplicable a las emociones básicas. En la Figura 3.2 se presenta una representación visual de este concepto

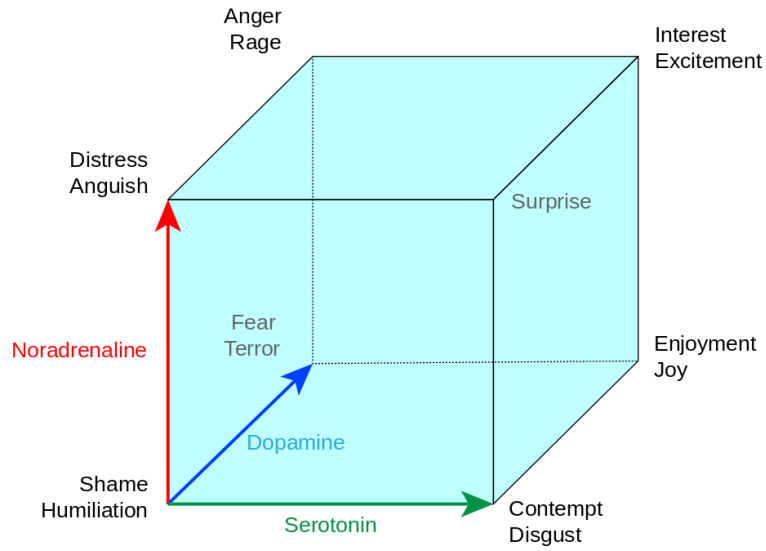


Figura 3.1: Ejemplo del cubo de las emociones Lövheim [26].

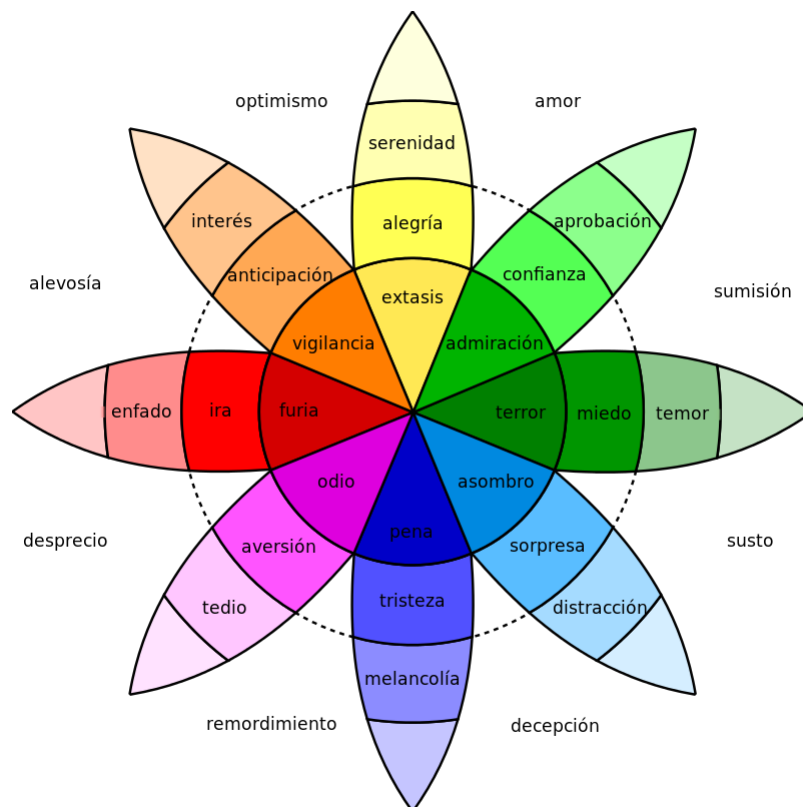


Figura 3.2: Rueda de las emociones de Plutchik [27].

3) Modelo circuplejo de Russel

James A. Russel propuso una representación bidimensional de las emociones basada en la valencia y activación [30].

La valencia se refiere a cómo percibimos la emoción en términos de agrado o desagrado. Por ejemplo, la felicidad se considera una emoción de valencia positiva, mientras que la tristeza se considera una emoción de valencia negativa. Por otro lado, la activación se refiere al nivel de excitación o activación asociado a una emoción. Algunas emociones pueden ser más activadoras y generar una mayor intensidad emocional, como el miedo o la ira, mientras que otras emociones pueden ser más calmadas y tener una menor activación, como la tranquilidad o la relajación [31]

Según el modelo circuplejo, cualquier emoción puede ser ubicada en un espacio bidimensional que combina estas dos dimensiones. Por lo tanto, como vemos en la Figura 3.3, las emociones se pueden representar en un gráfico con un eje horizontal que representa la valencia (positiva a la derecha, negativa a la izquierda) y un eje vertical que representa la activación (alta en la parte superior, baja en la parte inferior).

Este modelo permite visualizar y comparar diferentes emociones en función de su valencia y activación relativa. Además, también permite explorar cómo las emociones se relacionan entre sí y cómo pueden cambiar a lo largo del tiempo o en diferentes situaciones.

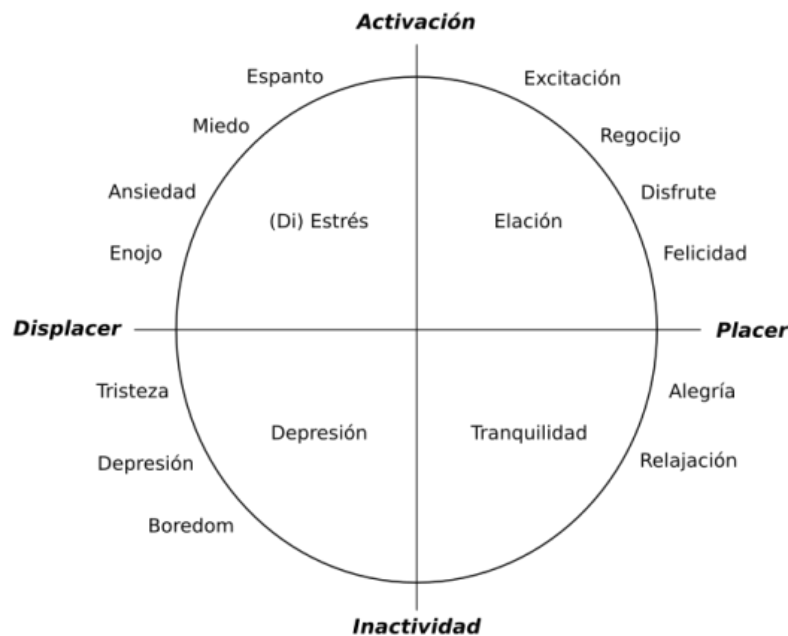


Figura 3.3: Diagrama del modelo circuplejo de Russel [28].

3.3. Emociones vs sentimientos

Aunque inicialmente pueden parecer términos similares, la "emoción" y "sentimiento" tienen significados diferentes.

La emoción, como explicamos anteriormente, es una respuesta del cerebro ante un estímulo externo o interno. Es una experiencia transitoria y no permanente que nos impulsa a la acción [27]. Para comprenderlo mejor, podemos utilizar un ejemplo: Imaginemos que recibes una sorpresa inesperada de un ser querido. En ese momento, puedes experimentar una emoción de alegría y sorpresa. Esta emoción surge como una respuesta automática de tu cuerpo ante la sorpresa y el afecto que recibes.

Pero ¿qué son realmente los sentimientos? Un sentimiento puede describirse como la combinación de una emoción y un pensamiento. Es decir, en los sentimientos intervienen tanto la reacción fisiológica como un componente subjetivo y cognitivo. A diferencia de las emociones, los sentimientos suelen tener una duración más prolongada, ya que persisten mientras los pensamos. No existe un sentimiento sin una emoción previa [27]. Siguiendo el ejemplo anterior de la sorpresa inesperada de un ser querido, a medida que reflexionas sobre el gesto y te sientes agradecido/a, desarrollas un sentimiento de gratitud hacia esa persona. Este sentimiento implica una combinación de la emoción inicial de alegría y sorpresa con pensamientos y reconocimiento de la bondad de la otra persona. A medida que sigues recordando el gesto y valorando la importancia de esa persona en tu vida, el sentimiento de gratitud puede perdurar y fortalecer tu conexión emocional con ella.

En resumen, la emoción inicial de alegría y sorpresa surge automáticamente en respuesta a la sorpresa y el afecto recibido, mientras que el sentimiento posterior de gratitud se desarrolla a través de una combinación de la emoción inicial y pensamientos cognitivos que reconocen y aprecian la bondad de la otra persona.

3.4. ¿Cómo se generan las emociones?

El sistema límbico es una de las redes neuronales más relevantes en el comportamiento humano, al punto que se le denomina frecuentemente como el "cerebro emocional". Este sistema no constituye una región cerebral separada, sino que está compuesto por un conjunto de estructuras ubicadas en diferentes áreas, principalmente a ambos lados del tálamo, es decir, debajo del cerebro. Estas estructuras establecen numerosas conexiones neuronales entre sí e incluyen áreas como la corteza cerebral, los núcleos basales, el tálamo y el hipotálamo [32].

El sistema límbico parece desempeñar un papel fundamental en nuestra vida emocional, ya que es la región del cerebro donde se generan las emociones. Por esta razón, es crucial desarrollar técnicas que faciliten el estudio de las emociones desde un enfoque objetivo. El gesto o expresión facial puede ser fácilmente manipulado, lo que genera confusión y dificulta la investigación, prevención y tratamiento de enfermedades mentales. Es por ello que, la ciencia busca eliminar estas barreras para poder avanzar en la comprensión y abordaje de las enfermedades psíquicas [32].

A continuación, describiremos algunas de las partes más importantes del sistema límbico que están involucradas en la generación de emociones, como el hipotálamo, el hipocampo y la amígdala. También hablaremos sobre otros sistemas menos conocidos pero implicados en la

generación de emociones, como el sistema nervioso parasimpático, el sistema nervioso simpático y el sistema nervioso entérico.

- **Hipotálamo:** Esta pequeña estructura se encuentra debajo del tálamo y desempeña un papel fundamental en la regulación de las emociones y los estados fisiológicos. Actúa como un puente entre el cerebro y el sistema endocrino, coordinando la comunicación entre neuronas y hormonas. El hipotálamo también coordina las respuestas del sistema nervioso autónomo, enviando órdenes a diferentes partes del cuerpo para adaptarse a diversas situaciones. Además, es responsable de mantener la homeostasis en el organismo, asegurando el equilibrio y el funcionamiento adecuado del cuerpo [33].
- **Hipocampo:** El hipocampo es una estructura curvada de la corteza ubicada en el lóbulo temporal. Juega un papel crucial en el aprendizaje y la consolidación de la memoria. Facilita la transformación de la memoria a corto plazo en memoria a largo plazo. Si el hipocampo resulta dañado, puede afectar la capacidad de una persona para reconstruir recuerdos. El hipocampo recibe información sensorial de diferentes vías y se conecta con el hipotálamo a través de una vía llamada fórnix [33].
- **Amígdala:** La amígdala es una estructura en forma de almendra que desempeña un papel central en la generación y procesamiento de las emociones. Es especialmente importante en situaciones de amenaza o miedo. La amígdala recibe información sensorial y desempeña un papel en la evaluación y respuesta emocional [33].

Como se ha explicado previamente, el sistema límbico desempeña un papel fundamental en la generación de emociones. Sin embargo, no es la única parte del sistema nervioso implicada en este proceso. Es importante destacar también la función del sistema nervioso autónomo que consta de dos partes que funcionan en oposición entre sí.

Por un lado, se encuentra el sistema nervioso simpático, que se origina en la médula espinal y se encarga de preparar al cuerpo para situaciones de peligro, como la huida o la violencia. El sistema nervioso simpático tiene diversos efectos, como la dilatación de las pupilas, la estimulación de las glándulas sudoríparas, la dilatación de los vasos sanguíneos en los músculos principales, el aumento de la presión cardíaca, la apertura de los tubos bronquiales en los pulmones y la liberación de epinefrina en la sangre (una hormona que provoca respuestas similares a las del sistema nervioso simpático en varias partes del cuerpo) [32].

Por otro lado, está el sistema nervioso parasimpático, que se origina en el tronco cerebral y la médula espinal de la zona lumbar. Su función principal es restablecer el estado del cuerpo previo a la activación del sistema nervioso simpático. Las respuestas del sistema nervioso parasimpático incluyen la activación de las glándulas salivares, la estimulación de las secreciones estomacales, el aumento de la actividad intestinal, la estimulación de las secreciones pulmonares y la disminución de la frecuencia cardíaca, entre otros [32].

Es importante mencionar que existe otra parte del sistema nervioso autónomo que a menudo se pasa por alto, conocido como el sistema nervioso entérico. Este sistema está estrechamente relacionado con la actividad del estómago y es responsable, por ejemplo, de la sensación de "mariposas en el estómago" que experimentamos cuando estamos nerviosos [32].

3.5. ¿Cómo se detectan las emociones?

Existen diversas técnicas utilizadas para detectar emociones en los seres humanos, siendo aquellas que utilizan biosensores las que suelen obtener mejores resultados. Como mencionamos previamente, las emociones generan cambios fisiológicos significativos, como el aumento de la

respiración o del flujo sanguíneo. Es por esto que, el uso de biosensores presenta ventajas en comparación con la detección de expresiones faciales, voz o posturas corporales. Debido al reducido tamaño de estos sensores, es posible realizar una monitorización constante de datos, lo que permite medir las emociones de forma natural y no inducida por medios externos como videos, imágenes o música. No obstante, el principal desafío radica en que no solo las emociones afectan los marcadores fisiológicos, y no todas las personas responden de la misma manera ante un estímulo similar, lo cual agrega complejidad al análisis. A continuación, se destacan algunos de los bioseñales que se pueden utilizar [34]:

- **Electrocardiograma (ECG):** Esta técnica mide la actividad eléctrica del corazón mediante electrodos colocados en la superficie del cuerpo, especialmente en el pecho. Con el ECG, es posible medir la frecuencia cardíaca, los intervalos entre latidos, la variabilidad de la frecuencia cardíaca y la arritmia respiratoria. Estos parámetros pueden asociarse a estados emocionales como la relajación o el estrés.
- **Volumen de pulso sanguíneo (BPS):** Esta bioseñal registra de forma no invasiva los cambios relativos en el volumen sanguíneo en las venas del dedo índice, lo que permite calcular la vasoconstricción, la dilatación vascular, la frecuencia cardíaca y la hipovolemia. Un alto volumen de pulso sanguíneo o vasoconstricción puede indicar furia o estrés, mientras que niveles reducidos pueden relacionarse con tristeza o relajación.
- **Medida de la respiración (RSP):** Esta bioseñal mide los cambios en la respiración a través de una banda colocada en el pecho. Se pueden registrar parámetros como la frecuencia respiratoria y la profundidad de la respiración, brindando información sobre situaciones de relajación, depresión o estrés en función de la rapidez o lentitud de la respiración.
- **Conductividad de la piel (SC):** También conocido como bioseñales de la actividad eléctrica de la dermis, mide la capacidad de la piel para transmitir corrientes eléctricas, la cual varía en función de la sudoración y los cambios en el organismo. La conductividad de la piel puede ser un indicador de estrés, furia o miedo, aunque puede verse afectada por factores ambientales.
- **Actividad electroencefalográfica (EEG):** Esta técnica utiliza electrodos colocados en el cuero cabelludo para medir la actividad eléctrica en la corteza cerebral. Existe una estrecha relación entre la actividad cerebral y la generación de emociones, ya que el sistema límbico, donde se generan las emociones, se encuentra en la corteza cerebral. Por lo tanto, las señales EEG son fundamentales para analizar las emociones.

En resumen, estas técnicas ofrecen la posibilidad de medir señales fisiológicas relacionadas con las emociones de manera no invasiva. Sin embargo, es importante tener en cuenta que no todos los individuos responden de la misma manera ante los mismos estímulos emocionales, lo que puede añadir complejidad al análisis de estos marcadores fisiológicos. En nuestro caso, hemos elegido el EEG como pieza fundamental en nuestro análisis sobre las emociones debido a su capacidad para proporcionar una medida directa y rápida de la actividad cerebral, considerada la fuente primaria de las emociones humanas.

Nuestra elección se centra en el hecho de que las emociones tienen su origen en el cerebro, lo que permite una detección precisa y temprana de los estados emocionales. A diferencia de otros bioseñales que podrían reflejar respuestas fisiológicas más tardías o indirectas, el EEG nos permite un acceso directo a la actividad eléctrica cerebral, lo que facilita la captura de las emociones en su punto de origen. Además, el EEG ofrece versatilidad en la detección de

emociones, lo que lo convierte también en una herramienta adecuada para explorar y comprender una amplia gama de experiencias emocionales humanas.

4. Capítulo IV

Metodología

Índice

| | |
|---|-----------|
| 4.1. Introducción | 33 |
| 4.2. Base de datos SEED..... | 33 |
| 4.3. Clasificador de emociones | 34 |
| 4.3.1. Extracción de datos..... | 34 |
| 4.3.2. EEG-Inception | 36 |
| 4.3.3. Modificaciones de la red..... | 38 |
| 4.4. Tipos de clasificación | 38 |
| 4.4.1. Clasificación multiclase..... | 38 |
| 4.4.2. Clasificación binaria..... | 40 |

4.1. Introducción

A continuación, se detalla la metodología aplicada en la clasificación de emociones. Desde el tratamiento de los datos, hasta los distintos tipos de clasificación de DL empleados en el presente estudio.

4.2. Base de datos SEED

Para llevar a cabo este estudio hemos solicitado acceso de la base de datos SEED, que consta de registros de EEG obtenidos en experimentos realizados a 15 sujetos. Estos experimentos consistieron en evocar emociones a los participantes mediante la presentación de 15 videos, de 4 minutos de duración. De los cuales, 5 videos estaban asociados a emociones negativas, 5 a positivas o felices y por último otros 5 que pretendían una emoción neutra del sujeto. Después de observar cada video, los participantes debían realizar un cuestionario para informar respecto de sus estados afectivos [35].

Con el fin de realizar estudios sobre las similitudes entre las señales durante las sesiones e individuos, se llevaron a cabo 3 sesiones con cada participante. Estas sesiones se realizaron en diferentes días, con al menos una semana de diferencia entre ellas. Cada sesión consistió en 15 experimentos, que se corresponden con los 15 videos, lo que resulta un total de 45 experimentos por participante.

En cada registro las señales se han obtenido mediante el sistema ‘ESI NeuroScan System’, que es un gorro acorde a la colocación de electrodos del sistema 10-10. Como vemos en la *Figura 4.1*, cuenta con 62 canales que son registrados a una frecuencia de muestreo de 1000 Hz [35]

En nuestro caso hemos utilizado unos datos ya preprocesados, por lo que no hemos tenido que aplicar ningún procesamiento previo al estudio. El procesamiento aplicado a la base de datos elegida es el siguiente: las señales que fueron registradas a 1000 Hz, se les aplicó un filtro paso bajo a 75 Hz, para eliminar las altas frecuencias. Posteriormente se submuestrearon a 200Hz y eliminaron manualmente los artefactos. De forma que, había estudios contaminados por EMG y EOG y tras comprobarlo visualmente, eliminaron manualmente los experimentos completos que resultaban muy ruidosos [36]. Además, en este paso se extrajeron los segmentos de EEG de interés, es decir, los que correspondían a la duración de cada video del experimento. [35]

El conjunto de datos consta de 45 archivos que contienen los registro de las señales de EEG. Estos archivos se nombran según el participante y la fecha del experimento, por ejemplo: “ml_20130709.mat”. Además, hay un archivo llamado “labels.mat” que contiene las etiquetas de las emociones asignadas a cada experimento.

Cada archivo.mat de cada participante consiste en una matriz de dimensiones 15x62x48000, que representa el número de videos (15), el número de canales (62) y los datos de la señal almacenados para cada canal. Este último varía un poco dependiendo de la duración exacta de los videos, ya que no todos los videos tienen una duración de 4 minutos exactos.

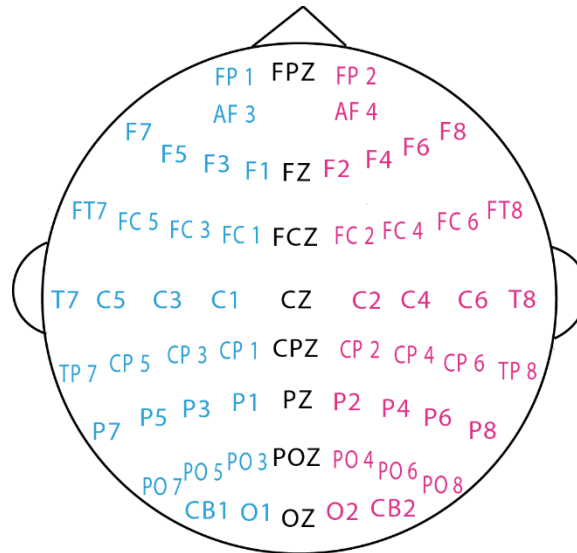


Figura 4.1: Localización de los electrodos en nuestro estudio [23].

4.3. Clasificador de emociones

4.3.1. Extracción de datos

Comenzamos con la base de datos SEED preprocesada, la cual consta de 45 archivos que representan registros de EEG de 15 sujetos. Estos registros fueron adquiridos durante 3 sesiones diferentes. En cada archivo hallamos 16 conjuntos de datos, siendo los primeros 15 cada uno un registro de 4 minutos. Cada registro se corresponde con cada uno de los 15 videos. El último conjunto de datos contiene las etiquetas, tomando la forma de una matriz con 1 fila y 15 columnas. Cada columna etiqueta secuencialmente la emoción de los videos mencionados anteriormente [35].

Los registros a estudiar tienen una duración de unos 4 minutos, aproximadamente la duración de los videos, fueron registrados a una frecuencia de muestreo de 1000 Hz y se submuestrearon a 200 Hz con el objetivo de reducir el ruido a altas frecuencias y disminuir la carga computacional del estudio. Aún así, contamos con conjuntos de datos de gran envergadura que no pueden ser directamente introducidos en nuestra red para su clasificación. Así que, el primer paso que emprendimos fue la segmentación de estos datos en muestras de 1 segundo cada una. Considerando que la frecuencia de muestreo es de 200 Hz, pudimos calcular la longitud de cada muestra de la siguiente manera:

$$\text{Longitud muestra} = \text{frecuencia muestreo} \times \text{duración muestra} = 200 \text{ Hz} \times 1 \text{ s}$$

Así, obtenemos la longitud de cada muestra de 1 segundo.

Los conjuntos de datos están compuestos por tres dimensiones [segmentos, canales, frecuencia]. Seleccionamos los canales que nos interesan, en este caso, todos, y recogemos todos los datos asociados en la segunda dimensión.

Es crucial también calcular la cantidad total de muestras que se generarán, lo cual se realiza de la siguiente manera:

$$\text{Total muestras} = \text{eeg.shape}[1](\text{tamaño del conjunto de datos}) // \text{longitud muestra}$$

Teniendo estos dos valores, estamos preparados para entrar en un bucle `'for'` que itera a través de un rango de valores que representan las diferentes muestras que se extraerán de la señal EEG (`'total_muestras'`). En cada iteración del bucle, se calcula el índice de inicio de la muestra actual multiplicando `'mu'` (el número de muestra actual) por la `'longitud_muestra'`. Esto determina el punto de inicio en la señal EEG desde donde se extraerá la muestra actual. A continuación, se calcula el índice de finalización de la muestra actual sumando la longitud de la muestra al índice de inicio. Esto define el punto de finalización de la muestra en la señal EEG. Y finalmente, se extrae la muestra actual de la señal EEG. Siendo `'eeg'` una matriz que almacena la señal EEG completa, utilizamos la notación `'eeg[:, inicio:fin]'` para seleccionar una porción de la matriz que abarca desde el índice de inicio hasta el índice de finalización. Esto crea una submatriz que representa la muestra EEG actual. Entonces, en cada iteración del bucle, se obtiene una nueva muestra de la señal EEG y se almacena en la variable `'muestra_eeg'`.

Luego, fuera de este ciclo, se utilizó el comando `'features.append(muestra_eeg)'` para acumularlas en una lista.

No únicamente requerimos guardar las muestras, sino también es esencial conocer a qué emoción, sujeto, video y sesión corresponde cada una. Por lo tanto, mediante un extenso ciclo que recorre la totalidad de la base de datos, generamos listas destinadas a conservar esta información deseada en conjunto con la muestra.

Por otro lado, la normalización de los datos es esencial para mejorar la estabilidad del entrenamiento, facilitar un aprendizaje más rápido y eficiente, promover la generalización y reducir el impacto de diferencias en la escala de las características [12]. Así que, es otro de los pasos que llevamos a cabo, normalizar todos los datos almacenados en la variable `features`.

Una vez que hemos dividido nuestra base de datos en segmentos de 1 segundo y les hemos asignado etiquetas correspondientes de las emociones a las que pertenecen, podemos comenzar a considerar cuál sería la manera más adecuada de distribuir los conjuntos de entrenamiento, validación y test.

La división destinada al conjunto de test estaba claramente establecida, llevando a cabo el método *Leave-One-Subject-Out* (LOSO) [12]. Desde un principio, reservamos a un sujeto completo fuera del conjunto de entrenamiento, con el propósito de llevar a cabo una clasificación intersujeto posteriormente. Sin embargo, al abordar la distribución del conjunto de validación, surgieron algunas dudas. Por lo tanto, pusimos a prueba tres métodos diferentes:

1. Utilizando la función `validation_split` de la biblioteca Keras, que consiste en dividir automáticamente el conjunto de datos de entrenamiento en conjuntos de entrenamiento y validación al inicio del proceso de entrenamiento del modelo. Esta función permite especificar la proporción de datos que se destinarán a la validación, lo que facilita la evaluación continua del rendimiento del modelo sin la necesidad de una división manual.
2. Asignamos una sesión completa de todos los sujetos al conjunto de validación. Es decir, todos los registros correspondientes a una sesión específica, como la sesión 3, se destinaron al conjunto de validación, mientras que el resto de los datos quedaron en el conjunto de entrenamiento.
3. Realizamos otra división manual en la cual seleccionamos esta vez 3 videos, cada uno ligado a una emoción específica (felicidad, neutralidad, tristeza), y los extraemos de los registros de cada sujeto para emplearlos como conjunto de validación. Los videos escogidos, son los mismos en todos los sujetos y en las tres sesiones, por lo que tenemos

135 (3 videos x 15 sujetos x 3 sesiones) registros como conjunto de validación. Mientras que, conservamos en el conjunto de entrenamiento los datos restantes.

4.3.2. EEG-Inception

EEG-Inception es una CNN innovadora inspirada en el trabajo de Eduardo *et al.*[37] para la clasificación de imágenes, adaptando sus conceptos para una arquitectura mejorada en el procesamiento de EEG y detección de potenciales relacionados con eventos (ERPs). Incorpora conceptos de clasificación de imágenes, como módulos Inception para capturar relaciones entre características a distintas escalas y convoluciones *depthwise*. La **¡Error! No se encuentra el origen de la referencia.** brinda una visión arquitectónica, mientras que la *Figura 4.3* detalla configuraciones. Cada bloque convolucional (2D y *depthwise* 2D) incluye normalización por lotes, activación y regularización para evitar sobreajuste [38].

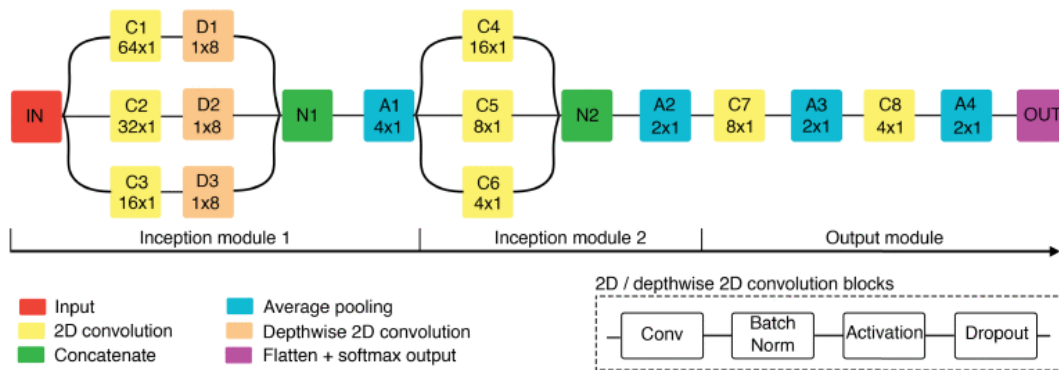


Figura 4.2: Resumen de la arquitectura de EEG-Inception. Los bloques de convolución 2D y los bloques de convolución 2D de profundidad incluyen normalización por lotes, activación y regularización mediante eliminación. El tamaño del núcleo se muestra para las capas de convolución y de promediado.

| Block | Type | Filters | Depth | Kernel | Padding | Output shape | Connected to | Role |
|-------|------------------|---------|-------|--------|---------|--------------|--------------|---------------------|
| IN | Input | - | - | - | - | 128 × 8 × 1 | C1, C2, C3 | Input |
| C1 | Conv2D | 8 | - | 64 × 1 | Same | 128 × 8 × 8 | D1 | Temporal analysis |
| D1 | DepthwiseConv2D | - | 2 | 1 × 8 | Valid | 128 × 1 × 16 | N1 | Spatial analysis |
| C2 | Conv2D | 8 | - | 32 × 1 | Same | 128 × 8 × 8 | D2 | Temporal analysis |
| D2 | DepthwiseConv2D | - | 2 | 1 × 8 | Valid | 128 × 1 × 16 | N1 | Spatial analysis |
| C3 | Conv2D | 8 | - | 16 × 1 | Same | 128 × 8 × 8 | D3 | Temporal analysis |
| D3 | DepthwiseConv2D | - | 2 | 1 × 8 | Valid | 128 × 1 × 16 | N1 | Spatial analysis |
| N1 | Concatenate | - | - | - | - | 128 × 1 × 48 | A1 | Concatenation |
| A1 | AveragePooling2D | - | - | 4 × 1 | - | 32 × 1 × 48 | C4, C5, C6 | Concatenation |
| C4 | Conv2D | 8 | - | 16 × 1 | Same | 32 × 1 × 8 | N2 | Temporal analysis |
| C5 | Conv2D | 8 | - | 8 × 1 | Same | 32 × 1 × 8 | N2 | Temporal analysis |
| C6 | Conv2D | 8 | - | 4 × 1 | Same | 32 × 1 × 8 | N2 | Temporal analysis |
| N2 | Concatenate | - | - | - | - | 32 × 1 × 24 | A2 | Concatenation |
| A2 | AveragePooling2D | - | - | 2 × 1 | - | 16 × 1 × 24 | C7 | Dimension reduction |
| C7 | Conv2D | 12 | - | 8 × 1 | Same | 16 × 1 × 12 | A3 | Temporal analysis |
| A3 | AveragePooling2D | - | - | 2 × 1 | - | 8 × 1 × 12 | C8 | Dimension reduction |
| C8 | Conv2D | 6 | - | 4 × 1 | Same | 8 × 1 × 6 | A4 | Temporal analysis |
| A4 | AveragePooling2D | - | - | 2 × 1 | - | 4 × 1 × 6 | C7 | Dimension reduction |
| OUT | Dense | - | - | - | - | 2 | - | Softmax output |

Column "Type" describes the class used to implement each block in Keras framework. It should be taken into account that this implementation may vary across different frameworks. All convolutional blocks (i.e., Conv2D and DethwiseConv2D) include batch normalization, activation and dropout regularization. The model has 15154 parameters, of which 14926 are fitted during training.

Figura 4.3: Esquema de la configuración de la red

La estructura original de EEG-Inception se divide en tres secciones principales:

- Módulo Inception 1:

En este módulo, la señal se procesa en tres escalas temporales distintas para cada canal EEG, según los tamaños de kernel de los bloques convolucionales C1, C2 y C3, que son 64×1 , 32×1 y 16×1 , respectivamente. Dado que la frecuencia de muestreo es 128 Hz, estos tamaños representan ventanas temporales de 500 ms, 250 ms y 125 ms. Luego, las capas D1, D2 y D3 realizan un procesamiento en el dominio espacial utilizando convoluciones *depthwise*. Estas convoluciones, inicialmente aplicadas en clasificación de imágenes, factorizan un kernel convolucional en kernels más pequeños al operar en cada canal de entrada por separado, reduciendo la cantidad total de parámetros. Cuando se emplean en el procesamiento de EEG, permiten aprender filtros espaciales óptimos (pesos de canales) para cada patrón temporal extraído por la capa anterior. Luego, la capa de concatenación N1 fusiona las características resultantes de D1, D2 y D3. Finalmente, se aplica una operación de promediado (average pooling) para reducir la dimensionalidad [38].

- Módulo Inception 2:

Este módulo sigue una estructura similar al anterior, compuesto por tres ramas que procesan la señal EEG en tres escalas. Es importante mencionar que, después de la capa de promediado en la primera etapa, estas escalas corresponden a tamaños de kernel de 16×1 , 8×1 y 4×1 . Este módulo extrae características adicionales en un nivel de abstracción superior, considerando todos los canales EEG. Al igual que antes, las salidas de los bloques convolucionales C4, C5 y C6 se concatenan. Luego, también se aplica un promediado para la reducción de dimensionalidad [38].

- Módulo de Salida:

Las dos últimas capas convolucionales están diseñadas para extraer los patrones más relevantes para la clasificación final, comprimiendo la información en unas pocas características. Importante señalar que el número de filtros disminuye progresivamente, lo que, junto con las capas de promediado, reduce la dimensionalidad para evitar el sobreajuste. En realidad, la información del EEG se concentra en 24 características antes de la capa de clasificación final. Finalmente, la salida *softmax* estima la probabilidad para cada clase (objetivo y no objetivo) [38].

El modelo en el estudio de Santamaria-Vázquez et al [38] fue entrenado con la siguiente configuración: optimizador Adam con hiperparámetros predeterminados $\beta_1=0.9$ y $\beta_2=0.999$; función de pérdida de entropía cruzada categórica; tamaño de *mini-batch* de 1024; y 500 épocas. Para acelerar el entrenamiento y evitar el sobreajuste, aplicaron '*early stopping*' cuando la pérdida del conjunto de validación no mejoraba durante 10 épocas consecutivas, restaurando los pesos que minimizaban esta métrica.

La elección de hiperparámetros en modelos de DL es crucial para obtener resultados adecuados. En particular, el *learning rate*, la función de activación y la tasa de *dropout* tienen un gran impacto. Con el fin de alcanzar una solución óptima, se optimizaron automáticamente en el conjunto de validación utilizando búsqueda en cuadrícula. El resto de los hiperparámetros (es decir, número de capas, número de ramas en los módulos Inception, número de filtros, tamaños de kernel y tamaños de *pooling*) se eligieron de manera heurística.

4.3.3. Modificaciones de la red

Para aplicar la red EEG-Inception a nuestro estudio tuvimos que llevar a cabo ciertas modificaciones en la red y en la elección de hiperparámetros, para optimizar al máximo la clasificación.

Apenas realizamos modificaciones en la arquitectura de la red. Dado que se trataba de una red ya publicada con resultados sólidos en la clasificación de EEG, optamos por utilizarla sin alteraciones significativas. No obstante, introdujimos un pequeño cambio al reemplazar la capa "flatten" por una capa de "globalmaxpooling2D", con el propósito de brindar mayor regularización a la red y explorar posibles mejoras en los resultados. Además, llevamos a cabo experimentos utilizando ambas configuraciones.

Cabe mencionar que también realizamos otro ajuste fundamental al cambiar la función de pérdida a "sparse_categorical_crossentropy". Con todas estas modificaciones y adaptaciones, obtuvimos un conjunto de experimentos exhaustivos para evaluar el rendimiento de la red en diversas condiciones.

Realizamos otra modificación importante en la red al ajustar los parámetros de entrada en función del conjunto de datos que se estaba utilizando. Los valores afectados incluyeron la frecuencia de muestreo, que se configuró en 200 Hz, el número de canales que se incrementó de 8 a 62 canales, y el número de clases, con experimentos realizados tanto con 2 como con 3 clases. Además, adaptamos el tamaño de entrada en base a la longitud de las muestras. En nuestros estudios, examinamos muestras de 1 segundo, 3 segundos y la media de 10 segundos, permitiéndonos evaluar cómo la red respondía a diferentes intervalos temporales.

Finalmente, en cuanto a la selección de hiperparámetros, durante las etapas iniciales de nuestra investigación, realizamos pruebas con diversas configuraciones de parámetros como *learning_rate*, *dropout* y *batch_size*. Sin embargo, al final decidimos mantener los valores originales que estaban presentes en la red, ya que mostraron un rendimiento adecuado. Los valores elegidos se pueden ver en *Tabla 4.1*:

| Learning rate | Dropout | Batch size |
|---------------|---------|------------|
| 0.001 | 0.25 | 128 |

Tabla 4.1: Hiperparámetros de la red

4.4. Tipos de clasificación

4.4.1. Clasificación multiclase

En el contexto de la clasificación de emociones en segmentos de EEG, donde buscamos etiquetar los videos como "feliz", "neutro" o "triste", es esencial contar con métodos de evaluación efectivos para medir el rendimiento del modelo. Aquí se describen varios métodos de evaluación comunes que hemos utilizado en la clasificación multiclase:

Accuracy: El accuracy es la proporción de predicciones correctas en relación con el total de predicciones realizadas por el modelo. En el caso de la clasificación multiclase, se calcula como la suma de las predicciones correctas para cada clase dividida por el total de predicciones, como vemos en la [4.1]:

$$Acc = \frac{VP + VN}{Total\ predicciones} \quad [4.1]$$

siendo *accuracy* (Acc), verdaderos positivos (VP) y verdaderos negativos (VN)

Es importante mencionar que la precisión puede ser engañosa en casos de desbalance de clases, donde una clase es mucho más frecuente que la otra. Por ello, nos aseguramos del balance de clases en nuestra base de datos, ya que de cada etiqueta contamos con el mismo número de videos.

Coefficiente Kappa: El coeficiente Kappa es una medida de concordancia entre las predicciones del modelo y las etiquetas reales, teniendo en cuenta la posibilidad de que las coincidencias se deban al azar. Es una métrica más robusta que la exactitud, especialmente cuando las clases están desbalanceadas. El valor del coeficiente Kappa puede variar de -1 a 1, donde 1 indica un acuerdo perfecto y 0 indica un acuerdo similar al azar. Se calcula como vemos en la Ecuación [4.2]:

$$Coeficiente\ Kappa = \frac{EO - EE}{1 - EE} \quad [4.2]$$

donde la "Exactitud Observada (EO)" es la proporción de coincidencias observadas entre las predicciones y las etiquetas reales, y la "Exactitud Esperada (EE)" es la proporción de coincidencias que se esperarían al azar.

Matriz de Confusión: La matriz de confusión es una herramienta que muestra el desempeño del modelo detallado para cada clase. Permite visualizar las predicciones verdaderas y falsas, tanto positivas como negativas, para cada categoría. Esto es útil para identificar dónde el modelo puede estar teniendo dificultades y qué clases podrían estar siendo confundidas entre sí. En la Figura 4.4 vemos una matriz de ejemplo. Además, utilizando los vectores de test y predicción generados, creamos una perspectiva adicional para visualizar los aciertos y fallos de nuestro modelo.

Porcentaje de acierto en cada sesión: evalúa el porcentaje de aciertos del modelo de clasificación en cada una de las 3 sesiones. Las sesiones en este pueden ser comprendidas como conjuntos de datos separados, en los que a cada sujeto se le ha realizado la misma prueba. Así con esta función se proporciona una visión de cómo el modelo se comporta en diferentes sesiones.

Porcentaje de clasificación en cada video: función que analiza el desempeño del modelo de clasificación en la predicción de emociones para cada video individual. Proporciona una visión detallada de cómo el modelo está clasificando cada video en términos de las emociones "triste", "neutro" y "feliz". Lo que permite analizar cómo el modelo está clasificando las emociones en cada caso.

| 3D | | ETIQUETA PREDICHA | | |
|--------------------|---------|-------------------|-----------------|---------------|
| | | SAD | NEUTRAL | HAPPY |
| ETIQUETA VERDADERA | SAD | SAD-SAD | SAD-NEUTRAL | SAD-HAPPY |
| | NEUTRAL | NEUTRAL-SAD | NEUTRAL-NEUTRAL | NEUTRAL-HAPPY |
| | HAPPY | HAPPY-SAD | HAPPY-NEUTRAL | HAPPY-HAPPY |

Figura 4.4: Esquema de matriz de confusión para nuestra clasificación multiclase.

4.4.2. Clasificación binaria

Tras examinar la clasificación multiclase y revisar los resultados obtenidos, llegamos a la conclusión de que los videos etiquetados como "neutros" podrían ser un tanto problemáticos para la clasificación. Estos videos tendían a confundirse con las categorías "feliz" y "triste", o viceversa. Así que, basados en esta observación, tomamos la decisión de eliminar los 5 videos etiquetados como neutros de nuestra base de datos para realizar una clasificación binaria, centrándonos únicamente en las categorías "feliz" y "triste". El propósito de esta modificación era verificar si el modelo que había demostrado el mejor desempeño en la clasificación multiclase podría lograr una mejora en sus resultados bajo estas condiciones.

Para esta tarea, hemos empleado el modelo de EEG-Inception modificado en su última capa y validado con un conjunto de 2 videos de cada sujeto. En esta ocasión, hemos variado la longitud de las muestras de entrada para evaluar con cual obtenemos mejores resultados. Hemos utilizado tres configuraciones de longitud: 3 segundos, para replicar el estudio de comparación seleccionado; 1 segundo, al igual que en el caso de EEG-Inception original; y, por último, hemos tomado muestras de 1 segundo y calculado el promedio de las probabilidades de salida en intervalos de 10 segundos. Esta decisión se basa en la idea de que las emociones pueden variar significativamente en períodos cortos y se pueden experimentar mejor en un intervalo de tiempo más extenso.

Este enfoque nos permitirá determinar si la clasificación binaria y la variación en la longitud de las muestras de entrada tienen un impacto positivo en la precisión de nuestros modelos en comparación con la clasificación multiclase original.

Por lo tanto, con el fin de evaluar la eficacia de nuestro modelo de clasificación binaria, hemos elegido cinco métricas fundamentales que nos permitieron analizar su desempeño. De manera similar a lo que hicimos en la clasificación multiclase, calculamos la precisión, el coeficiente kappa y la matriz de confusión, métricas que ya fueron explicadas en la sección anterior. Y, consideramos otras métricas como:

Matriz de confusión: Permite visualizar las predicciones verdaderas y falsas, tanto positivas como negativas, para cada categoría. Esto es útil para identificar dónde el modelo puede estar teniendo dificultades y qué clases podrían estar siendo confundidas entre sí. En este caso el esquema difiere un poco con respecto a la clasificación multiclase y lo vemos explicado en la Figura 4.5.

F1-Score: Esta métrica nos proporcionará un equilibrio entre precisión y recall, ayudándonos a comprender cómo nuestro modelo maneja las predicciones precisas. En este contexto, nos permite evaluar cuántos de los videos clasificados como "felices" y "tristes" son verdaderamente felices y tristes, respectivamente, y cómo el modelo maneja los falsos positivos y falsos negativos en esta tarea de clasificación binaria. En la *Ecuación* [4.3] vemos como se calcula el F1-Score:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad [4.3]$$

siendo precisión y recall, lo equivalente a sensibilidad y especificidad, respectivamente.

- **Sensibilidad (feliz):** indica la capacidad del modelo para detectar una emoción como feliz, es decir, expresa cuan "sensible" es el modelo cuando la emoción es feliz. En la *Ecuación* [4.4] vemos como se calcula.

$$Sensibilidad = \frac{VP}{VP + FP} \quad [4.4]$$

Una alta sensibilidad indica que el modelo tiene pocas falsas alarmas al clasificar como "feliz".

- **Especificidad (triste):** La especificidad indica la capacidad que tiene el modelo de identificar como emoción triste a los segmentos que efectivamente lo son. En la *Ecuación* 4.5 vemos como se calcula:

$$Recall = \frac{VN}{VN + FN} \quad [4.5]$$

El F1-Score tiende a ser más bajo cuando una de las dos métricas (precisión o recall) es muy baja.

Gráfico acierto emociones: gráfico que permite visualizar cómo las probabilidades de predicción evolucionan en el tiempo para videos específicos y en diferentes sesiones, lo que puede ayudar a comprender cómo el modelo está tomando decisiones en diferentes contextos temporales y de sesiones.

Asimismo, empleamos las funciones previamente desarrolladas para la clasificación multiclase, ya que dan una visión más detallada sobre el rendimiento de nuestro modelo.

| 2D | | ETIQUETA PREDICHA | |
|--------------------|-------|--|--|
| | | SAD | HAPPY |
| ETIQUETA VERDADERA | SAD | <p>VERDADEROS NEGATIVOS</p> <p>SAD - SAD</p> | <p>FALSOS NEGATIVOS</p> <p>SAD - HAPPY</p> |
| | HAPPY | <p>FALSOS POSITIVOS</p> <p>HAPPY - SAD</p> | <p>VERDADEROS POSITIVOS</p> <p>HAPPY - HAPPY</p> |

Figura 4.5: Esquema matriz de confusión para clasificación binaria.

5. Capítulo V

Resultados

Índice

| | |
|---|----|
| 5.1. Introducción | 44 |
| 5.2. Resultados de clasificación multiclase | 44 |
| 5.3. Resultados de clasificación binaria..... | 51 |

5.1. Introducción

La siguiente sección se centra en la presentación y análisis de los resultados obtenidos a través de nuestro clasificador de emociones. En esta etapa del estudio, exploramos en profundidad cómo los modelos de clasificación de emociones se desempeñaron en la tarea de reconocer y categorizar las emociones de los sujetos en respuesta a diversos estímulos visuales.

5.2. Resultados de clasificación multiclase

Después de completar nuestro estudio y llevar a cabo varias pruebas con diferentes parámetros y variaciones de nuestro modelo, vamos a presentar las tres situaciones en la que obtuvimos mejores resultados en nuestra investigación:

1. **MODELO 1:** Modelo EEG-Inception en su forma original.
2. **MODELO 2:** Modelo EEG-Inception con modificación de la última capa a global max pooling y validación con 3 videos.
3. **MODELO 3:** Modelo EEG-Inception con modificación de la última capa a global max pooling y validación durante la sesión 3.

Los primeros valores que calculamos para seguir el rendimiento de nuestro modelo es la *accuracy* y el coeficiente kappa. En la *Tabla 5.1* y *Tabla 5.2* mostraremos los resultados de estas métricas respectivamente

| SUJETO | ACC 1 | ACC 2 | ACC 3 |
|--------------|--------|--------|--------|
| 1 | 0.3900 | 0.4106 | 0.3851 |
| 2 | 0.4548 | 0.4244 | 0.3831 |
| 3 | 0.5845 | 0.6326 | 0.6228 |
| 4 | 0.4332 | 0.4401 | 0.4568 |
| 5 | 0.5678 | 0.5756 | 0.5187 |
| 6 | 0.4460 | 0.5363 | 0.4499 |
| 7 | 0.3448 | 0.3861 | 0.3605 |
| 8 | 0.5963 | 0.555 | 0.6395 |
| 9 | 0.5953 | 0.6326 | 0.6198 |
| 10 | 0.4597 | 0.4096 | 0.3635 |
| 11 | 0.6513 | 0.5953 | 0.6189 |
| 12 | 0.4764 | 0.4666 | 0.4676 |
| 13 | 0.5452 | 0.5195 | 0.4695 |
| 14 | 0.6081 | 0.6444 | 0.6523 |
| 15 | 0.6081 | 0.6287 | 0.6660 |
| MEDIA | 0.5091 | 0.5158 | 0.4995 |

Tabla 5.1: Valores obtenidos de *accuracy* en los 3 modelos de clasificación multiclase. ACC 1 (modelo 1), ACC 2 (modelo 2) y ACC 3 (modelo 3).

| SUJETO | KAPPA 1 | KAPPA 2 | KAPPA 3 |
|--------------|---------|---------|---------|
| 1 | 0.0921 | 0.1219 | 0.0851 |
| 2 | 0.1881 | 0.1432 | 0.0826 |
| 3 | 0.3771 | 0.4483 | 0.4328 |
| 4 | 0.1507 | 0.1602 | 0.1826 |
| 5 | 0.3481 | 0.3604 | 0.2725 |
| 6 | 0.1636 | 0.3008 | 0.1661 |
| 7 | 0.0258 | 0.0869 | 0.0490 |
| 8 | 0.3903 | 0.3269 | 0.4567 |
| 9 | 0.3916 | 0.4480 | 0.4293 |
| 10 | 0.1942 | 0.1191 | 0.0510 |
| 11 | 0.4753 | 0.3912 | 0.4278 |
| 12 | 0.2159 | 0.2041 | 0.2037 |
| 13 | 0.3203 | 0.2830 | 0.2074 |
| 14 | 0.4104 | 0.4675 | 0.4788 |
| 15 | 0.4130 | 0.4437 | 0.4994 |
| MEDIA | 0.2284 | 0.2506 | 0.2066 |

Tabla 5.2: Valores obtenidos del coeficiente Kappa en los 3 modelos de clasificación multiclase. KAPPA 1 (modelo 1), KAPPA 2 (modelo 2) y KAPPA 3 (modelo 3).

Los resultados de la *Tabla 5.1* representan las *accuracy* para los tres modelos de clasificación de emociones (triste, neutro y feliz). Cada columna de la tabla corresponde a un modelo diferente y cada fila representa un sujeto diferente en el estudio.

Podemos observar que estos resultados varían entre los sujetos y entre los modelos, viendo también un seguimiento en los que los sujetos que mejor clasifican en un modelo también lo suelen hacer en los otros dos. En promedio de todos los sujetos, el modelo 2 tiene la tasa de precisión más alta con una media de 51.58%, siguiéndole el modelo original con 50.91% y finalmente el modelo 3 con una precisión promedio de 49.95%.

Esto indica que, el modelo EEG-Inception con variación de la última capa a *global max pooling* y su conjunto de validación 3 videos, es ligeramente mejor en términos de precisión que las últimas dos variantes. Además, continuaremos estudiando el coeficiente kappa para determinar si estas diferencias en precisión son significativas.

En general, estos valores de kappa indican que los modelos tienen un rendimiento variable en la clasificación de emociones. Los valores de kappa cercanos a 0 indican una concordancia débil, mientras que valores más cercanos a 1 indican una concordancia más fuerte. En promedio, los valores obtenidos son bastante cercanos a 0, por lo que sugiere que aún hay margen de mejora en la capacidad de estos modelos para clasificar emociones de manera precisa y consistente.

Una métrica adicional de gran relevancia en nuestro análisis es la matriz de confusión, la cual proporciona una comprensión más completa del desempeño de los modelos. Nos permite identificar hacia qué clases tienden a inclinarse más y en cuáles tienen dificultades. En las ilustraciones 10, 11 y 12 presentamos las matrices de confusión de nuestros tres modelos. Estas matrices han sido calculadas con la suma de las predicciones de los 15 sujetos.

4 Resultados

La matriz de confusión para todos los experimentos y sujetos del modelo 1, 2 y 3 se presenta en la Figura 5.1, 5.2 y 5.3.

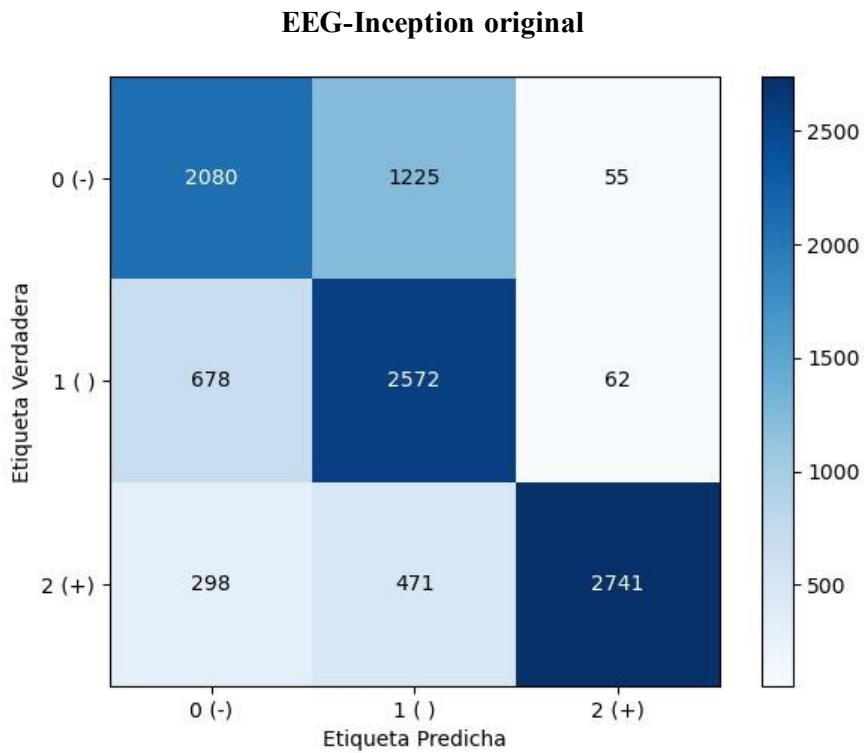


Figura 5.1: Matriz de confusión con la suma de predicciones del modelo 1

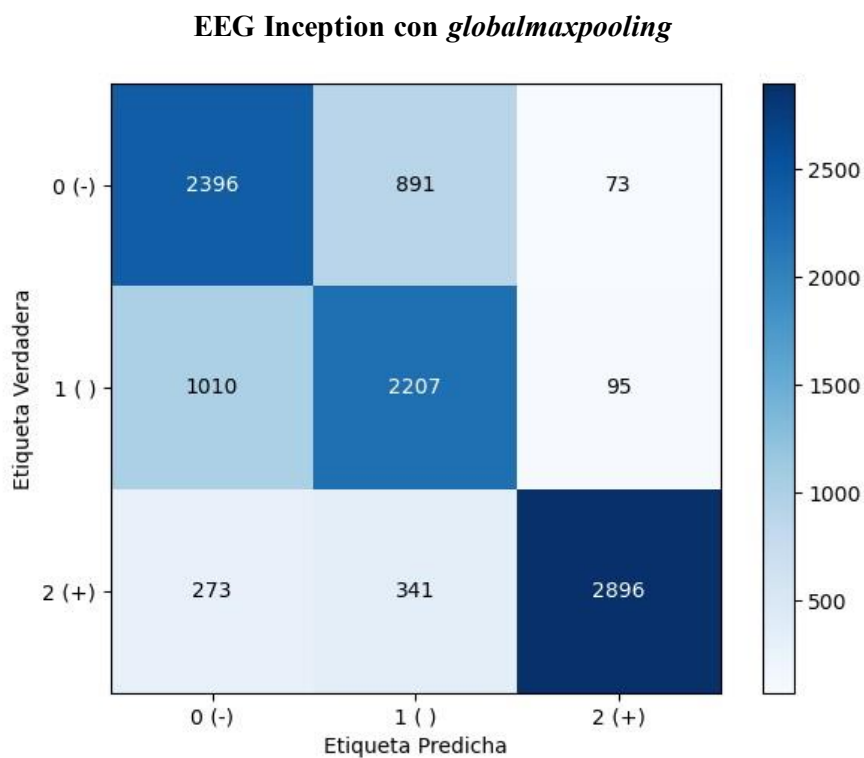


Figura 5.2: Matriz de confusión con la suma de predicciones del modelo 2

EEG Inception con *globalmaxpooling* y validación sesión 3

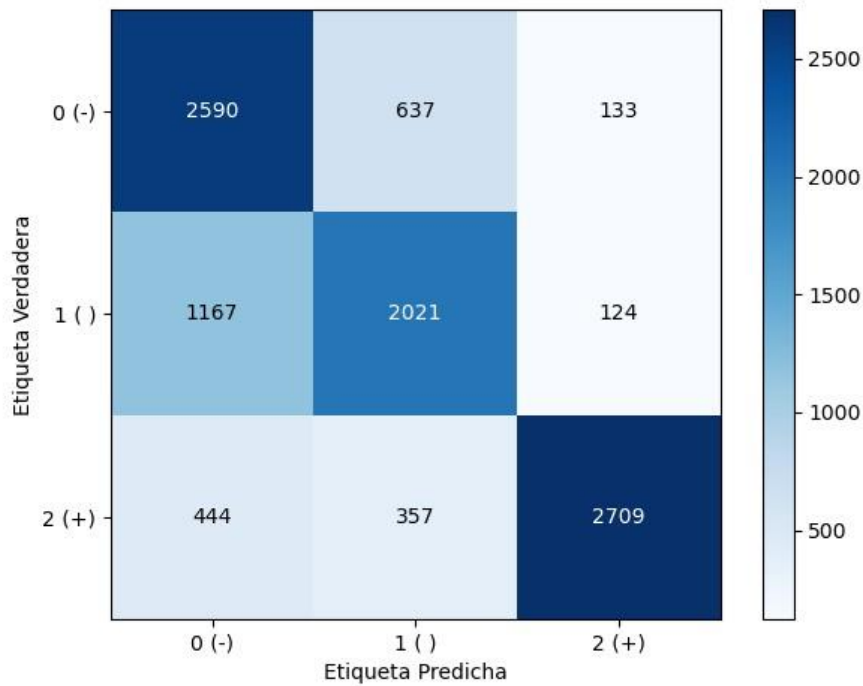


Figura 5.3: Matriz de confusión con la suma de predicciones del modelo 3.

En las tres matrices de confusión, se observa una tendencia común de equivocaciones entre las clases "feliz" y "neutro". En el Modelo 1, se tiende a clasificar erróneamente los videos felices como neutros. En el Modelo 3, ocurre lo contrario, con los videos neutros etiquetados incorrectamente como felices. En el Modelo 2, esta tendencia es menos pronunciada en ambos casos y se cometen menos errores en comparación con los otros modelos. Estos errores representan las discrepancias más notables en el desempeño de nuestros modelos, ya que la diagonal principal, que representa los aciertos, es bastante destacada en todas las matrices y demuestra que, en general, los tres modelos tienen un alto número de predicciones correctas.

El Modelo 2 continúa destacándose como el modelo con un mejor rendimiento, como mencionamos previamente al analizar las métricas anteriores. Para realizar un análisis más detallado de esta métrica, seleccionaremos varios sujetos que puedan proporcionar información significativa sobre el desempeño de nuestro Modelo 2, considerado el mejor. Elegiremos a 3 sujetos en función de su *accuracy* y coeficiente kappa obtenidos, seleccionando al sujeto con el peor rendimiento, al que obtuvo el mejor rendimiento y a uno que esté cerca del promedio.

En la *Tabla 5.3* presentamos a los sujetos elegidos con sus respectivos valores asociados y en las ilustraciones 13 y 14 sus matrices de confusión:

| SUJETO | ACCURACY | KAPPA |
|--------|----------|--------|
| 7 | 0.3861 | 0.0869 |
| 14 | 0.6444 | 0.4675 |

Tabla 5.3: Tabla con los valores de *accuracy* y coeficiente kappa de los sujetos 7 y 14, siendo estos el sujeto con peor y mejor rendimiento en nuestro modelo respectivamente.

4 Resultados

Se presentan en las figuras 5.5 y 5.4 las matrices de confusión de los sujetos 7 y 14 respectivamente.

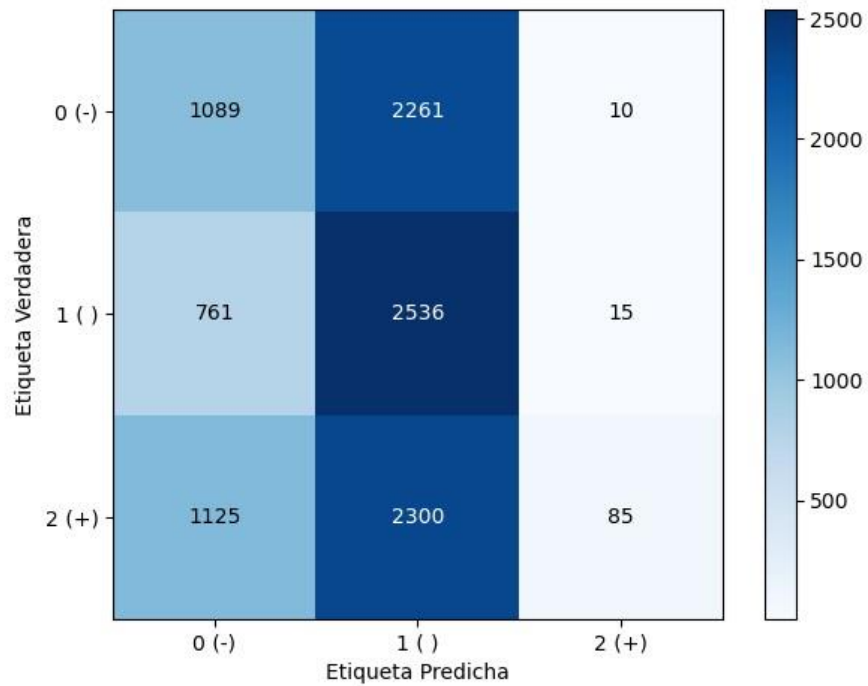


Figura 5.5: Matriz de confusión del sujeto 7 en el modelo 2

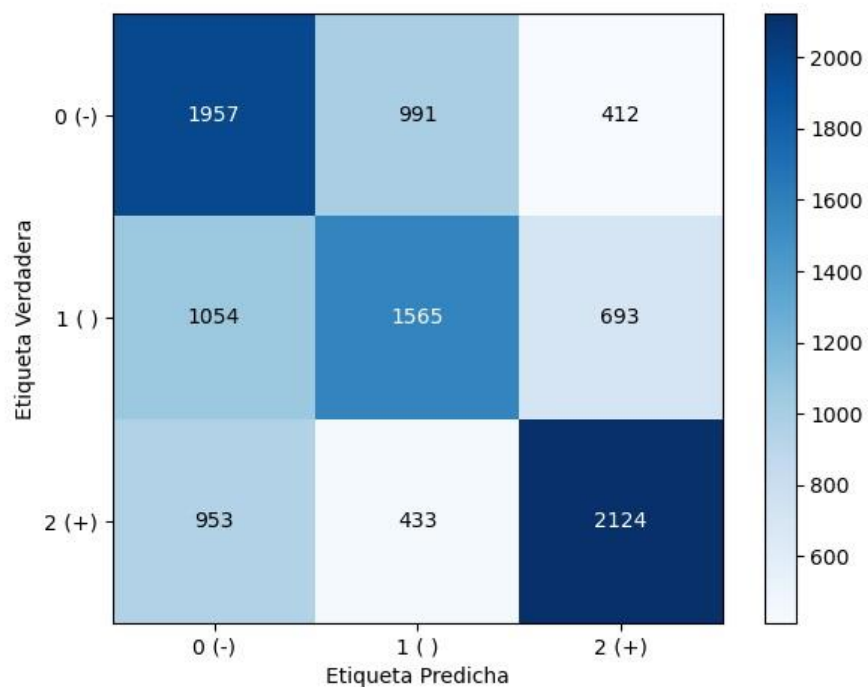


Figura 5.4: Matriz de confusión del sujeto 14 en el modelo 2

4 Resultados

El propósito de estas matrices es resaltar las variaciones en el desempeño de nuestro modelo en diferentes sujetos. En un sujeto, observamos una clasificación coherente, donde la mayoría de las emociones se etiquetan correctamente. Sin embargo, en el caso del sujeto 7, notamos que la clasificación tiende a ser mayoritariamente neutra; el modelo tiene dificultades para distinguir las emociones en este sujeto y comete numerosos errores de clasificación. Observamos que el sujeto bien clasificado tiene la misma tendencia que la matriz de confusión general de nuestro modelo, lo que tiene sentido.

Con el objetivo de profundizar en la evaluación de nuestros modelos, consideramos valioso analizar cuántos videos cada modelo clasifica correctamente en cada sujeto, así como identificar los videos en los que los modelos cometen más errores. En la *Tabla 5.4*, presentamos la cantidad de videos correctamente clasificados en cada sujeto en el modelo 2, que es con el que hemos conseguido mejor rendimiento. Además, hemos incluido una representación visual utilizando colores: verde para la emoción feliz, blanco para la emoción neutra y rojo para la emoción triste. Donde encuentres una "X," indica que la clasificación fue correcta para ese video y ese sujeto.

Modelo 2 - Modelo EEG-Inception con modificación de la última capa a global max pooling y validación con 3 videos.

| SUJETO | Nº VIDEOS ACERTADOS | VIDEOS QUE ACIERTA | | | | | | | | | | | | | | |
|--------|---------------------|--------------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 1 | 8 | X | X | X | X | X | | | X | | | X | | X | | |
| 2 | 5 | | X | | | X | | | X | | | X | | X | | |
| 3 | 13 | X | X | X | X | X | X | X | X | X | X | X | | X | X | |
| 4 | 9 | X | X | | | X | | X | X | | X | X | | X | X | |
| 5 | 11 | X | | X | X | X | X | X | X | X | X | | | X | X | |
| 6 | 9 | X | | X | X | | X | | | X | X | | X | X | X | |
| 7 | 5 | | X | | | X | | | X | | | X | | X | | |
| 8 | 9 | X | X | | X | | X | X | | X | X | | | | X | X |
| 9 | 10 | X | | X | X | | X | X | | X | X | | X | | X | X |
| 10 | 6 | | | X | X | | | X | | | | X | X | | | X |
| 11 | 10 | X | | | | X | X | X | X | X | X | X | | X | X | |
| 12 | 10 | X | X | X | X | X | | X | X | | | X | | X | | X |
| 13 | 9 | X | X | | | X | | | X | X | X | X | | X | | X |
| 14 | 14 | X | X | X | X | X | X | X | X | X | X | X | X | X | | X |
| 15 | 10 | X | | X | X | | X | X | | X | X | | X | | X | X |

Tabla 5.4: Número de videos acertados por cada sujeto en el modelo 2. Se muestra con una X los videos que cada sujeto acierta. Los videos en verde corresponden a feliz, los blancos a neutro y los rojos a triste.

De estos resultados sacamos otro valor relevante, que es el % de acierto en videos de cada modelo. Presentando en la *Tabla 5.5* Porcentaje de aciertos de videos totales. los porcentajes de los 3 modelos de clasificación:

| % ACIERTO EN VIDEOS | 1 | 2 | 3 |
|---------------------|-------|-------|-------|
| | 60,4% | 61,3% | 56,4% |

Tabla 5.5 Porcentaje de aciertos de videos totales.

A partir de estos resultados, podemos notar que nuestros modelos a menudo cometen errores en los mismos videos para un sujeto dado. Con la *Tabla 5.4* también destacamos otra vez la gran diferencia de clasificación entre sujetos. Observamos que algunos sujetos logran una precisión cercana al 100% en la clasificación de videos, mientras que otros apenas alcanzan un 30%. Esto resulta en una precisión promedio global del 61.3% en el mejor modelo de clasificación, que, una vez más, es el Modelo 2.

Finalmente, vamos a introducir una métrica que hemos desarrollado para evaluar el rendimiento de nuestro modelo a lo largo del estudio. Nos centraremos en presentar el Modelo 2, ya que es el que hemos seleccionado como el mejor. Esta métrica es el "% de acierto por sesión" que presentamos en la *Tabla 5.6*.

A partir de estos resultados, no identificamos una tendencia clara entre los sujetos. Observamos casos en los que algunos sujetos tienen un mejor desempeño en la primera sesión y luego empeoran, mientras que otros comienzan con un rendimiento inferior y mejoran en la tercera sesión. Además, el hecho de que las sesiones tengan un espacio de tres semanas entre ellas parece afectar a las clasificaciones, ya que notamos que, en algunos sujetos, la precisión varía en casi un 10% entre sesiones, ya sea mejorando o empeorando.

| SUJETO | SESIÓN 1 (%) | SESIÓN 2 (%) | SESIÓN 3 (%) |
|--------------|--------------|--------------|--------------|
| 1 | 42.46 | 39.69 | 33.18 |
| 2 | 43.43 | 39.42 | 39.19 |
| 3 | 44.61 | 58.01 | 60.78 |
| 4 | 43.87 | 43.25 | 41.19 |
| 5 | 47.44 | 52.39 | 49.97 |
| 6 | 42.37 | 49.91 | 46.55 |
| 7 | 40.93 | 30.32 | 38.07 |
| 8 | 60.61 | 51.62 | 61.58 |
| 9 | 60.37 | 59.19 | 54.95 |
| 10 | 43.58 | 33.03 | 33.77 |
| 11 | 58.99 | 54.6 | 63.05 |
| 12 | 50.94 | 40.01 | 38.45 |
| 13 | 45.93 | 53.54 | 42.31 |
| 14 | 55.45 | 54.33 | 56.57 |
| 15 | 53.12 | 63.17 | 64.88 |
| MEDIA | 48.94 | 48.17 | 48.30 |

Tabla 5.6: Tabla de los porcentajes de aciertos que cada sujeto tuvo en cada una de las tres sesiones

5.3. Resultados de clasificación binaria

Presentamos ahora los resultados correspondientes a la clasificación binaria. En primer lugar, en la *Tabla 5.7* y *Tabla 5.8* tenemos las métricas de accuracy y f1-score.

Estos resultados muestran que el *accuracy* y el F1-score son similares en las longitudes de 3 y 1 segundo, mientras que se nota una gran mejora cuando realizamos la media de 10 segundos, lo que indica que la capacidad de los modelos para clasificar videos como "felices" o "tristes" mejora cuando se considera un intervalo de tiempo más largo. Sin embargo, las variaciones entre sujetos sugieren que el rendimiento puede depender de las diferencias individuales en las respuestas emocionales.

Accuracy

| SUJETO | 3 s | 1 s | Media 10s |
|--------------|--------|--------|-----------|
| 1 | 0.7965 | 0.7913 | 0.8196 |
| 2 | 0.6794 | 0.6632 | 0.6774 |
| 3 | 0.7684 | 0.7921 | 0.8182 |
| 4 | 0.6973 | 0.7036 | 0.7067 |
| 5 | 0.6851 | 0.6918 | 0.7199 |
| 6 | 0.7859 | 0.7234 | 0.7170 |
| 7 | 0.7952 | 0.7866 | 0.7991 |
| 8 | 0.7289 | 0.8434 | 0.8768 |
| 9 | 0.8013 | 0.8068 | 0.8138 |
| 10 | 0.7018 | 0.7345 | 0.7317 |
| 11 | 0.9246 | 0.9157 | 0.9457 |
| 12 | 0.7075 | 0.6514 | 0.6598 |
| 13 | 0.8184 | 0.8268 | 0.8402 |
| 14 | 0.7969 | 0.7818 | 0.8123 |
| 15 | 0.9561 | 0.9550 | 0.9692 |
| MEDIA | 0.7762 | 0.7778 | 0.7938 |

Tabla 5.7: Valores obtenidos de *accuracy* en las distintas configuraciones de longitud de muestras de entrada. 3 segundos (3s), 1 segundo (1s) y media de 10 segundos (10s)

F1-score

| SUJETO | 3 s | 1 s | Media 10s |
|--------------|--------|--------|-----------|
| 1 | 0.7638 | 0.7572 | 0.7912 |
| 2 | 0.5583 | 0.5233 | 0.5473 |
| 3 | 0.8044 | 0.8132 | 0.8390 |
| 4 | 0.5868 | 0.6213 | 0.6183 |
| 5 | 0.7304 | 0.7360 | 0.7651 |
| 6 | 0.8154 | 0.7844 | 0.7824 |
| 7 | 0.7644 | 0.7474 | 0.7575 |
| 8 | 0.7897 | 0.8656 | 0.8920 |
| 9 | 0.8300 | 0.8339 | 0.8414 |
| 10 | 0.5923 | 0.6526 | 0.6460 |
| 11 | 0.9241 | 0.9172 | 0.9455 |
| 12 | 0.6517 | 0.5916 | 0.5958 |
| 13 | 0.7851 | 0.7978 | 0.8162 |
| 14 | 0.8086 | 0.7835 | 0.8156 |
| 15 | 0.9576 | 0.9571 | 0.9707 |
| MEDIA | 0.7575 | 0.7588 | 0.7749 |

Tabla 5.8: Valores obtenidos de coeficiente kappa en las distintas configuraciones de longitud de muestras de entrada. 3 segundos (3s), 1 segundo (1s) y media de 10 segundos (10s)

Otra métrica importante a tener en cuenta son las matrices de confusión. Presentamos en las Figuras 5.6, 5.7 y 5.8 las matrices de confusión correspondientes a los resultados de clasificación binaria.

Segmentos de 3 segundos de longitud

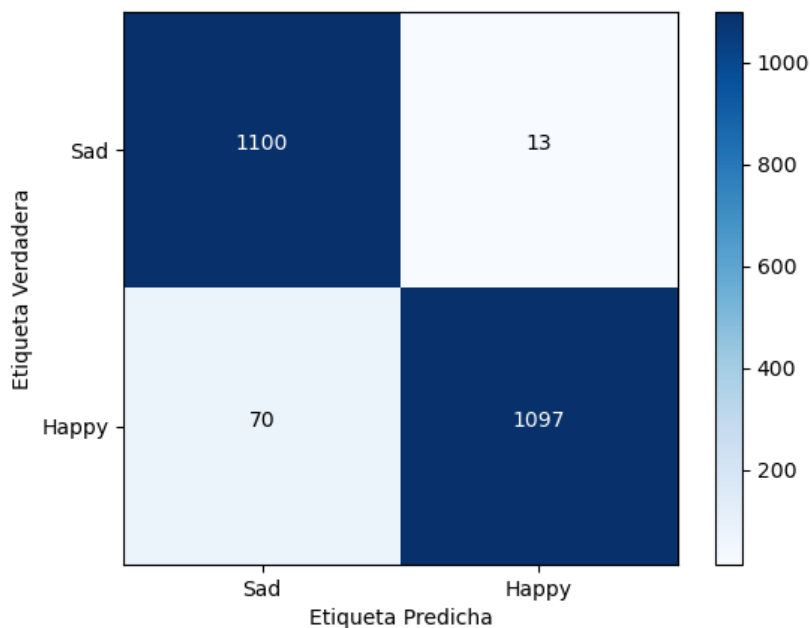


Figura 5.6: Matriz de confusión del modelo de clasificación donde la longitud de los segmentos de entrada es de 3 segundos.

Segmentos de 1 segundo de longitud

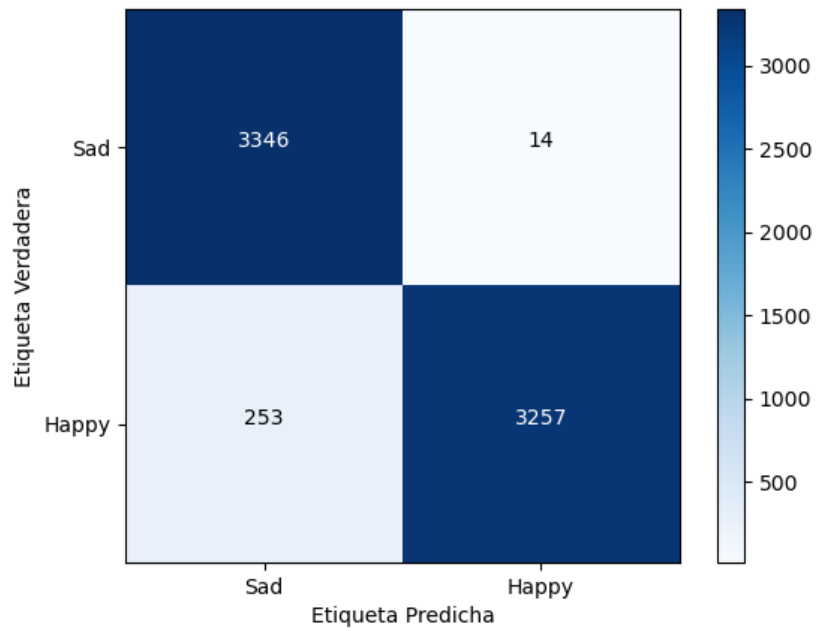


Figura 5.7: Matriz de confusión del modelo de clasificación donde la longitud de los segmentos de entrada es de 1 segundo.

Segmentos de media segmentos de 1 segundo en intervalos de 10 segundos.

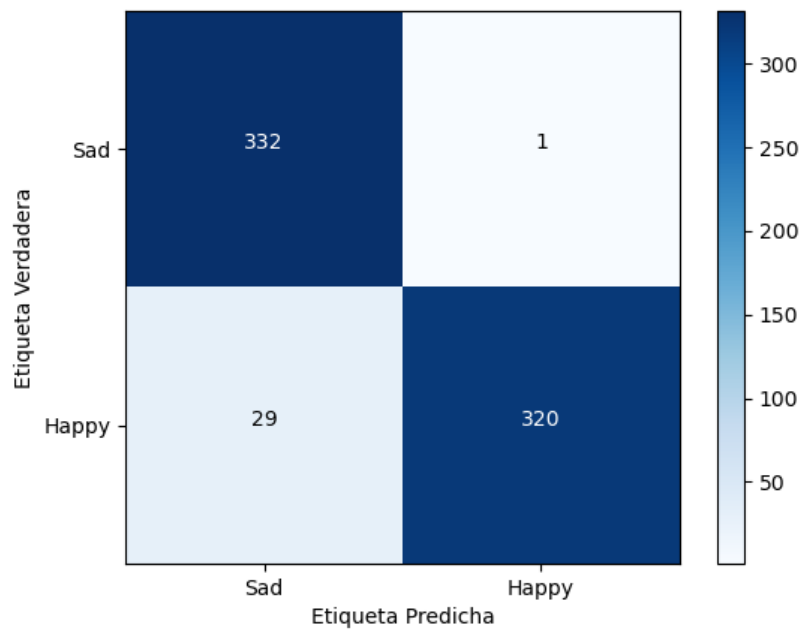


Figura 5.8: Matriz de confusión del modelo de clasificación donde la longitud de los segmentos de entrada es la media de los segmentos de 1 segundo en intervalos de 10 segundos.

En estas matrices, podemos notar que, a rasgos generales, las clasificaciones están equilibradas, lo que significa que los modelos tienen un alto nivel de precisión tanto en la clasificación de videos "felices" como "tristes". Luego, observamos que, en algunos sujetos, existe una tendencia a clasificar de manera predominante una emoción, lo que a veces lleva a errores al intentar clasificar la emoción opuesta. Sin embargo, en general, se aprecian resultados satisfactorios en la capacidad de los modelos para distinguir entre las emociones "feliz" y "triste".

Además, en las Figura 5.9Figura 5.10, presentaremos las matrices de confusión más relevantes del modelo con segmentos de entrada de media de 1 segundo en intervalos de 10 segundos. Mostraremos tanto el sujeto con el mejor rendimiento en términos de *accuracy* y puntuación *F1-score*, como el sujeto con el peor rendimiento. Esto nos permitirá visualizar las diferencias entre sujetos y evaluar el desempeño individual de nuestro modelo.

Matriz de confusión del sujeto 12, con peor rendimiento

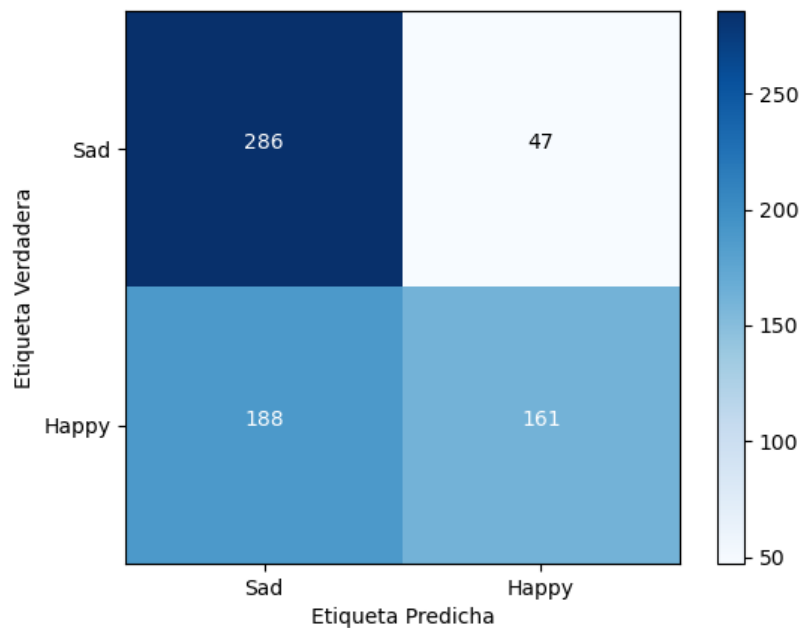


Figura 5.9: Matriz de confusión del sujeto 12 en el modelo elegido. Sujeto con peor rendimiento en este modelo.

Matriz de confusión del sujeto 15, sujeto con mejor rendimiento

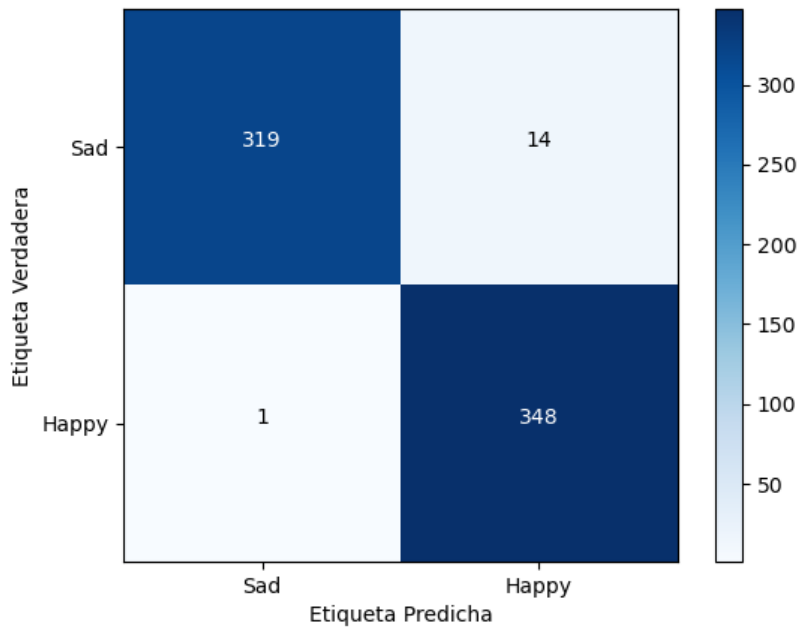


Figura 5.10: Matriz de confusión del sujeto 15 en el modelo elegido. Sujeto con mejor rendimiento en este modelo.

De manera similar a la clasificación multiclase, también queremos mostrar las métricas relacionadas con el porcentaje de aciertos en videos y la cantidad de videos clasificados correctamente. Presentamos en la *Tabla 5.10* los videos que ha acertado nuestro modelo en la configuración de entrada que mejor rendimiento hemos obtenido, la media de segmentos de 1 segundo en intervalos de 10 segundos. Presentamos estos valores ya que es donde obtenemos mejores resultados y podemos observar la gran mejora de clasificación de videos.

Al analizar estos resultados, se evidencia que en la clasificación binaria se obtiene una mayor tasa de aciertos, llegando incluso al punto en el que, en la configuración de media de 10 segundos, el modelo logra clasificar correctamente el 100% de los videos para 8 de los sujetos del estudio. Este resultado es particularmente destacable, ya que indica una alta mejora en la capacidad de los modelos para discernir entre las emociones "feliz" y "triste" en comparación con la clasificación multiclase. La reducción en la complejidad al trabajar con una clasificación binaria y en combinación con la media intervalos de 1 segundo parece haber contribuido significativamente a este éxito.

También presentamos en la *Tabla 5.9* el porcentaje de aciertos calculado de las tablas previas para las 3 configuraciones, lo cual refleja la mejora previamente mencionada.

| % ACIERTO EN VIDEOS | 1 | 2 | 3 |
|---------------------------|--------|--------|--------|
| | 83.33% | 84.00% | 87.33% |

Tabla 5.9: Tabla de porcentaje de aciertos de videos en total

Media de 10 segundos de segmentos de 1 segundo

| SUJETO | Nº VIDEOS ACERTADOS | VIDEOS QUE ACIERTA | | | | | | | | | |
|--------|---------------------|--------------------|---|---|---|---|---|----|----|----|----|
| | | 1 | 3 | 4 | 6 | 7 | 9 | 10 | 12 | 14 | 15 |
| 1 | 10 | X | X | X | X | X | X | X | X | X | X |
| 2 | 6 | | X | X | | X | | X | X | | X |
| 3 | 10 | X | X | X | X | X | X | X | X | X | X |
| 4 | 7 | | X | X | | X | | X | X | X | X |
| 5 | 8 | X | X | X | X | | X | X | X | X | |
| 6 | 7 | X | | X | X | | X | X | X | X | |
| 7 | 8 | X | X | X | X | X | X | | X | | X |
| 8 | 10 | X | X | X | X | X | X | X | X | X | X |
| 9 | 10 | X | X | X | X | X | X | X | X | X | X |
| 10 | 7 | X | X | X | | X | X | | X | | X |
| 11 | 10 | X | X | X | X | X | X | X | X | X | X |
| 12 | 8 | X | X | X | X | X | | X | X | | X |
| 13 | 10 | X | X | X | X | X | X | X | X | X | X |
| 14 | 10 | X | X | X | X | X | X | X | X | X | X |
| 15 | 10 | X | X | X | X | X | X | X | X | X | X |

Tabla 5.10: Número de videos acertados por cada sujeto en el modelo 2 con segmentos de media de 10 segundos. Se muestra con una X los videos que cada sujeto acierta. Los videos en verde corresponden a feliz y los rojos a triste.

Por último, una de las métricas más innovadoras son los gráficos en los que vemos hacia donde tiende la predicción de nuestro modelo según el sujeto y la sesión. Vamos a presentar en las *Figura 5.11* *Figura 5.13* *Figura 5.12* *Figura 5.14*, dos videos de los sujetos 12 y 15, que son los que hemos presentado anteriormente con información relevante para nuestro modelo. Los videos son el 3 y el 6, triste y feliz, respectivamente.

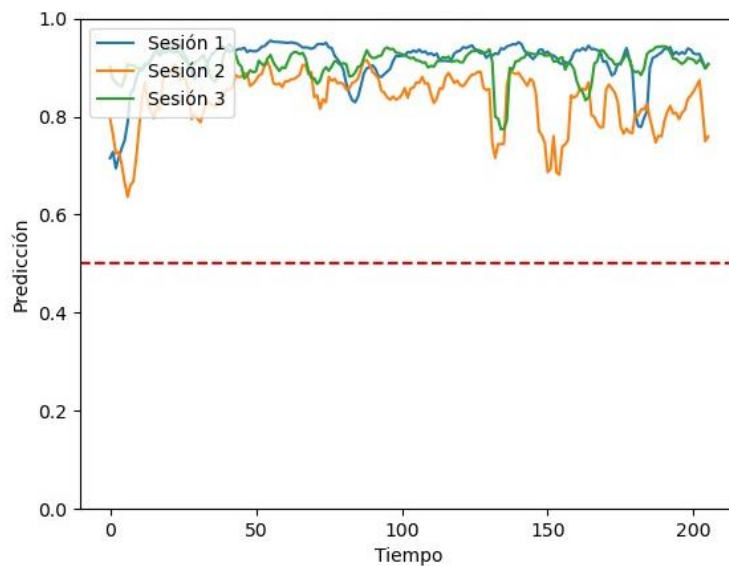


Figura 5.11: Gráfico de acierto de emoción feliz en video 3 del sujeto 15.

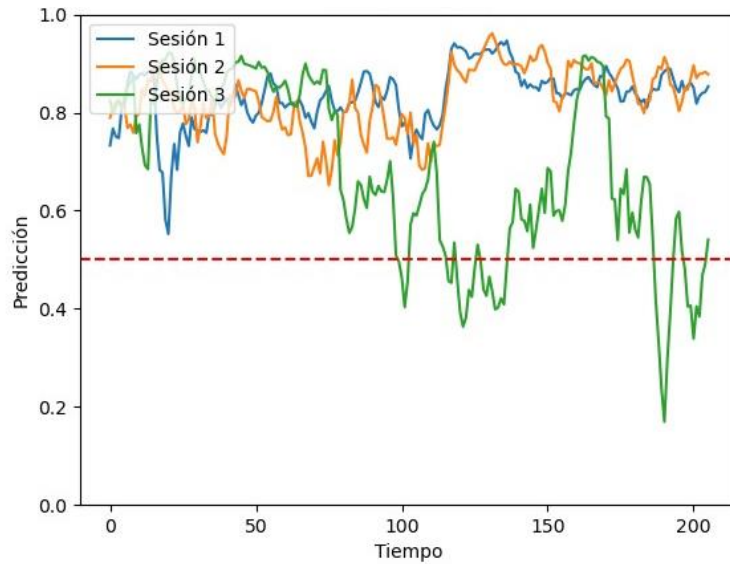


Figura 5.13: Gráfico de acierto de la emoción feliz en el video 3 del sujeto 12.

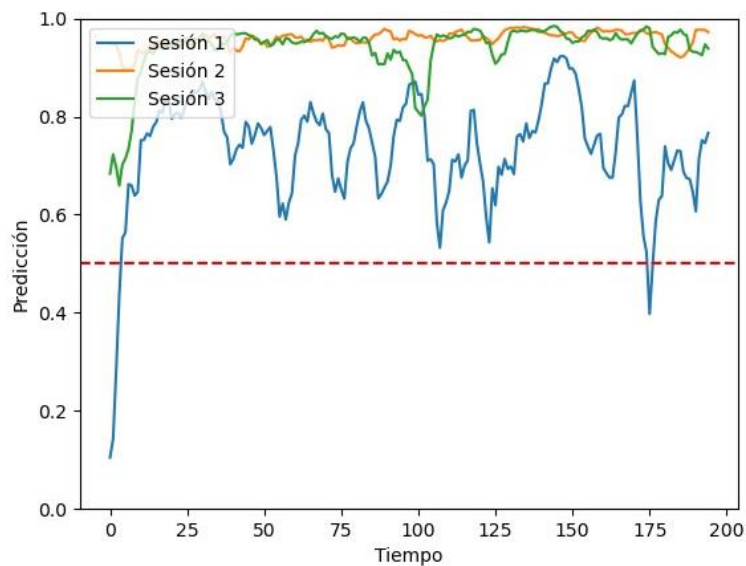


Figura 5.12: Gráfico de acierto de la emoción triste en el video 6 del sujeto 15.

Hemos presentado gráficos que representan el rendimiento del sujeto con el peor y el mejor desempeño en nuestro modelo, lo que explica las notables diferencias en los resultados entre ambos. Estos gráficos nos permiten visualizar las predicciones por segmento de emociones en videos específicos de cada sujeto.

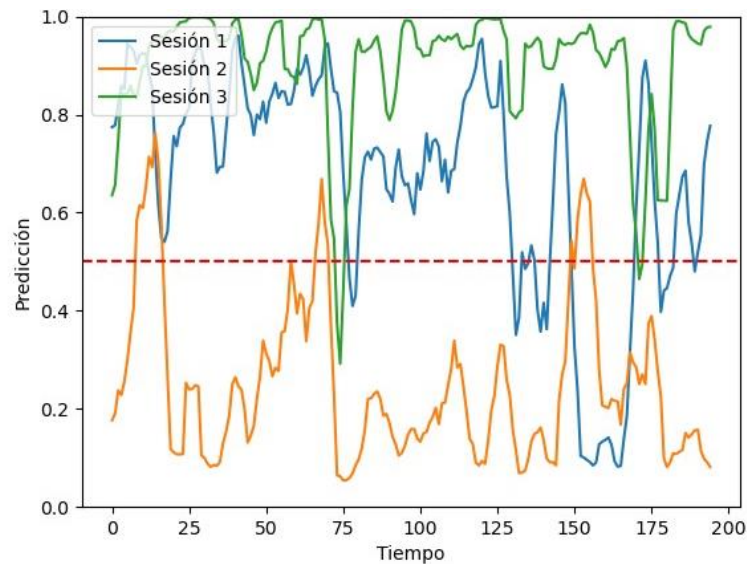


Figura 5.14: Gráfico de acierto de la emoción triste en el video 6 del sujeto 12.

En general, observamos que la emoción "feliz" tiende a clasificarse un poco mejor que la emoción "triste". Sin embargo, hay una discrepancia significativa entre el sujeto 15, que logra casi un 100% de precisión en estos dos videos, y el sujeto 12, que en alguna de las tres sesiones comete errores de clasificación y se confunde con respecto a la emoción que se está expresando en el video.

Observamos un patrón interesante en los resultados: en la mayoría de los sujetos, el modelo clasifica correctamente a lo largo de todo el video durante 1 o 2 sesiones. Sin embargo, hay en otra sesión donde notamos que en algunos sujetos la precisión disminuye, y existen segmentos del video donde el modelo comete errores en la clasificación de la emoción.

6. Capítulo VI

Discusión

Índice

| | |
|---|----|
| 6.1. Introducción | 60 |
| 6.2. Discusión | 60 |
| 6.3. Comparación con otros estudios similares | 61 |
| 6.4. Limitaciones | 63 |

6.1. Introducción

A lo largo de este capítulo, se abordarán aspectos esenciales relacionados con la evaluación y discusión de los resultados obtenidos en el marco de este TFG. Se analizarán en detalle los resultados alcanzados en la clasificación de emociones utilizando técnicas de DL aplicadas a EEG. Asimismo, se explorarán las comparaciones con investigaciones previas que han empleado enfoques similares o diferentes para la clasificación emocional. Además, se expondrán las limitaciones que se han encontrado durante el desarrollo del TFG, lo que proporcionará un contexto crítico para la comprensión de los resultados y sus implicaciones.

6.2. Discusión

Tras un análisis detenido de los resultados, se observan aspectos significativos que pueden explicar la influencia de la subjetividad y la variabilidad inter-sujeto en la clasificación de emociones basada en señales EEG. Estas observaciones tienen implicaciones importantes para la interpretación de los resultados y ofrecen una perspectiva más completa de la dinámica de las respuestas emocionales capturadas en el estudio.

Es notable que la experiencia subjetiva de las emociones por parte de los sujetos puede variar considerablemente, incluso cuando se les presenta el mismo estímulo emocional objetivo. Esto plantea preguntas acerca de la naturaleza de la emoción y cómo se refleja en las señales EEG. Por ejemplo, es posible que un sujeto experimente diferentes emociones al ver un video que objetivamente se clasifica como "feliz", pero la experiencia subjetiva del sujeto no se alinee con esta clasificación. Esto podría deberse a factores personales, estados de ánimo cambiantes o incluso la percepción individual de lo que constituye una emoción "feliz". En este contexto, es relevante considerar que la clasificación de emociones basada en señales EEG podría verse fuertemente influenciada por el estado emocional y subjetivo del sujeto en un momento particular durante los registros.

Además, se destaca la variabilidad en la experiencia emocional a lo largo del tiempo. Como hemos observado en la *Figura 5.14* estudiando el sujeto 2. Los sujetos podrían no experimentar las mismas emociones en sesiones diferentes. Por ejemplo, en la primera visualización de un video, el sujeto podría sentir felicidad, pero en una tercera visualización, debido a la familiaridad con la trama, es posible que no llegue a experimentar la misma emoción. O, al contrario, podría ocurrir que en la primera visualización el sujeto no experimente una emoción específica, pero tras repetir la visualización del video en la tercera sesión, podría comprender mejor la trama y sentir la emoción "correcta". Esta perspectiva subraya la importancia de considerar las diferencias en las respuestas emocionales a lo largo de múltiples visualizaciones y cómo estas variaciones pueden influir en la capacidad del modelo para clasificar correctamente las emociones en diferentes momentos.

En relación con las diferencias inter-sujeto, es relevante mencionar que los usuarios en los que el modelo multiclase solo acierta la categoría "neutro" podrían estar experimentando emociones menos marcadas en comparación con otros sujetos. Esto podría ser un indicativo de que algunos individuos tienen una respuesta emocional menos pronunciada o menos distintiva en términos de señales EEG cuando se enfrentan a estímulos emocionales. Estas diferencias individuales pueden influir en la capacidad del modelo para detectar y clasificar emociones en estos sujetos específicos.

Finalmente, se plantea la pregunta de por qué, a pesar de los errores en la clasificación multiclase, el modelo de clasificación binaria logra clasificar correctamente a algunos de los usuarios que habían sido erróneamente categorizados en el enfoque multiclase. Una posible explicación podría ser que el modelo binario se beneficia de la simplificación de la tarea al solo tener que discernir entre dos estados emocionales en lugar de múltiples categorías. Esto puede hacer que las diferencias sutiles en las respuestas emocionales entre usuarios sean más evidentes y, por lo tanto, más fácilmente identificables en el enfoque binario.

Estas observaciones resaltan la complejidad de la experiencia emocional y la variabilidad inter-sujeto en la clasificación de emociones basada en señales EEG. Los resultados sugieren que la subjetividad y las fluctuaciones temporales en las respuestas emocionales son factores importantes a considerar al interpretar los resultados de este estudio.

6.3. Comparación con otros estudios similares

Después de una exhaustiva revisión bibliográfica, identificamos un artículo que se centraba en una clasificación binaria (feliz y triste) utilizando la base de datos SEED. En dicho estudio, se evaluaron varios modelos de DL mediante la introducción de segmentos de 3 segundos en la red neuronal y la aplicación del método de validación cruzada *"leave one subject's data out"*[39].

Escogimos este estudio, debido a que sigue la misma metodología a la hora de dividir los conjuntos de datos. Así que, en base en estas condiciones, replicamos este enfoque utilizando nuestro propio modelo de clasificación EEG-Inception. El artículo de referencia utilizaba el F1-score como métrica de rendimiento, por lo que también calculamos esta métrica en nuestro estudio y procedimos a comparar los resultados obtenidos.

Replicando idénticamente las condiciones de entrada y método de validación cruzada, comparamos los resultados en la Tabla 6.1:

| SUJETO | RBM + LSTM | 3 s |
|--------------|------------|--------|
| 1 | 0.7668 | 0.7924 |
| 2 | 0.7721 | 0.5072 |
| 3 | 0.8018 | 0.7669 |
| 4 | 0.6915 | 0.7292 |
| 5 | 0.7552 | 0.6554 |
| 6 | 0.8054 | 0.7804 |
| 7 | 0.7953 | 0.7714 |
| 8 | 0.7897 | 0.8347 |
| 9 | 0.7759 | 0.7461 |
| 10 | 0.7452 | 0.4189 |
| 11 | 0.8026 | 0.8538 |
| 12 | 0.7718 | 0.6169 |
| 13 | 0.7962 | 0.5971 |
| 14 | 0.7876 | 0.8145 |
| 15 | 0.8079 | 0.8491 |
| MEDIA | 0.7777 | 0.7155 |

Tabla 6.1: Tabla de comparación F1-score de estudio de referencia con modelo de clasificación binaria RBM+LSTM y nuestro estudio con configuración de longitud de entrada de las muestras de 3 segundos (3s).

Podemos notar una distinción notable entre ambos estudios. En el estudio de referencia, se observa una tendencia marcada donde todos los sujetos obtienen valores de F1-Score que oscilan entre 0.6915 y 0.8079, mientras que, en nuestro estudio, se presentan variaciones significativas entre sujetos. Por ejemplo, el sujeto 11 alcanza un F1-score considerablemente alto de 0.8538, pero en contraste, el sujeto 2 obtiene un valor bajo de 0.5072. Estas diferencias contribuyen a que la media en nuestro estudio se sitúe en 0.7155. Así que, nuestro modelo, utilizando las mismas características de entrada que el estudio de referencia, logra un rendimiento inferior.

Tras realizar algunos cambios en nuestro estudio, como la variación de la longitud de las muestras. En la tabla *Tabla 6.2* comparamos nuestro mejor resultado de la media de 10 segundos de segmentos de 1 segundo. Donde obtenemos una media de 0.7749 lo que se asemeja al penúltimo prueba del estudio de referencia.

Observamos que con la mejora de nuestro modelo conseguimos una media de F1-Score casi igual al estudio de referencia. Pero encontramos alguna discrepancia entre los dos modelos. Por ejemplo, en el artículo que referencia consiguen una clasificación muy equilibrada, todos los sujetos tienen unos valores que oscilan en una variación del 2%, mientras que nuestro modelo lleva a variar un 20% entre los sujetos que mejor y peor clasifica. Esto también lleva a que nuestro modelo consigue mucha mayor precisión en los sujetos que consigue clasificar bien, observamos valores bastante mayores en esos casos.

| SUJETO | RBM + LSTM | Media 10s |
|--------------|------------|-----------|
| 1 | 0.7668 | 0.7912 |
| 2 | 0.7721 | 0.5473 |
| 3 | 0.8018 | 0.8390 |
| 4 | 0.6915 | 0.6183 |
| 5 | 0.7552 | 0.7651 |
| 6 | 0.8054 | 0.7824 |
| 7 | 0.7953 | 0.7575 |
| 8 | 0.7897 | 0.8920 |
| 9 | 0.7759 | 0.8414 |
| 10 | 0.7452 | 0.6460 |
| 11 | 0.8026 | 0.9455 |
| 12 | 0.7718 | 0.5958 |
| 13 | 0.7962 | 0.8162 |
| 14 | 0.7876 | 0.8156 |
| 15 | 0.8079 | 0.9707 |
| MEDIA | 0.7777 | 0.7749 |

Tabla 6.2: Tabla de comparación F1-Score de estudio de referencia con modelo de clasificación binaria RBM+LSTM y nuestro estudio con configuración de longitud de entrada de las muestras de media de segmentos de 1 segundo en intervalos de 10 segundos

6.4. Limitaciones

Durante el desarrollo del Trabajo Fin de Grado (TFG), se han identificado diversas limitaciones que han influido en el proceso de investigación y en los resultados obtenidos. Estas limitaciones incluyen:

1. Base de Datos Limitada: Una de las limitaciones más notables ha sido la disponibilidad limitada de bases de datos públicas adecuadas para la clasificación de emociones mediante EEG. La elección de una base de datos de calidad es esencial para el éxito de la investigación, y la falta de opciones diversificadas podría haber restringido la variabilidad de los datos utilizados.
2. Número de Sujetos: El número de sujetos disponibles en las bases de datos utilizadas puede haber sido insuficiente para capturar la variabilidad interindividual en la respuesta emocional y la actividad cerebral. Una mayor cantidad de sujetos podría haber mejorado la generalización de los resultados.
3. Preprocesamiento de Datos: Aunque el preprocesamiento de datos es necesario para garantizar la calidad de estos, es una etapa delicada y sensible que puede afectar los resultados finales. La falta de información detallada sobre el proceso de preprocesamiento previo de los datos utilizados podría haber introducido incertidumbre en los resultados. Además, teníamos limitado el hecho de aplicar otro tipo de procesamientos que igual hubiera afectado menos a la clasificación.
4. Varianza Intersujeto: La variabilidad entre sujetos en la respuesta emocional y la actividad cerebral es un desafío conocido en esta área. Nuestro modelo ha tenido dificultades para generalizar de manera efectiva debido a estas diferencias interindividuales.
5. Arquitectura de la red: Utilizamos la red EEG-Inception con alguna variación. Podríamos haber realizado más modificaciones a los hiperparámetros de esta red, haber probado con otras longitudes de entrada o incluso haber creado una red específica para la clasificación de emociones.

Estas limitaciones son importantes para considerar al interpretar los resultados y las conclusiones del TFG. Sin embargo, también proporcionan oportunidades para futuras investigaciones, como la adquisición de bases de datos más diversificadas, el desarrollo de técnicas de preprocesamiento más avanzadas y la exploración de enfoques de modelado más sofisticados para abordar estos desafíos.

7. Capítulo VII

Conclusión

Índice

| | |
|---------------------------------|----|
| 7.1. Introducción | 65 |
| 7.2. Objetivos propuestos | 65 |
| 7.3. Aportaciones..... | 66 |
| 7.4. Conclusiones..... | 66 |
| 7.5. Líneas futuras..... | 67 |

7.1. Introducción

En este capítulo, profundizaremos en la evaluación de los resultados de nuestro trabajo de fin de grado, enfocándonos en los objetivos previamente establecidos, las conclusiones extraídas y las valiosas aportaciones logradas a lo largo de este proyecto. Además, consideraremos las posibles líneas de desarrollo futuro que se desprenden de este estudio, tomando en cuenta las limitaciones encontradas durante su ejecución.

7.2. Objetivos propuestos

Este Trabajo Fin de Grado se ha llevado a cabo de manera exitosa, cumpliendo cada uno de los objetivos específicos establecidos para el desarrollo de un clasificador de emociones basado en la actividad eléctrica cerebral mediante técnicas de DL. A continuación, destacamos los logros alcanzados en relación con los objetivos mencionados:

1. Revisión Bibliográfica: Se ha realizado una exhaustiva revisión de la literatura científica relacionada con el reconocimiento de emociones utilizando EEG y técnicas de DL. Esta revisión proporcionó la base teórica y conceptual necesaria para abordar el problema de manera informada y eficaz.
2. Recopilación de Bases de Datos: Se llevaron a cabo labores de investigación y recopilación de bases de datos públicas relacionadas con la actividad cerebral y las emociones. Esta recopilación permitió disponer de conjuntos de datos adecuados para entrenar y evaluar el modelo propuesto, como es la base de datos SEED.
3. Preprocesamiento de Datos: Se aplicaron técnicas de preprocesamiento a los datos para garantizar su calidad y utilidad en el desarrollo del modelo. Esto incluyó la normalización de los datos y la división en épocas, entre otros procesos esenciales.
4. Desarrollo del Modelo DL: Se propuso un modelo de DL adecuado para el reconocimiento de emociones a partir de la actividad cerebral. Se seleccionó una arquitectura de red, se definieron las capas y se configuraron los hiperparámetros de manera óptima.
5. Evaluación del Rendimiento: Se evaluó exhaustivamente el rendimiento del modelo propuesto utilizando métricas apropiadas, como la precisión o las matrices de confusión. Los resultados obtenidos se compararon con trabajos previos y se discutieron las ventajas y limitaciones del enfoque desarrollado.

En resumen, este trabajo ha contribuido al campo del reconocimiento de emociones a partir de la actividad eléctrica cerebral, proporcionando una herramienta sólida basada en DL. Los resultados obtenidos respaldan la efectividad de esta aproximación y abren la puerta a futuras investigaciones y aplicaciones en este emocionante campo interdisciplinario.

7.3. Aportaciones

Las aportaciones obtenidas tras haber realizado este TFG y haber analizado cada una de sus partes, se muestran a continuación.

En primer lugar, se ha demostrado la viabilidad de utilizar técnicas de DL en combinación con datos EEG para la clasificación de emociones. Los modelos desarrollados en este estudio han arrojado resultados prometedores en la identificación de estados emocionales basados en los patrones de actividad cerebral.

La elección de la arquitectura de red neuronal se ha revelado como un factor crucial en el éxito del modelo, destacando la influencia de la selección adecuada de capas y hiperparámetros en el rendimiento general del sistema. Además, se ha modificado una herramienta de clasificación de registros de EEG basada en DL para utilizarla en clasificación de emociones, que podría servir como punto de partida para investigaciones futuras en neurociencia emocional.

Este TFG también ha contribuido al análisis de datos EEG mediante la implementación de técnicas de preprocesamiento y exploración de datos. Estas estrategias han ampliado nuestra comprensión de la interacción entre los patrones de actividad cerebral y estados emocionales particulares.

En resumen, este TFG no solo ha generado conocimientos valiosos sobre la clasificación de emociones a partir de EEG, sino que también ha sentado las bases para futuras investigaciones en este campo multidisciplinario. Las aportaciones obtenidas contribuyen al avance en la comprensión de las complejas interacciones entre el cerebro humano y las emociones.

7.4. Conclusiones

Tras la ejecución del TFG centrado en el reconocimiento de emociones a partir de la actividad eléctrica cerebral (EEG), se han extraído las siguientes conclusiones:

1. La variabilidad emocional entre individuos y su dependencia del estado en tiempo real del sujeto representan desafíos significativos en la clasificación de emociones a partir de registros de EEG. Esta dificultad en la generalización se debe a las diferentes formas en que cada persona puede expresar sus emociones a nivel cerebral.
2. El uso de EEG-Inception ha demostrado ofrecer una alta capacidad de detección y clasificación de patrones asociados a las emociones en los registros de EEG, especialmente en el contexto de la clasificación binaria.
3. Disponer de una base de datos con relevancia científica y validar rigurosamente los datos para lograr resultados precisos y fiables.
4. Importancia del análisis de datos de EEG, aplicación de técnicas de procesamiento y exploración de datos que proporcionan una comprensión más profunda de cómo los patrones de actividad cerebral se relacionan con estados emocionales específicos.

7.5. Líneas futuras

A pesar de las limitaciones encontradas durante el desarrollo del TFG, existen diversas oportunidades de continuación y aplicaciones para la herramienta desarrollada:

1. Ampliación de la Base de Datos: La adquisición y ampliación de una base de datos específica y diversa podría mejorar la robustez y generalización del modelo. Obtener datos de más sujetos y situaciones emocionales podría ser beneficioso.
2. Exploración de Datos Crudos: Trabajar directamente con datos EEG sin preprocesar podría ser una dirección futura. Esto podría incluir el desarrollo de técnicas de preprocesamiento más avanzadas que minimicen el impacto de artefactos.
3. Transferencia de Aprendizaje: Utilizar modelos preentrenados en datos de EEG de amplia disponibilidad para el reconocimiento de características generales antes de adaptarlos a tareas específicas de reconocimiento de emociones.
4. Aplicaciones Clínicas: La herramienta podría tener aplicaciones en la detección temprana de trastornos emocionales o neuropsiquiátricos, así como en la mejora de la interacción hombre-máquina en contextos de atención médica.

Además, para mejorar la exactitud de la herramienta, podrían considerarse las siguientes mejoras:

5. Validación de Base de Datos: Realizar una verificación más exhaustiva de la calidad y consistencia de la base de datos utilizada, identificando y abordando posibles anomalías.
6. Exploración de Modelos Avanzados: Probar modelos de DL más avanzados, como redes neuronales convolucionales 3D (CNN 3D), que puedan capturar mejor las relaciones temporales en los datos EEG.
7. Ensemble de Modelos: Implementar un ensemble de múltiples modelos puede aumentar la robustez y precisión del sistema.

En resumen, a pesar de las limitaciones, este TFG ha sentado las bases para futuras investigaciones y aplicaciones en el campo del reconocimiento de emociones a partir de EEG. Las mejoras propuestas y la exploración de enfoques más avanzados podrían conducir a herramientas más precisas y aplicaciones más efectivas en el futuro.

Bibliografía

- [1] “Qué es la ingeniería biomédica + Plan de estudios | Areandina.” <https://www.areandina.edu.co/blogs/que-es-ingenieria-biomedica> (accessed Sep. 17, 2023).
- [2] J. D. Bronzino, Ed., *Medical Devices and Systems*. CRC Press, 2006. doi: 10.1201/9781420003864.
- [3] M. X. Cohen, “Where Does EEG Come From and What Does It Mean?,” *Trends Neurosci*, vol. 40, no. 4, pp. 208–218, Apr. 2017, doi: 10.1016/j.tins.2017.02.004.
- [4] J. Poza Crespo, “Análisis tiempo-frecuencia de la actividad magnetoencefalográfica espontánea en la enfermedad de Alzheimer,” Universidad de Valladolid, 2008. doi: 10.35376/10324/136.
- [5] Y.-S. Lee and W.-H. Park, “Diagnosis of Depressive Disorder Model on Facial Expression Based on Fast R-CNN,” *Diagnostics*, vol. 12, no. 2, p. 317, Jan. 2022, doi: 10.3390/diagnostics12020317.
- [6] C. I. Hernández-Vargas and M. E. Dickinson-Bannack, “Importancia de la inteligencia emocional en Medicina,” *Investigación en Educación Médica*, vol. 3, no. 11, pp. 155–160, Jul. 2014, doi: 10.1016/S2007-5057(14)72742-5.
- [7] E. María, M. Guerrero, S. A. Cruces, and Á. Catedrático, “Análisis e identificación de potenciales evocados en el electroencefalograma.”
- [8] D. Dadebayev, W. W. Goh, and E. X. Tan, “EEG-based emotion recognition: Review of commercial EEG devices and machine learning techniques,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4385–4401, Jul. 2022, doi: 10.1016/j.jksuci.2021.03.009.
- [9] A. Cristian, F. Blanco Díaz, A. F. Ruiz, O. Profesor, and A. Nariño, “Estudio comparativo de métodos para el reconocimiento frecuencial de señales de EEG relacionadas a potenciales evocados visuales en estado estacionario.”
- [10] R. C. Deo, “Machine learning in medicine,” *Circulation*, vol. 132, no. 20, pp. 1920–1930, Nov. 2015, doi: 10.1161/CIRCULATIONAHA.115.001593.
- [11] A. L. Samuel, “Some Studies in Machine Learning Using the Game of Checkers,” *IBM J Res Dev*, vol. 3, no. 3, pp. 210–229, Jul. 1959, doi: 10.1147/rd.33.0210.
- [12] T. M. (Tom M. Mitchell, *Machine Learning*.
- [13] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction,” *BMC Med Res Methodol*, vol. 19, no. 1, Mar. 2019, doi: 10.1186/S12874-019-0681-4.
- [14] A. Núñez Reiz, M. A. Armengol de la Hoz, and M. Sánchez García, “Big Data Analysis y Machine Learning en medicina intensiva,” *Med Intensiva*, vol. 43, no. 7, pp. 416–426, Oct. 2019, doi: 10.1016/j.medin.2018.10.007.
- [15] P. P. Shinde and S. Shah, “A Review of Machine Learning and Deep Learning Applications,” in *2018 Fourth International Conference on Computing Communication*

- Control and Automation (ICCUBEA)*, IEEE, Aug. 2018, pp. 1–6. doi: 10.1109/ICCUBEA.2018.8697857.
- [16] “A.I. technical - Machine Learning vs. Deep Learning.” <https://lawtomated.com/a-i-technical-machine-vs-deep-learning/> (accessed Jul. 15, 2023).
- [17] X. Li *et al.*, “EEG Based Emotion Recognition: A Tutorial and Review,” *ACM Comput Surv*, vol. 55, no. 4, pp. 1–57, Apr. 2023, doi: 10.1145/3524499.
- [18] A.M.L. Coenen, “Pioneros de la electroencefalografía (EEG),” *Brainclinics insights*, 2022. <https://brainclinics.com/pioneros-del-eeg/> (accessed Sep. 05, 2023).
- [19] M. Balcells and V. Cisteré, “Imagen de archivo Historia de la Electroencefalografía en España: introducción y evolución,” 2014.
- [20] Rafael Barea Navarro, “Electroencefalografía,” *Universidad de Alcalá*.
- [21] “Historia y avances de la Electroencefalografía (EEG),” *Instituto Valenciano de Neurología Pediátrica (INVANEP)*, 2018. https://invanep.com/blog_invanep/historia-y-avances-de-la-electroencefalografia-eeg (accessed Sep. 05, 2023).
- [22] J. R. , Wolpaw, *Brain-computer interfaces : principles and practice*. New York: Oxford University Press, 2012.
- [23] Javier Gómez Pilar, “Apuntes Neurociencia Computacional. 4º Ingeniería Biomédica UVa.”
- [24] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain*. Oxford University Press, 2006. doi: 10.1093/acprof:oso/9780195050387.001.0001.
- [25] Adrian Eleazar Contreras Martínez, Dr. Jorge Luis Rosas Trigueros, and Dra. Blanca Esther Carvajal Gámez, “Sistema clasificador de señales EEG para el control de una neuroprótesis por medio de un dispositivo móvil,” INSTITUTO POLITECNICO NACIONAL ESCUELA SUPERIOR DE COMPUTO. Accessed: Jul. 07, 2023. [Online]. Available: https://www.escom.ipn.mx/posgrado/tesis/Tesis_064_Adrian_Eleazar_Contreras_Martinez.pdf
- [26] Carlos Javier Bless Sastre and Carlos Gómez Peña, “Análisis de la actividad EEG durante una sesión de estimulación multisensorial en una sala Snoezelen,” Universidad de Valladolid, Valladolid. Accessed: Jul. 07, 2023. [Online]. Available: <https://uvadoc.uva.es/bitstream/handle/10324/42498/TFG-G4147.pdf?sequence=1&isAllowed=y>
- [27] “El Observatorio FAROS Sant Joan de Déu (www.faroshjd.net) es la plataforma de promoción de la salud y el bienestar infantil del Hos-pital Sant Joan de Déu (HSJD) de Barcelona”, Accessed: Jul. 09, 2023. [Online]. Available: <http://www.facebook.com/ObservatorioFAROS>
- [28] H. Lövheim, “A new three-dimensional model for emotions and monoamine neurotransmitters,” 2nd ed. Medical hypotheses, 2012, pp. 341–348.
- [29] R. Plutchick, “The Nature of Emotions,” 4th ed. American Scientist, 2001, pp. 344–350. Accessed: Jul. 09, 2023. [Online]. Available: https://www.academia.edu/43620307/The_Nature_of_Emotions_Plutchik_2001_

- [30] J. A. Russel, “A circumplex model of affect,” 6th ed. *Journal of Personality and Social Psychology*, 1980, pp. 1161–1178.
- [31] P. Duque, J. M. Quintillá, J. A. Varela, P. Garrido, Ó. Valencia, and J. M. Maestre, “Respuesta emocional a una simulación de alta fidelidad en trauma: Un estudio observacional,” *Rev Esp Anestesiol Reanim*, vol. 66, no. 10, pp. 521–527, Dec. 2019, doi: 10.1016/j.redar.2019.07.006.
- [32] “Apuntes Fisiología 2º Ingeniería Biomédica ,” *Universidad de Valladolid*. 2021.
- [33] L. J. M. Rothkrantz, R. Horlings, and D. Dacu, “Emotion recognition using brain activity,” in *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, CompSysTech '08*, 2008. doi: 10.1145/1500879.1500888.
- [34] Dña. Raquel Martínez Rodríguez, “DISEÑO DE UN SISTEMA DE DETECCIÓN Y CLASIFICACIÓN DE CAMBIOS EMOCIONALES BASADO EN EL ANÁLISIS DE SEÑALES FISIOLÓGICAS NO INTRUSIVAS,” *Escuela Universitaria de Ingeniería Técnica de Bilbao*, Bilbao, 2016.
- [35] “SEED Dataset.” <https://bcmi.sjtu.edu.cn/home/seed/seed.html> (accessed Jul. 07, 2023).
- [36] Wei-Long Zheng and Bao-Liang Lu, “Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks,” *IEEE Trans Auton Ment Dev*, vol. 7, no. 3, pp. 162–175, Sep. 2015, doi: 10.1109/TAMD.2015.2431497.
- [37] G. E. Ingeniería De Tecnologías De, D. Eduardo Santamaría Vázquez, and D. D. Roberto Hornero Sánchez Víctor Martínez Cagigal, “UNIVERSIDAD DE VALLADOLID ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN TRABAJO FIN DE Máster BRAIN COMPUTER INTERFACE (BCI) ORIENTADA A PERSONAS CON GRAVE DISCAPACIDAD.”
- [38] E. Santamaria-Vazquez, V. Martinez-Cagigal, F. Vaquerizo-Villar, and R. Hornero, “EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-Based Brain-Computer Interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2773–2782, Dec. 2020, doi: 10.1109/TNSRE.2020.3048106.
- [39] X. Li *et al.*, “Latent Factor Decoding of Multi-Channel EEG for Emotion Recognition Through Autoencoder-Like Neural Networks,” *Front Neurosci*, vol. 14, Mar. 2020, doi: 10.3389/fnins.2020.00087.