



Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS DE LA COMPUTACIÓN
DOCTORADO EN INGENIERÍA DEL LENGUAJE Y DEL
CONOCIMIENTO

**“Caracterización de señales electroencefalográficas para la
identificación de conceptos asociados al proceso de
comunicación humana”**

T E S I S

Presentada para obtener el grado de:

Doctor en ingeniería del lenguaje y del conocimiento

Presenta:

M. C. Rigoberto Cerino Jiménez

Director:

Dr. Sergio Vergara Limon

Co-director:

Dr. David Eduardo Pinto Avendaño

Puebla, México

NOVIEMBRE 2022

AGRADECIMIENTOS

Al ver culminado un proyecto con tantas ambiciones, con tantos alcances y emociones, únicamente me resta dar gracias.

Gracias porque el proceso para culminar y plasmar una idea en un trabajo tangible, requiere la participación y apoyo de más de una persona, y en este caso, conté con ellas. El proceso de finalizar un doctorado es una etapa compleja en múltiples sentidos, pero el más notorio es en el estado de ánimo, en esas fortalezas que los buenos comentarios y las buenas personas plasman en nuestro ser, ese apoyo incondicional y esa confianza puesta en mí, y esas ocasiones que por la mente pasa el titubear sobre lo que haces, pero siempre tuve a alguien dispuesto a escuchar, a llenar de fortaleza y a motivar. En consecuencia, me es satisfactorio expresar mis agradecimientos a las personas e instituciones que hicieron posible el logro de este trabajo. Agradezco en primera instancia a Dios por permitirme la vida y salud, así como la fortaleza para culminar una etapa más a lo largo de mi vida personal y profesional, por mantener a mi familia unida y permitirme esa dicha de disfrutarla.

A mi familia que siempre ha estado en completa disposición de apoyarme en todos los sentidos, en las decisiones realizadas, en esos momentos que únicamente los lazos familiares pueden reparar cualquier daño y llenar cualquier vacío. Por los consejos otorgados, por la educación y valores que me dan a diario, que sin ellos no podría estar en este punto ni en este lugar. Ese amor cálido de familia que me llena de seguridad, esa formación que me ha permitido avanzar en esta vida en múltiples ámbitos, trazarme nuevas metas y cumplirás, eternamente gracias a ellos.

A mis directores de tesis, al Dr. Sergio Vergara Limon y al Dr. David Eduardo Pinto Avendaño, quienes depositaron su confianza en mí, en mis habilidades, y que a pesar de las múltiples equivocaciones siempre estuvieron para apoyar, para guiar, para corregir, para enseñar. Gracias por permitirme trabajar en conjunto, por los consejos, por las correcciones y por el conocimiento adquirido, pero sobre todo gracias por la motivación y el ejemplo.

Agradezco especialmente a mis compañeros de doctorado, a José de Jesús Moya Mora y Francisco José López Cortés, con quienes compartí múltiples experiencias, y que a pesar de lo complicado que resulta para cada uno el desarrollar un tema pos doctoral, siempre tuvieron la gentileza y el tiempo de escuchar, apoyar, y esa solidaridad en aspectos personales así como laborales. Este proyecto no hubiera sido posible sin ellos.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico para cubrir los gastos necesarios de manutención en este lapso.

Finalmente agradezco a mi alma mater, la Benemérita Universidad Autónoma de Puebla por permitirme crecer personal y profesionalmente, así como al Doctorado en Ingeniería del Lenguaje y del Conocimiento, el cual me permitió obtener un grado más, adquiriendo conocimientos y nuevos retos día a día, experiencias únicas que me han formado como profesionista, así como por todo el apoyo otorgado hacia el proyecto, que gracias a éste fue posible concluirlo.

DEDICATORIA

A mis padres

Dedico esta tesis a Efraín Cerino López y Esperanza de Jesús Jiménez Cantoral, quienes siempre han estado para mí, permitiéndome sus consejos, su confianza, su amor y su apoyo ante cualquier circunstancia, preparándome para la vida, para los retos, para sobresalir, que sin su guía no podría haber crecido como persona, que sin su amor incondicional no estaría donde me encuentro ahora, quienes entregan todo sin pedir nada a cambio y me permiten crecer con humildad.

A mis hermanos

Erick Cerino Jiménez y Eduardo Efraín Cerino Jiménez, quienes han sido un ejemplo y un complemento de vida, por el amor incomparable y las vivencias adquiridas, por ese crecimiento en todos los sentidos de los que han sido partícipes, y que sin su compañía y apoyo no me permitiría vivir con plenitud.

ÍNDICE

1. Introducción.....	12
1.1. Antecedentes.....	13
1.1.1. La discapacidad en México	16
1.2. Justificación	17
2. Marco epistémico.....	19
2.1. Problema de investigación.....	19
2.2. Objetivos de la tesis.....	20
2.2.1. Objetivo general.....	20
2.2.2. Objetivos específicos	20
2.3. Preguntas de investigación	20
2.4. Límites de la tesis	21
3. Marco teórico.....	22
3.1. El cerebro.....	22
3.2. Ondas cerebrales.....	26
3.2.1. Origen fisiológico de las señales.....	26
3.2.2. Tipos de ondas.....	27
3.3. Adquisición de señales cerebrales.....	30
3.4. Electroencefalografía (EEG)	31
3.4.1. Ventajas e inconvenientes de la Electroencefalografía	33
3.4.2. Aplicaciones.....	33
3.5. Sistema Internacional 10-20	34
3.6. Potenciales relacionados con eventos.....	34
3.6.1. Ondas P3 o P300	37
3.7. Interfaz cerebro-computadora (BCI)	39
3.7.1. Elementos de una BCI	40
3.7.2. Tipos de interfaces BCI	41
3.8. Sistemas aumentativos y alternativos de comunicación (saac).....	43
3.9. Los sistemas de símbolos.....	44
3.9.1. Los sistemas pictográficos de comunicación.....	44
3.10. Preprocesamiento de las señales EEG.....	45
3.10.1. Artefactos EEG	45
3.10.2. Filtrado de las señales EEG	48

3.11.	Extracción de características de señales EEG	49
3.12.	Clasificación de las señales EEG.....	55
4.	Estado del arte	61
4.1.	Historia y desarrollo del campo	61
4.2.	La representación de conceptos de objeto en el cerebro.....	63
4.3.	Identificación de categorías de objetos a partir de EEG relacionado con eventos	71
4.3.	Identificación de categorías de objetos a partir de fMRI	77
4.4.	Un enfoque de aprendizaje profundo de extremo a extremo para la clasificación de señales EEG para BCI.....	79
5.	Diseño metodológico	87
5.1.	Entrenamiento	87
5.2.	Identificación automática de conceptos.....	89
5.3.	Población.....	91
5.4.	Hardware	91
5.5.	Software	91
6.	Resultados	95
6.1.	La lengua española.....	95
6.2.	Lenguaje Toki Pona	96
6.3.	Diseño del experimento EEG	105
6.4.	Dispositivo para la adquisición de las señales EEG	109
6.5.	Base de datos pública de señales EEG.....	110
6.6.	Preprocesamiento de las señales EEG.....	111
6.6.1.	Implementación de filtros.....	112
6.6.2.	Eliminación de artefactos	116
6.7.	Extracción de características.....	125
6.8.	Selección de clasificador de señales EEG	129
6.9.	Redes neuronales convolucionales (CNN).....	133
6.9.1.	Preparación de los datos para la entrada a la CNN.....	135
6.9.2.	Arquitectura del modelo.....	136
6.9.3.	Evaluación del modelo	138
6.9.4.	Overfitting o sobreajuste.....	139
6.10.	Pruebas de laboratorio	148
7.	Conclusiones.....	152
	Referencias	155

ÍNDICE DE FIGURAS

Figura 3.1. a) Áreas cerebrales según principios frenológicos del siglo XIX. b) Áreas funcionales cerebrales según conceptos neurológicos actuales (Vargas, 2012).....	23
Figura 3.2. Especialización de los hemisferios cerebrales. En general las funciones cerebrales están más deslocalizadas de lo que se creía, pero hay algunas funciones que se realizan con más intensidad en una mitad que en otra (Argopsique, 2014), (Vargas, 2012).....	24
Figura 3.3. La corteza cerebral marcada muestra diferentes áreas funcionales (AAMODH, 2016).	25
Figura 3.4. Clasificación de ondas cerebrales respecto a su frecuencia de onda: Gamma, alfa, beta, theta y delta (Damasio, 2010).....	28
Figura 3.5. Relaciones comunes entre subcampos EEG. Las aplicaciones clínicas están relacionadas principalmente con enfermedades neurológicas. La investigación de EEG es llevada a cabo por neurólogos, neurocientíficos cognitivos, físicos e ingenieros que tienen un interés especial en EEG (Nunez, 2017).	32
Figura 3.6. Sistema internacional 10-20 de colocación de electrodos (Oñate, 2018).	35
Figura 3.7. Forma de onda del ERP auditivo promediado por computadora a un breve sonido. La visualización del tiempo logarítmico permite la visualización de las respuestas tempranas del tronco encefálico (ondas I – VI), componentes evocados intermedios y componentes endógenos de latencia larga (Nd, N2, P300 y SW, u onda lenta). Los componentes de ERP se etiquetan de acuerdo con su polaridad, P (positiva) o N (negativa), seguidos de un número o letra (P3, Pa, etc.) que especifica la posición ordinal de la onda en una secuencia particular o por un número más grande (P300, etc.) que se refiere a la latencia de la onda en milisegundo (Hillyard, 2017).	36
Figura 3.8. Comparación de la forma de onda típica del ERP visual con un estímulo transitorio (arriba) con el potencial evocado visual en estado estacionario (SSVEP) provocado por un estímulo parpadeante repetitivo (abajo). En este ejemplo, el SSVEP es provocado por un estímulo visual que parpadea a 16 Hz, y el SSVEP aparece como una respuesta sinusoidal de la corteza visual a la misma frecuencia. Siguiendo una señal para atender el estímulo parpadeante, se ve que el SSVEP crece en amplitud, y el curso del tiempo de este crecimiento proporciona una estimación de la velocidad a la que se dedica la atención (Hillyard, 2017)	37
Figura 3.9. La onda P3 (- - no objetivo, - objetivo) (Fabien L. , 2014), (Apdullah, 2017).....	38
Figura 3.10. Electrodos comúnmente usados para detectar los potenciales P300 (Fabien L. , 2014).	39
Figura 3.11. (A) Ejemplo de matriz de electrodos utilizada para ECoG y (B) su disposición sobre el cerebro (Schalk & Leuthardt, 2011)	42
Figura 3.12. Posicionamiento de los electrodos sobre el córtex en la Electrocorticografía.	42
Figura 3.13. El casco Emotiv EPOC+ es un ejemplo de instrumento no invasivo para tomar datos (Emotiv, 2013).	43
Figura 3.14. Ejemplos de formas de onda de prueba única que muestran artefactos comunes. Cada artefacto tiene un conjunto interactivo de propiedades y, por lo tanto, puede detectarse de forma precisa con los algoritmos reales que se ajustan a estas propiedades (Delorme, EEGLAB, 2018), (Lopez & J.Luck, 2014).	46
Figura 3.15. Proceso típico para extraer características de un canal de datos EEG para un diseño BCI basado en P300. En esta imagen podemos ver que el P300 se vuelve más visible con los diferentes pasos de procesamiento (Fabien L. , 2014).	52

Figura 3.16. El SVM encuentra el hiperplano óptimo para generalización. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).....	57
Figura 3.17. El SVM encuentra el hiperplano óptimo para generalización. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).....	57
Figura 3.18. El SVM encuentra el hiperplano óptimo para generalización. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).....	57
Figura 3.19. El SVM encuentra el hiperplano óptimo para generalización. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).....	57
Figura 3.20. El algoritmo k vecino más cercano comienza en los datos de prueba y aumenta la región circular hasta que cubre el número K de datos de entrenamiento. Los datos de prueba se etiquetan por mayoría de votos. Fuente: (Apdullah, 2017).....	58
Figura 3.21. El algoritmo k vecino más cercano comienza en los datos de prueba y aumenta la región circular hasta que cubre el número K de datos de entrenamiento. Los datos de prueba se etiquetan por mayoría de votos. Fuente: (Apdullah, 2017).....	58
Figura 3.22. El algoritmo k vecino más cercano comienza en los datos de prueba y aumenta la región circular hasta que cubre el número K de datos de entrenamiento. Los datos de prueba se etiquetan por mayoría de votos. Fuente: (Apdullah, 2017).....	58
Figura 3.23. El algoritmo k vecino más cercano comienza en los datos de prueba y aumenta la región circular hasta que cubre el número K de datos de entrenamiento. Los datos de prueba se etiquetan por mayoría de votos. Fuente: (Apdullah, 2017).....	58
Figura 3.24. Un hiperplano que separa dos clases: los círculos y las cruces. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).....	59
Figura 4.1. Vista lateral esquemática del hemisferio izquierdo (A) y vista ventral del lóbulo frontal y temporal derecho (B). El giro fusiforme se muestra con mayor detalle en (C). Las regiones en rojo muestran la ubicación aproximada de áreas típicamente involucradas en tareas de procesamiento conceptual, especialmente con objetos concretos. ITG, giro temporal inferior; LG, circunvolución lingual. Ver texto para otras abreviaturas. Fuente: (Martin A. , 2007).....	66
Figura 4.2. Las regiones cerebrales se activan más consistentemente durante el procesamiento de las características emocionales en los estudios de procesamiento de palabras y oraciones. Los círculos rojos denotan áreas que se han asociado principalmente con el procesamiento afectivo, mientras que los círculos azules indican regiones que sirven principalmente a procesos relacionados con el lenguaje. Abreviaturas: ACC, corteza cingulada anterior; PCC, corteza cingulada posterior; OFC, corteza prefrontal orbital; DLPFC, corteza prefrontal dorsolateral; mPFC, corteza prefrontal medial; IC, corteza insular; SPL, lóbulo parietal superior; CE, corteza extra-estriada; Am, amígdala; BG, ganglios basales; PC, corteza premotora; MTC, corteza temporal media; IFG, giro frontal inferior; FG, giro fusiforme; MTG, giro temporal medio; LTP, polo temporal izquierdo, ITG, giro temporal inferior; ILFC, corteza frontal inferior izquierda; MLFC, corteza frontal media izquierda; ASTG, giro temporal superior anterior (Hinojosa, et al., 2019).	67
Figura 4.3. (A) Ubicación del área cortical de la corriente ventral en el cerebro del mono macaco y flujo de información visual desde la retina. (B) Cada área se traza de manera que su tamaño sea proporcional a su área de superficie cortical. El número total aproximado de neuronas (ambos hemisferios) se muestra en la esquina de cada área (M = millones). La dimensionalidad aproximada de cada representación (número de neuronas de proyección) se muestra encima de cada área, según las densidades neuronales, la fracción neuronal de capa 2/3 y la porción (color)	

dedicado al procesamiento de los 10 grados centrales del campo visual. La latencia de respuesta media aproximada se enumera a la derecha (James, Davide, & Nicole, 2012).	68
Figura 4.4. Modelo conceptual de la corteza visual artificial. La imagen en color se descompone en cuatro dimensiones (color, orientación, forma e intensidad). Luego, una estructura jerárquica se encarga de resolver el problema de clasificación de objetos a través de un paradigma dirigido por funciones (Olague, et al., 2019).	69
Figura 4.5. Clasificación del desempeño para las tres modalidades. Las barras oscuras indican resultados significativos (Simanova, Gerven, Oostenveld, & Hagoort, 2010).	72
Figura 4.6. Diseño de presentación visual en serie rápida. Se presenta la secuencia de los objetivos visuales que observaran las personas para el proceso de identificación de conceptos (Grootswagers, Zhou, Robinson, Hebart, & Carlson, 2022).	76
Figura 4.7. Paradigma experimental e imágenes de los estímulos. Se presentan a los participantes imágenes de objetos sobre un fondo natural utilizando un paradigma RSVP.....	77
Figura 4.8. Representación esquemática del tiempo de presentación de cada concepto (Shinkareva, et al., 2008).	78
Figura 4.9. Marco de aprendizaje profundo bidireccional a nivel de región para la clasificación de objetos visuales. El enfoque consta de tres etapas, incluida la etapa de extracción de información a nivel de región, la etapa de codificación de características y la etapa de clasificación.	82
Figura 5.1. Diagrama a bloques de la metodología propuesta para la solución del problema de entrenamiento del sistema.	89
Figura 5.2. Diagrama a bloques de la metodología propuesta para la creación de un sistema de identificación automática de conceptos.	90
Figura 6.1. Subclases de taxonomía. b) Palabras ambiguas contenidas en el corpus de pictogramas	98
Figura 6.2. a) Clasificación del total de pictogramas en el corpus. b) Clasificación de pictogramas sin repetir localizados en el corpus (OLIVARES, 2019).	98
Figura 6.3. Formulario realizado y enviado en línea a un público de más de 160 personas, para obtener respuestas de lo que significa para ellos cada pictograma.	99
Figura 6.4. Ejemplos de pictogramas presentados a la comunidad, con el objetivo de saber que es lo primero que llega a su mente cuando observan los pictogramas.	100
Figura 6.5. Relación de las necesidades básicas según Virginia Henderson.	103
Figura 6.6. Alteración en el porcentaje de clasificación de señales EEG basado en el número de presentaciones repetidas, el número de sujetos y la cantidad de datos de entrenamiento.	107
Figura 6.7. Diseño del experimento para la presentación de estímulos objetivos al usuario.	108
Figura 6.8. Localización de los electrodos en el kit EMOTIV (Alejandro, 2016)	109
Figura 6.9. Estímulo visual presentado ante los usuarios (Kaneshiro, M, H-S, AM, & P, 2015).....	111
Figura 6.10. Señales EEG representadas en EEGLAB, generadas por los estímulos visuales presentados al usuario, donde las líneas verticales representan el momento en que el usuario concluyó de ver el estímulo.	113
Figura 6.11. Diagrama de bloques del sistema de identificación de la estructura del filtro.....	114
Figura 6.12. Densidad espectral de potencia de las señales EEG de todos los canales de la diadema utilizada.	114
Figura 6.13. Densidad espectral de potencia. Se presentan las señales después de aplicar el FIR diseñado para eliminar el desplazamiento de DC y el ruido de la señal de la red eléctrica.	116

Figura 6.14. Diseño de filtro pasa bandas FIR para la eliminación de ruido indeseado.	116
Figura 6.15. Porcentajes de tipos de algoritmos de eliminación de artefactos. El gráfico de la izquierda muestra la división de ciertas técnicas a las que se hace referencia en (Urigüen & Garcia-Zapirain, 2015). El gráfico de la derecha muestra la división de los algoritmos ICA.....	117
Figura 6.16. Método ICA para desacoplar señales EEG y trabajar sobre las fuentes originales. (Delorme, EEGLAB, 2020)	118
Figura 6.17. Comparación entre ICA y PCA, donde ICA intenta hacer las salidas estadísticamente, sin colocar restricciones en la matriz W (Delorme, EEGLAB, 2020).	119
Figura 6.18. Detección automática de artefactos implementando toolbox EEGLAB en MATLAB..	120
Figura 6.19. Detección automática de artefactos implementando toolbox EEGLAB en MATLAB..	121
Figura 6.20. Señales EEG reconstruidas después de aplicar el método ICA para detección de artefactos y la eliminación de estos y segmentación de la señal en épocas.	121
Figura 6.21. Análisis de la actividad cerebral en cada uno de los electrodos de la diadema con la que fueron capturados los datos, para observar a través del procesamiento ICA, el comportamiento de dicha actividad.	121
Figura 6.22. Espectro de potencia de la actividad cerebral en el electrodo 21.	122
Figura 6.23. Análisis del espectro de potencia de la actividad cerebral del electrodo 21, graficado en el tiempo.	122
Figura 6.24. Intensidad de actividad en las distintas regiones cerebrales a lo largo del tiempo. ..	123
Figura 6.25. Señal graficada de un solo canal en el dominio del tiempo de la base de datos, en las diferentes bandas clásicas de comportamiento de la señal EEG.	124
Figura 6.26. Potencia de las señales EEG, donde se observa la mayor amplitud de esta, alrededor de los 400 ms, que representa la mayor actividad cerebral.	124
Figura 6.27. Análisis espectral mediante el método de Welch de un solo canal de la base de datos, en las diferentes bandas clásicas de comportamiento de la señal EEG.....	125
Figura 6.28 Métrica de exactitud y matriz de confusión de la clasificación utilizando el método de SVM.	131
Figura 6.29. a) Localización espacial de los electrodos implementados en la adquisición de las señales EEG. b) Exactitud individual de los electrodos.	132
Figura 6.30. Métrica de exactitud y matriz de confusión de la clasificación utilizando el método de LDA.	132
Figura 6.31. Gráfica de exactitud en la clasificación a nivel categoría.....	133
Figura 6.32. Red convolucional: Capas convolucionales (azul), Pooling (verde), y Fully-connected (rojo). Fuente: (Dagallier & Campillos, 2018).....	134
Figura 6.33. Arquitectura general: Las capas de la CNN se pueden describir como diferentes tipos de pasos de filtrado y procesamiento antes de la clasificación real. Los rectángulos rojos indican el núcleo del filtro y las dimensiones de la agrupación. Fuente: (Dosea, Møllera, Iversenb, & Puthusserypady, 2018).....	135
Figura 6.34. Representación de las convoluciones aplicadas a la matriz de entrada a la CNN, donde el eje “y” representa el número de canales, y el eje “x” representa los puntos de tiempo de la ventana. Fuente: (Schirrmeister R. T., et al., 2017).	136
Figura 6.35. Arquitectura del modelo original de EEGNet.	137
Figura 6.36. Arquitectura EEGNet adaptada a los parámetros previamente definidos de la base de datos.....	137

Figura 6.37. Separación del conjunto de datos para un flujo de trabajo eficaz. Fuente: (Developers, 2021)	139
Figura 6.38. Gráficas con señal de sobreajuste: la pérdida de entrenamiento está disminuyendo, pero la pérdida de validación está aumentando. Fuente: (Schluter, 2019).....	140
Figura 6.39. Gráfica de exactitud y pérdida implementando una red EEGNet.....	140
Figura 6.40. a) Red neuronal estándar. b) Red neuronal después de aplicar “Dropout”. Fuente: (Pelayo & Gordillo, 2018).	141
Figura 6.41. Gráficas de exactitud y pérdida de la arquitectura EEGNet.....	142
Figura 6.42. Gráficas de exactitud y pérdida de la arquitectura DeepConvNet.	143
Figura 6.43. Gráficas de exactitud y pérdida de la arquitectura ShallowConvNet.	143
Figura 6.44. a) Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: EEGNet.b) Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: DeepConvNet.	144
Figura 6.45. Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: ShallowConvNet.....	144
Figura 6.46. Esquema del modelo convolucional doble.	146
Figura 6.47. Esquema del modelo convolucional simple.	146
Figura 6.48. a) Gráfica de exactitud de la arquitectura de una red convolucional simple para datos EEG. b) Gráfica de exactitud para la arquitectura de una red convolucional doble.	147
Figura 6.49. Diseño del experimento en la interfaz psychopy a través de bloques de programación y componentes de Emotiv.	149
Figura 6.50. Visualización del paradigma diseñado para la presentación de estímulos a los usuarios.	149
Figura 6.51. Adquisición de las señales EEG de los 5 canales de la diadema Insight, a través del software EmotivPro.....	150
Figura 6.52. Representación de los canales EEG segmentados en inicio y fin de cada estímulo, para un procesamiento más eficiente.....	151

ÍNDICE DE TABLAS

Tabla 3.1. Ventajas y desventajas de interfaces cerebro-máquina atendiendo a la naturaleza de la señal. Fuente: (Nicolas & Gómez, 2012)	27
Tabla 3.2. Ventajas y desventajas de métodos de extracción de características (Bacary, Alice, & Amine, 2020).	55
Tabla 4.1. Principales métodos de análisis, filtros y procesos de extracción de características en la literatura para el proceso de categorización de objetos.	83
Tabla 4.2. Clasificadores mayormente implementados en la literatura, basados en el mecanismo neural medido y la región cerebral activada.....	84
Tabla 4.3. Campos semánticos más clasificados en la literatura.	86
Tabla 5.1. Características principales de los dispositivos de adquisición de señales EEG de la empresa EMOTIV.	92
Tabla 5.2. Lista alfabética de las contribuciones de software destacados en la literatura. Las características principales de todos los paquetes de software se detallan en términos de análisis de series de tiempo (desde el preprocesamiento básico de datos sin procesar hasta métricas de sistemas dinámicos), análisis de tiempo-frecuencia y descomposición, modelado directo MEG / EEG y / o modelado de fuente (ajuste de dipolo, formación de haces / filtros espaciales e imágenes de origen), medidas y modelos de conectividad funcionales o efectivos, análisis estadístico e inferencia, interfaz gráfica de usuario versus bibliotecas de secuencias de comandos, interoperabilidad con otros paquetes de software / la posibilidad de que los usuarios desarrollen complementos, el entorno informático y el sistema operativo (s) (OS) requerido, y características definitivas que caracterizan cada paquete de software Fuente: adaptado de (Sylvain, Karl, & Oostenveld, 2011).	93
Tabla 6.1. Uso del vocabulario español.....	96
Tabla 6.2. Estructura del diccionario Toki Pona con pictogramas obtenidos del portal de ARASAAC correspondientes al significado	98
Tabla 6.3. Necesidades básicas humanas.	103
Tabla 6.4. Diseño de los experimentos o paradigmas más implementados en la literatura, para la presentación de estímulos visuales a una población determinada.	106
Tabla 6.5. Análisis de parámetros de evaluación, que son determinados por el tiempo disponible de la muestra (Zheng & Chen, 2021).....	107
Tabla 6.6. Tipo de características comúnmente extraídas en señales EEG.	126
Tabla 6.7. Ejemplo de codificación one-hot.....	136
Tabla 6.8. Parámetros evaluados en las CNN y exactitud del entrenamiento.....	142
Tabla 6.9. Comparación entre los tiempos de ejecución de una CPU y diferentes GPU's.....	147
Tabla 6.10. Parámetros evaluados en las CNN y exactitud del entrenamiento.....	148

CAPÍTULO 1

1. Introducción

La implementación de la tecnología de última generación como apoyo significativo de las actividades básicas del ser humano genera una indiscutible optimización de las tareas diarias y se han convertido en un elemento indispensable para el desarrollo dentro de la sociedad. Múltiples estudios han demostrado que al integrar la tecnología en procesos tempranos de educación y como apoyo en el tratamiento de deficiencias relacionadas con los trastornos del lenguaje, visión, audición, déficit de atención y la deficiencia mental, provocan altos impactos en el mejoramiento de los resultados terapéuticos e inclusión social (Garzón, Rojas, Sanabria, & Montenegro, 2013).

Con el desarrollo de nuevas tecnologías y la mejora de los sistemas de encefalografía, las investigaciones en esta área han logrado una interpretación de las ondas cerebrales que a simple vista podrían parecer totalmente aleatorias, para ello se buscó tratar las señales electroencefalográficas (EEG) crudas y aplicar distintos métodos matemáticos, que permiten obtener información que haga posible entender que suceden con las ondas cerebrales ante distintos estímulos externos, accionamientos cognitivos, enfermedades neurológicas, concentración, influencia de sustancias, etc. con el objetivo de lograr una comprensión más detallada de la actividad cerebral (Gómez, García, & Roque, 2016).

Para una correcta lectura de las señales producidas por el cerebro es necesario tener un dispositivo robusto que permita la lectura de las ondas que se requieren medir. Desde la invención de la electroencefalografía (EEG) se han realizado numerosas investigaciones con el objetivo de estudiar la relación existente entre las ondas cerebrales y los diferentes estados de conciencia (Morillo, 2014). Se conoce que los diferentes patrones de ondas cerebrales están directamente relacionados con diferentes estados de conciencia, estados que se pueden generar en un día cotidiano, como concentración intensa, estado de alerta (despierto), sueño profundo, sueños vívidos, somnolencia, relajación, hipnosis, estados alterados de conciencia, etc. (Aramayo, 2016).

Existen múltiples dispositivos relacionados a la exploración electroencefalográfica (EEG). Para el caso particular de esta investigación, se considera el uso de una diadema *Emotiv EPOC* (Emotiv, 2020), la cual basa su funcionamiento en un conjunto de sensores ubicados estratégicamente en distintas áreas de la cabeza, esto con el fin de interpretar las frecuencias producidas para detectar los pensamientos, los sentimientos y las expresiones del usuario. El

dispositivo detecta las señales primarias emitidas en el cerebro, estos tipos de onda interpretados por el sistema EEG se han clasificado en las siguientes bandas de frecuencia:

Gamma: superior a 30 (Hz); Beta (13-30 Hz); Alpha (8-12 Hz), Theta (4-8 Hz) y Delta (menor a 4 Hz).

Dispone de 14 canales que se encargan de recibir, cifrar y enviar las señales a través del puerto de comunicación para su pos procesamiento, el cual cuenta con un protocolo de abstracción lógica denominada *EmoEngine* (Garzón, Rojas, Sanabria, & Montenegro, 2013). Para acceder a los registros de las señales y monitorizarlas en el equipo se hace uso de las herramientas SDK proporcionadas por el fabricante EmotivSystems SDK (Emotiv, 2020). Cuenta con múltiples herramientas que puede usarse para simular las señales de entrada, un emulador opcional encargado de enviar la información, con estas herramientas se puede simular pruebas de comportamiento del usuario al momento de usar la aplicación.

El Emotiv EPOC posee 3 Suites para la detección de las señales de entrada: *Expressiv*, que descifra las expresiones faciales; *Affectiv*, cuya descripción es el estado emocional del usuario, y *Cognitiv*, cuya interpretación se basa sobre el uso consciente de los pensamientos (Garzón, Rojas, Sanabria, & Montenegro, 2013).

Una de las grandes ventajas que proporciona el Emotiv Epoc es que posee estas y otras funciones que permiten dar inicio a nuevas propuestas para obtener resultados más completos y funcionales para un margen mayor de usuarios.

El objetivo final de la interacción y cooperación entre humanos y máquinas es que pueda dar una respuesta razonable directamente a la intención del usuario. La tecnología de las interfaces hombre-máquina (BCI por sus siglas en inglés) se basa en el acceso directo a la información de la actividad física en los procesos de pensamiento del cerebro humano, lo que genera un neuro-camino eficaz que conecta el pensamiento y la inteligencia de la máquina. Esto se ha convertido en una importante dirección de desarrollo en el campo de la interacción persona-computadora. Las BCI basados en EEG, caracterizados por ventajas como la detección de señales no invasivas y la facilidad de uso, las investigaciones recientes concretan que tienen un gran potencial de aplicaciones prácticas.

Sin embargo, aunque se ha logrado mucho progreso para las BCI no invasivos, las señales cerebrales disponibles que pueden usarse como señales de control a la vez son todavía limitadas (Jia, et al., 2012).

1.1. Antecedentes

El sistema nervioso se encarga de organizar, controlar y dirigir las funciones de todos los sistemas del organismo; como nadar, caminar, percibir aromas o realizar actividades cotidianas como conducir un automóvil. Del mismo modo, el cuerpo humano tiene mecanismos de alto reflejo que se dan cuando estamos frente a posibles situaciones peligrosas, como tocar accidentalmente una superficie caliente (Dugarte, et al., 2018).

El cerebro está compuesto por millones de neuronas, y cada neurona tiene entre 1500 y 9000 conexiones con otras neuronas. La actividad cerebral está determinada específicamente, como el procesamiento de la información que existe entre los billones de conexiones nerviosas. Esta complicada red nerviosa está asociada a procesos como el pensamiento abstracto y la inteligencia. Las estructuras neuronales dan lugar a emociones más primitivas como el miedo o el placer, o sentimientos como el amor y el altruismo (Arango, 2004). Sin

embargo, en una red tan compleja y grande es posible que se presenten fallas que desencadenen patologías cognitivas.

En patologías como la parálisis cerebral y la esclerosis múltiple, el paciente es excluido del uso independiente de elementos tecnológicos debido a la total incapacidad motora, lo que ocasiona un agravamiento en la integridad y el desarrollo de la persona. En patologías como el retraso leve, el autismo y el síndrome de down, el factor determinante es la limitación cognitiva y de igual manera conlleva a un impedimento en el uso de herramientas de la vida cotidiana (Garzón, Rojas, Sanabria, & Montenegro, 2013).

La interacción BCI, es el estudio de la relación que existe entre los usuarios humanos y los sistemas de cómputo que usan para realizar distintas actividades. Su objetivo es entender cómo se comportan los usuarios, las tareas que requieren desempeñar y la manera en que los sistemas informáticos necesitan estar estructurados para facilitar el cumplimiento de estas mismas, es decir sistemas seguros y usables (García, Cruz, & Coria, 2015).

La literatura reporta una gran cantidad de algoritmos de clasificación de señales cerebrales que han sido probados para el desarrollo de sistemas BCI, presentando algunos inconvenientes en la toma y la clasificación de las señales. Existen investigaciones que discuten sobre las ventajas de los diferentes clasificadores, presentando propuestas con resultados significativos en la modelación de sistemas no lineales, además de sistemas que se aproximan a la forma en que los humanos procesan la información, demostrando una alta eficiencia (Bedoya, Giraldo, Torres, & Becerra, 2013).

Durante el desarrollo de dispositivos tecnológicos que interactúan con el ser humano, lo fundamental es diseñar sistemas que sean eficientes, efectivos y con facilidad de control, con el objetivo de superar las problemáticas que surgen en el proceso BCI relacionadas con el tipo de señales a procesar, lenguaje de programación o el hardware utilizado (Manzano, Placencia, Ayala, Chicaiza, & Carrera, 2017).

Dentro de la interacción humano-computadora, hay líneas que han tomado relevancia como lo son, las interfaces BCI y la inteligencia artificial. Una tarea en particular que tiene suma importancia es la de realizar análisis de ondas EEG para determinar emociones a partir de estímulos visuales, donde integran al usuario y sus emociones como datos importantes para inferir o apoyar en la toma de decisiones, es decir, involucrar los aspectos cognitivos, esta integración es conocida como computación afectiva (García, Cruz, & Coria, 2015).

De igual forma, se ha demostrado la efectividad de lograr un control a través de una BCI, mediante estímulos externos oculares, debido a que la actividad ocular tiene un rol dominante en los estudios con EEG, lo cual no ha sido muy común en los trabajos de investigación y publicaciones estudiadas.

Otra de las áreas altamente citadas están orientadas a la identificación y categorización de objetos, en las cuales los autores analizan la áreas del cerebro que se activan al momento de que los usuarios perciben estímulos visuales, ya sea a través de imágenes a color, pictogramas o textos (verbos, sustantivos, nombres de herramientas, etc.), y a través de un proceso completo de procesamiento, es posible saber la categoría semántica a la que pertenece el estímulo que el usuario está observando en intervalos de tiempo sumamente cortos (milisegundo).

La identificación de los procesos neurales que subyacen a las representaciones semánticas es un desafío clave en la neurociencia cognitiva. Se han propuesto diferentes hipótesis sobre

cómo las representaciones de conceptos particulares establecen un sistema de conocimiento conceptual. La aceptación general es que las propiedades de los objetos compartidos se reflejan en la organización del sistema semántico y que el sistema se generaliza a través de conceptos que pertenecen a una categoría particular (como animales, herramientas o edificios).

Uno de los principales problemas al desarrollar una BCI es decidir cómo se van a preprocesar las señales adquiridas del EEG, de manera de poder analizar y clasificar información útil a partir de la información que a simple vista puede parecer aleatoria (Gómez, García, & Roque, 2016).

Más de mil millones de personas viven en todo el mundo con discapacidad en alguna de sus variantes; de ellas, casi 200 millones tienen dificultades considerables en su funcionamiento. En los próximos años, la discapacidad será un motivo de preocupación aún mayor, pues su prevalencia está aumentando. Esto debido a que la población está envejeciendo y el riesgo de discapacidad aumenta entre los adultos mayores. Otra causa es el aumento mundial de enfermedades crónicas como la diabetes, los trastornos de salud mental, el cáncer y las enfermedades cardiovasculares.

En todo el mundo, las personas que sufren una discapacidad tienden a tener malos resultados sanitarios, malos resultados académicos, menor participación económica y tasas de pobreza más altas que las personas sin discapacidad. En parte, esto se debe a que existen ciertos obstáculos que limitan el acceso de las personas con discapacidad a servicios que muchos de nosotros consideramos obvios, en particular la educación, la salud, el transporte, la información o el empleo (Chan, 2011).

Se estima internacionalmente que la discapacidad intelectual (DI) afecta del 2 al 3% de la población general del mundo desarrollado. A pesar del éxito de las intervenciones médicas específicas en los últimos cincuenta años en cuanto al diagnóstico y tratamiento de procesos que conllevan discapacidad intelectual, hoy en día, en la mitad de los casos su causa es desconocida y por lo tanto, la disminución de estas, o un diagnóstico y solución requeridos han sido escasos (Fernández, et al., 2010).

De acuerdo a la Organización mundial de la salud (OMS) la “discapacidad es un término general que abarca las deficiencias, las limitaciones de la actividad y las restricciones de la participación. Las deficiencias son problemas que afectan a una estructura o función corporal; las limitaciones de la actividad son dificultades para ejecutar acciones o tareas, y las restricciones de la participación son problemas para participar en situaciones vitales”.

Por consiguiente, “la discapacidad es un fenómeno complejo que refleja una interacción entre las características del organismo humano y las características de la sociedad en la que vive” (OMS, s.f.).

Se calcula que más de mil millones de personas, es decir, un 15% de la población mundial presentan algún tipo de discapacidad. Entre 110 millones (2,2%) y 190 millones (3,8%) de personas mayores de 15 años, tienen dificultades importantes para funcionar (OMS., 2018).

La accesibilidad de los medios electrónicos es la facilidad de uso de las tecnologías de la información y la comunicación (TIC), tales como Internet, por personas con discapacidad. La presentación de los sitios web debe permitir que los usuarios discapacitados tengan acceso a la información (OMS-, 2013).

Datos y cifras

- Más de mil millones de personas, o sea, un 15% de la población mundial, padece alguna forma de discapacidad.
- Entre 110 millones y 190 millones de adultos tienen dificultades considerables para funcionar.
- Las tasas de discapacidad están aumentando a causa del envejecimiento de la población y el aumento de las enfermedades crónicas, entre otras causas.
- Las personas con discapacidad tienen menos acceso a los servicios de asistencia sanitaria y, por lo tanto, necesidades insatisfechas a este respecto (OMS., 2018).

1.1.1. La discapacidad en México

El tema de la discapacidad ha cobrado un significativo incremento de atención por parte de los profesionales y la sociedad en su conjunto en los últimos años. La problemática, dificultades y barreras que enfrentan las personas con discapacidad han sido expuestas y analizadas públicamente y cada vez son más las personas, instituciones y asociaciones que se incorporan a la tarea de mejorar las condiciones de vida y bienestar personal de este grupo poblacional.



Caminar o moverse. Hace referencia a la dificultad de una persona para moverse, caminar, desplazarse o subir escaleras debido a la falta de toda o una parte de sus piernas; incluye también a quienes teniendo sus piernas no tienen movimiento o presentan restricciones para moverse, de tal forma que necesitan ayuda de otras persona, silla de ruedas u otro aparato, como andadera o pierna artificial.



Ver. Abarca la pérdida total de la vista en uno o ambos ojos, así como a los débiles visuales y a los que aun usando lentes no pueden ver bien por lo avanzado de sus problemas visuales.



Mental. Abarca cualquier problema de tipo mental como retraso, alteraciones de la conducta o del comportamiento.



Escuchar. Incluye a las personas que no pueden oír, así como aquellas que presentan dificultad para escuchar (debilidad auditiva), en uno o ambos oídos, a las que aun usando aparato auditivo tiene dificultad para escuchar debido a lo avanzado de su problema.



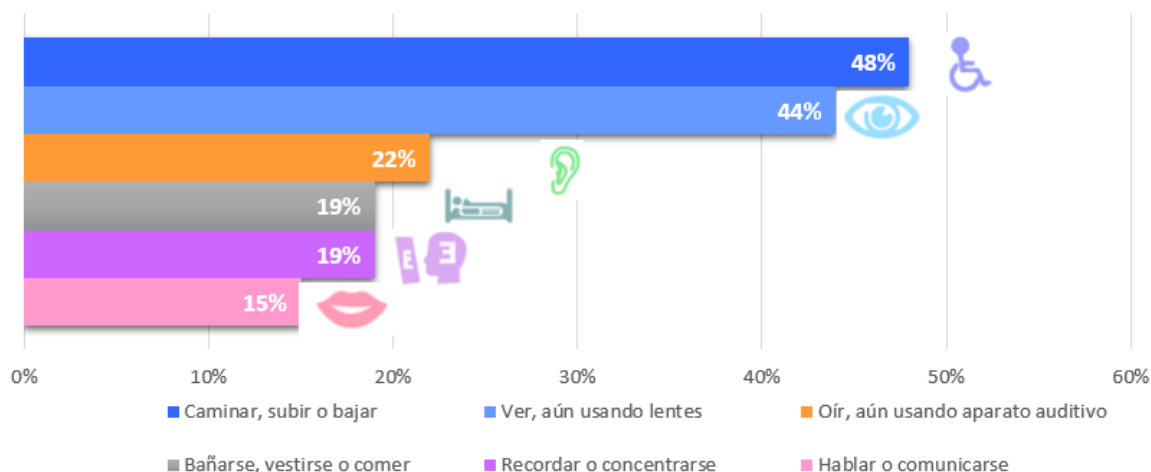
Hablar o comunicarse. Hace referencia a los problemas para comunicarse con los demás, debido a limitaciones para hablar o porque no pueden platicar o conversar de forma comprensible.



Atención y aprendizaje. Incluye las limitaciones o dificultades para aprender una nueva tarea o para poner atención por determinado tiempo, así como limitaciones para recordar información o actividades que se deben realizar en la vida cotidiana.



Autocuidado. Hace referencia a las limitaciones o dificultades para atender por sí mismo el cuidado personal, como bañarse, vestirse o tomar alimentos (INEGI, 2020).



Gráfica 1.1 Porcentaje de la población con discapacidad según dificultad en la actividad (Año 2020). Nota: La suma de porcentajes es mayor a 100 por la población que presenta más de una dificultad. (INEGI, 2020).

Los motivos que producen discapacidad en las personas pueden ser variados, pero el INEGI los clasifica en cuatro grupos de causas principales: nacimiento, enfermedad, accidente y edad avanzada.

De cada 100 personas con discapacidad:

- 39 la tienen porque sufrieron alguna enfermedad.
- 23 están afectados por edad avanzada.
- 16 la adquirieron por herencia, durante el embarazo o al momento de nacer.
- 15 quedaron con lesión a consecuencia de algún accidente.
- 8 debido a otras causas (INEGI, 2020).

1.2. Justificación

Las limitaciones y dependencia que presenta y demanda una persona con una discapacidad física eventual, parcial o total, generan dificultades para llevar a cabo las actividades básicas

de su vida diaria, lo que implica que estas personas desarrollen sus propios medios para enfrentarse a este tipo de situaciones y que le permiten una dependencia parcial.

Actualmente se están desarrollando aplicaciones orientadas hacia las BCI las cuales tienen como objetivo generar interactividad entre personas con discapacidad y cualquier dispositivo que facilite su condición de vida, es por esto que el estudio y la generación de control interpretando señales cerebrales electroencefalográficas (EEG) presenta una potencial área de investigación, ya que es posible convertirlas en un medio de comunicación que facilite o permita a los usuarios una adaptación real con el medio exterior, así como su aplicabilidad a sistemas robóticos, automatizados, que se convierten en sistemas de asistencia para las personas. Este tipo de interfaces tienen dos características que las hacen únicas ante todos los sistemas existentes. La primera es su potencial para construir un canal de comunicación natural con la persona, la segunda su potencial acceso a la información emocional y cognitiva del usuario (Minguez, 2010).

Sin embargo, las BCI tienen varios requisitos para su adecuado funcionamiento que van evolucionando con el transcurso de los avances y métodos tecnológicos, siendo mayores los requerimientos que hacen a estos sistemas cada vez más complejos y donde se centran las investigaciones para aportar en desarrollos más notorios; uno de estos requerimientos es garantizar que las señales describan de manera muy acertada y real los distintos estados cerebrales que estén disponibles en el proceso de adquisición, aunado a esto, la facilidad de extracción de las señales, para que se puedan diferenciar respectivamente y realizar una correcta clasificación (Oñate, 2018).

Esta investigación brindará soluciones tecnológicas en el área de las BCI, que pueden ser orientadas a las personas con discapacidad física, con el objetivo de realizar una adquisición de las diferentes señales que son emitidas por el cerebro al percibir conceptos de manera visual. Estas señales pueden verse afectadas y variar de acuerdo a factores que se encuentran inmersos en el cerebro de la persona, como el estado emocional, físico y anímico; y esta investigación permitirá tener un panorama amplio sobre la relación que existe entre un concepto externo y el cerebro humano, para lograr alcanzar un beneficio social, y reducir la dependencia personal en el desarrollo de actividades de la vida diaria, que por consiguiente permiten una mejora en la calidad y estilo de vida.

El presente trabajo se organiza de la siguiente manera: en el capítulo 2 se expone el marco epistémico, en el cual se aborda el problema de investigación, los objetivos de la investigación, la hipótesis que ayudará a la resolución del problema y los alcances y limitaciones de la tesis. En el capítulo 3 se presenta el marco teórico de los conceptos requeridos para la comprensión del problema, así como los conceptos fundamentales de la teoría relacionada a los objetivos. En el capítulo 4 se presenta el estado del arte, en donde se analizan los trabajos existentes en la literatura relacionados al tema expuesto, y las limitantes encontradas en las investigaciones actuales. En el capítulo 5 se menciona la metodología propuesta que será implementada para la conclusión de esta investigación. En el capítulo 6 se muestran los resultados de la metodología propuesta sobre el tema en cuestión, los métodos que se siguieron para concluir los objetivos y los resultados obtenidos al realizar los experimentos. Finalmente en el capítulo 7 se expresan las conclusiones de los resultados obtenidos de este trabajo de tesis, así como la discusión de estos, donde se exponen los puntos más sobresalientes de la investigación, los procesos por mejorar y que se pueden realizar como un trabajo a futuro.

CAPÍTULO 2

2. Marco epistémico

En este capítulo se describe el marco epistémico de la investigación, donde se expone cual es el problema de investigación a resolver, dar a conocer los objetivos, general como particulares que este trabajo busca culminar, las preguntas de investigación que se hacen para entender la problemática, se define una hipótesis que se buscará aceptar o refutar a lo largo del trabajo y las limitantes de esta tesis.

2.1. Problema de investigación

El ser humano a través de su evolución ha implementado distintos mecanismos para su comunicación, uno de ellos es el habla, otro es el lenguaje de señas y mensajes a través del cuerpo. Esos mecanismos de comunicación permiten establecer un mensaje hacia el receptor, y que el receptor obtenga de manera satisfactoria ese mensaje por parte del emisor.

Hoy en día, existe una cantidad significativa de la población que tiene problemas al comunicarse, ya sea porque ha perdido el habla, personas que no escuchan, o personas que no pueden comunicarse a través de algún protocolo establecido. Los trastornos del habla y del lenguaje en numerosos casos han significado que muchas personas corran el riesgo de exclusión social, ya que se encuentran sin acceso a formas de comunicación clave, como escribir, leer y hablar (Bautista, et al., 2017). Además, estas formas de comunicación se han vuelto más relevantes en relación con las nuevas tecnologías, con los métodos de búsqueda de información y los medios de difusión de información actuales.

El cerebro humano genera ondas eléctricas para realizar cualquiera de sus funciones básicas y elementales, así como para recibir, enviar y procesar información. Estas señales pueden ser adquiridas de manera externa y manipuladas a través de un sistema de lecturas electroencefalográficas (EEG) (Bedoya, Giraldo, Torres, & Becerra, 2013).

Se requiere caracterizar las señales EEG generadas por el cerebro al percibir un concepto, e identificar si existen asociaciones entre estas señales y tales conceptos. Es necesario verificar si existe un conjunto de valores discriminantes asociados a cierto concepto, es decir, un patrón regular en la persona en diferentes instantes del tiempo, y más interesante, si ese

patrón existe con regularidad en diferentes personas. Se puede validar si elementos externos afectan la generación de ondas cerebrales al exponer al usuario ante un ambiente distinto, como por ejemplo, diferentes temperaturas, música, intensidad de luz en el ambiente, etc.

Este tipo de sistemas en los que se puede medir la actividad cerebral e interaccionar con el entorno de la forma deseada por el usuario, pueden tener aplicaciones en la vida real muy concretas, por ejemplo, como apoyo a personas que perdieron el habla, personas que no escuchan, personas que han tenido un derrame cerebral y se necesita verificar si la identificación de los conceptos aún se encuentran presentes en su mente.

2.2. Objetivos de la tesis

A continuación se presenta el objetivo general de esta investigación, lo que pretende resolver, así como los objetivos particulares que apoyarán a la solución del problema general.

2.2.1. Objetivo general

El objetivo de esta investigación es generar un modelo computacional que sea capaz de identificar automáticamente las características o rasgos discriminantes que generan las ondas cerebrales a través de EEG, cuando el ser humano percibe conceptos, utilizando un conjunto de estímulos relacionados con el proceso de comunicación humana.

2.2.2. Objetivos específicos

1. Construir un conjunto de datos de conceptos basado en estímulos para su uso en los experimentos.
2. Definir un protocolo para realizar el correcto procesamiento de las señales cerebrales generadas por el usuario al percibir conceptos.
3. Diseñar un proceso de recopilación y almacenamiento de señales electroencefalográficas (EEG) asociadas a los conceptos, sobre una población estadísticamente significativa.
4. Evaluar los resultados obtenidos con la finalidad de determinar el rendimiento basado en la metodología propuesta.
5. Implementar un sistema de identificación automática de conceptos.

2.3. Preguntas de investigación

A continuación se presentan las preguntas de investigación relacionadas al problema expuesto, las cuales ayudarán a exponer un panorama general de lo que se busca solucionar.

¿Cuál es la metodología más eficiente para la recopilación de las señales generadas por el cerebro, cuando este identifique un concepto?

¿Cómo se puede elaborar una base de datos de conceptos asociados al proceso de comunicación y que esta funcione de manera genérica en una población determinada?

¿Cuál es el protocolo a seguir para identificar los patrones generales presentes en las señales EEG correspondientes a un concepto?

¿Mediante qué métricas se realizará la evaluación y confiabilidad de las señales medidas?

¿Cómo se pueden comparar los valores discriminantes en ondas EEG de diferentes personas asociadas al mismo concepto?

¿De qué manera se puede clasificar la acción que pensamos y asociarla a su significado correspondiente?

2.4. Límites de la tesis

Esta investigación se enfoca en conceptos orientados al proceso de comunicación humana, por lo que no se utilizarán todos los conceptos, ya que existe una infinidad de estos, entonces se trabajará con conceptos limitados que estarán asociados a un dominio en específico, como acciones específicas, verbos o sustantivos previamente seleccionados, que generan un modelo conceptual basado en las necesidades humanas.

El tiempo de entrenamiento para cada persona de la población puede ser un factor limitante, debido a que la correcta preparación de los usuarios, y la presentación de los conceptos objetivo requiere determinado tiempo. Es posible que los usuarios no se encuentren en un estado de ánimo necesario o en concentración requerida para la correcta medición de las señales cerebrales.

En primera instancia se trabajará con personas sin discapacidad, con el objetivo de tener una retroalimentación adecuada y garantizar que los resultados obtenidos sean los correctos, para poder identificar como primer paso, las características específicas asociadas a los conceptos. Se realizará en colaboración con investigadores expertos en el área, con la intención de definir una población y muestra adecuada, con base a la experiencia en trabajos previos con interfaces BCI.

CAPÍTULO 3

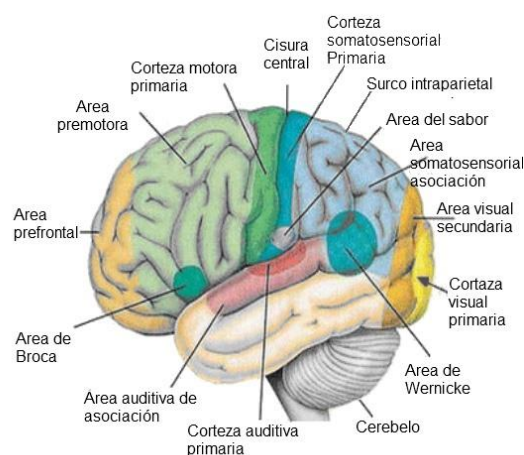
3. Marco teórico

En este capítulo se presentan conceptos fundamentales asociados a las BCI, al cerebro humano, su funcionamiento, los tipos de ondas que genera, e información requerida para adentrar y relacionar al lector con el tema expuesto.

3.1. El cerebro

El cerebro es la parte principal del sistema nervioso central del ser humano, y se encuentra ubicado dentro del cráneo. En la especie humana tiene un peso aproximado de 1,3 kg y es una masa de tejido gris-rosáceo que está compuesto por aproximadamente 100,000 millones de células nerviosas (cantidad de células en un cerebro adulto), conectadas entre sí y son las encargadas del control de todas las funciones mentales. Además de las funciones sensoriales, motoras, sensitivas y neuro humorales, en el cerebro se controlan todas las funciones elevadas del ser humano, como lo es la lectura, la escritura, el habla, el pensamiento, el juicio, la comprensión, controla las acciones como el movimiento, el sueño, el hambre, la sed y casi todas las actividades vitales necesarias para la supervivencia del ser humano. Asimismo, es el encargado de generar las emociones humanas como el miedo, la ira, la tristeza, el amor, el odio y la alegría. También es el encargado de recibir e interpretar las innumerables señales que se envían desde el organismo y del exterior. La gran superficie del cerebro y su compleja estructura interna justifican el nivel de superioridad de inteligencia que posee el hombre sobre otras especies vivientes (Oñate R. , 2018).

La comprensión de la función cerebral, especialmente de la corteza cerebral, se inició a finales del siglo XIX. Fue en 1861 cuando el médico francés Paul Pierre Broca presentó a la Sociedad Francesa de Antropología el caso de un paciente quien padecía un trastorno del lenguaje, cuando el paciente falleció, el Dr. Broca encontró en la autopsia una lesión en la circunvolución prefrontal inferior del hemisferio cerebral izquierdo. El Dr. Broca llegó a la conclusión que existía una relación entre lenguaje y función cerebral, y que lesiones en el área prefrontal izquierda causaban una alteración del lenguaje. Con este trabajo se inicia la neurología moderna y en su honor se denominó a esta zona el área de Broca y la patología



La corteza cerebral está dividida por una fisura longitudinal, en una parte derecha y otra izquierda, llamadas hemisferios cerebrales. Ambos hemisferios (derecho-izquierdo), son simétricos y se encuentran interconectados por medio del denominado "cuerpo calloso" que es un conglomerado de fibras nerviosas blancas que los conectan y permiten la transferencia de información de un hemisferio a otro.

23

aproximaciones o estimaciones matemáticas, así como también controla el lado izquierdo de nuestros cuerpos (Argopsique, 2014). En la Fig. 3.2 se presenta un diagrama con las funciones principales de cada hemisferio del cerebro humano.

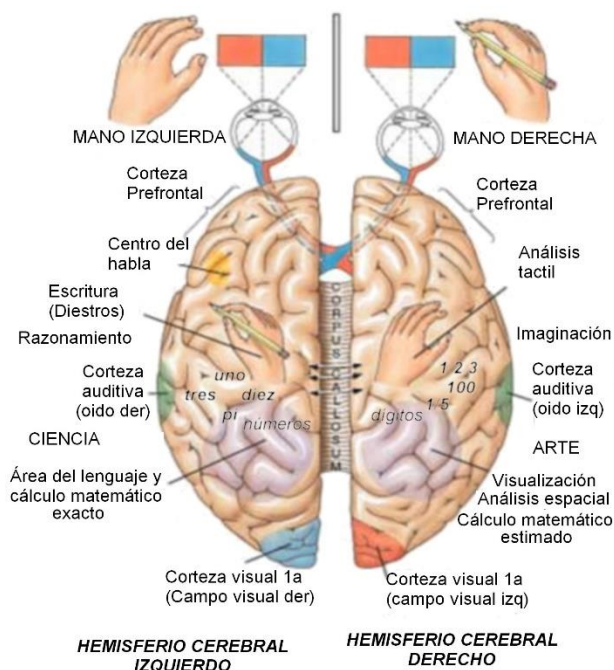


Figura 3.2. Especialización de los hemisferios cerebrales. En general las funciones cerebrales están más deslocalizadas de lo que se creía, pero hay algunas funciones que se realizan con más intensidad en una mitad que en otra (Argopsique, 2014), (Vargas, 2012).

El cerebro es el órgano del cuerpo que más trabaja, ya que todo aquello que se hace, se siente o se piensa, es gracias al cerebro, si se realiza una comparación con una computadora, la diferencia es muy interesante y notoria, por ejemplo, la computadora realiza sus operaciones por medio de procesos secuenciales y lógicos, mientras que el cerebro es multidireccional, funcionando en una forma mucho más compleja ya que al procesar la información lo hace de manera paralela, puede sintetizar e integrar esta información de manera simultánea. Toda la información que puede recibir el cerebro del mundo exterior, se obtiene por medio de los sentidos (gusto, tacto, olfato, vista y oído) cuyas sensaciones se reciben a través de los órganos respectivos (lengua, piel, nariz, ojos y oídos) que a su vez están controlados por el cerebro; estos órganos receptores y sensibles envían información de manera constante para saber todo lo que sucede a nuestro alrededor, aunque a pesar de toda esta información controlada por el cerebro, los estudios demuestran que el ser humano utiliza únicamente un diez por ciento (10 %), aproximadamente, de la capacidad de este órgano.

De igual manera se ha determinado en investigaciones, que cada hemisferio se interrelaciona íntimamente con su homólogo, aunque ejercen funciones diferentes y cada uno es responsable de un lado del cuerpo, en forma especular o de espejo, es decir que las funciones realizadas por el lado izquierdo del cuerpo son dirigidas y controladas por el hemisferio derecho, de manera similar para el hemisferio izquierdo.

Aunque existen varios aspectos del cerebro, la electroencefalografía (EEG) se limita a la corteza, que es una estructura plegada que forma la capa externa del cerebro como se muestra en la figura 3.3. Cada una de las áreas en la corteza tiene diferentes funciones y habrá diferentes señales que emanan de ellas.

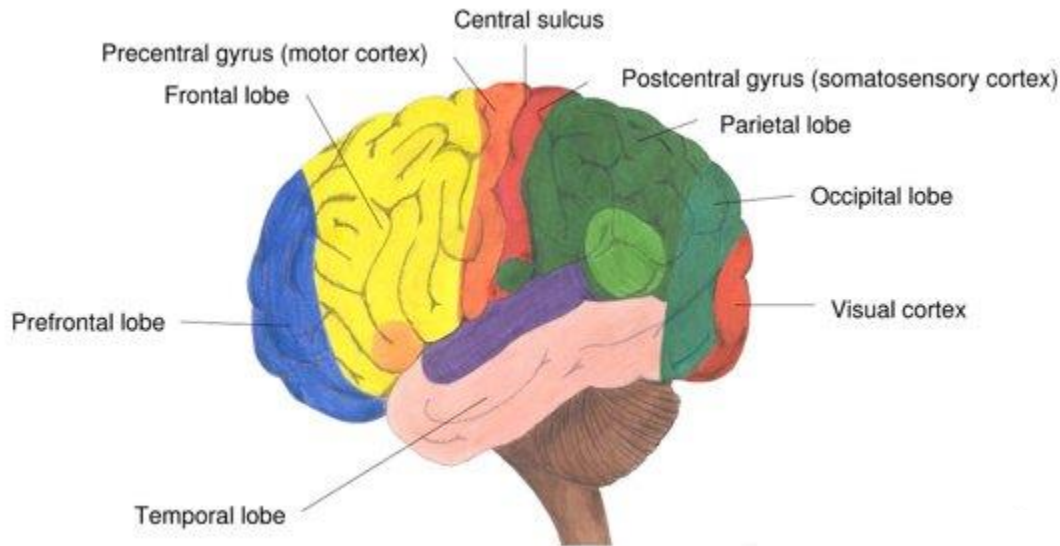


Figura 3.3. La corteza cerebral marcada muestra diferentes áreas funcionales (AAMODH, 2016).

El lóbulo frontal, que ocupa un lugar importante en el cerebro, es responsable de la resonancia, la resolución de problemas, el control del movimiento, la planificación y el control de los movimientos voluntarios.

El lóbulo temporal participa en el procesamiento del lenguaje, la memoria y las funciones del sistema auditivo,

El lóbulo parietal está interviniendo en el tratamiento de la información sensorial,

El lóbulo occipital que permite el reconocimiento de formas, colores e información visual.

Las señales grabadas generalmente tienen amplitudes que varían de unos pocos microvoltios a 100 μV aproximadamente con frecuencias que varían de 0.5 a 30–40 Hz [8].

La señal EEG se caracteriza por:

- Rango de amplitud: Despierto ($V_{pp} = 100 \mu\text{V}$), dormido ($V_{pp} = 300 \mu\text{V}$)
- Rango de frecuencia: de 0.01 a 100 Hz
- Artefactos EEG comunes: parpadeo de los ojos (EOG), actividad muscular (EMG), ruido ambiental, movimiento de electrodos (Bacary, Alice, & Amine, 2020).

3.2. Ondas cerebrales

El cerebro humano tiene la capacidad de procesar 64,000 pensamientos por minuto. En la comunidad científica, es poco conocido el impacto que tiene la meditación sobre el cerebro humano, pero se sabe con certeza que el ser humano, cuando percibe y está consciente, emite un patrón de ondas cerebrales que permiten identificar su estado.

El cerebro produce impulsos eléctricos (potenciales de acción) que viajan a través de las neuronas. Estos impulsos eléctricos producen ritmos diferentes que son conocidos como ondas cerebrales. Los impulsos eléctricos son información que viaja de neurona a neurona haciendo uso de cientos de miles de ellas para lograr transportarse y ejecutar una función determinada. Están inmersos procesos bioquímicos muy complejos que permiten el pasaje de la actividad eléctrica de una neurona hacia otra. Este fenómeno es conocido como sinapsis. Esta actividad eléctrica es constante y tiende a variar dependiendo de las circunstancias, los momentos del día y los estados del individuo. Esta actividad eléctrica es medida en unidades de micro voltios, y conforme la frecuencia de descarga eléctrica de los grupos de neuronas la actividad cerebral será diferente.

Esta actividad eléctrica cerebral puede ser registrada a través de un electroencefalógrafo, que es un equipo que la registra y representa a través de ondas que tiene diferentes propiedades. Las propiedades básicas de esas ondas cerebrales son: frecuencia y potencia.

La frecuencia es el número de ondas que se registran en un segundo, y se mide en ciclos por segundos (cps) o Hertzios. La potencia hace referencia a la altura que adquiere cada onda en el registro. La potencia se mide en micro voltios. A nivel cerebral la potencia se traduce como la suma de una mayor o menor cantidad de neuronas activadas en un determinado momento (Oñate R. , 2018).

3.2.1. Origen fisiológico de las señales

Al momento de analizar la actividad cerebral de los usuarios se tiene en consideración que existen dos tipos principales de señales que pueden ser monitorizadas. Otra característica importante a considerar sobre los fenómenos neurológicos de interés para realizar interfaces BCI es el origen de las señales deseadas y saber si son ocasionadas por una estimulación externa o bien de una actividad inherente al propio individuo (Morán, 2015).

Señales evocadas o exógenas

Las señales evocadas o exógenas son aquellas que se generan a través de una estimulación exterior. Muchas de las señales comúnmente utilizadas al realizar interfaces BCI dependen de estímulos visuales exteriores. Sin embargo, la estimulación exterior no es únicamente visual sino que puede implicar alguna otra actividad sensorial como la auditiva (Gao, Wang, Gao, & Hong, 2014). Las interfaces exógenas son BCI que utilizan este tipo de señales.

Señales no evocadas o endógenas

Las señales endógenas, en divergencia con las exógenas, son aquellas que se generan en el interior del cerebro sin que exista alguna estimulación externa. Dependen de la habilidad del usuario para controlar su actividad electrofisiológica (Hornero, Corralejo, & Álvarez, 2012). Las interfaces endógenas utilizan este tipo de señales.

Para el diseño de interfaces que implementen señales endógenas o exógenas, es bueno tener en consideración la siguiente tabla de ventajas y desventajas asociadas a la naturaleza de las oscilaciones:

Tabla 3.1. Ventajas y desventajas de interfaces cerebro-máquina atendiendo a la naturaleza de la señal. Fuente: (Nicolas & Gómez, 2012)

	Ventajas	Desventajas
Interfaces exógenas	Poco entrenamiento, rápida configuración, alta tasa de transferencia y funcionalidad con pocos canales de información.	Necesario prestar atención continuada al estímulo, lo que puede producir cansancio en el usuario.
Interfaces endógenas	No dependen de estímulos, lo que lo hace útil para personas con órganos sensoriales afectados.	Largo entrenamiento que no garantiza obtener control sobre el sistema. Baja tasa de transferencia y requiere de múltiples canales de información.

3.2.2. Tipos de ondas

En función de la frecuencia, la actividad eléctrica cerebral puede clasificarse en cinco tipos de ondas eléctricas básicas: gamma, alfa, beta, theta y delta (ver Fig. 3.4) (Damasio, 2010). A continuación se describen estos diferentes tipos de ondas cerebrales, en orden de mayor a menor actividad.

Ondas Beta

Se producen cuando el cerebro está despierto y actuando en actividades mentales. Son ondas amplias y las de mayor velocidad de transmisión de las cuatro. Originan un campo electromagnético y su frecuencia oscila entre 14 y 30 o 35 Hz (ciclos por segundo o cps). Denotan una actividad mental intensa. Cuando una persona está dando un discurso, estudiando, realizando un problema de matemáticas, etc. su cerebro se encuentra emitiendo este tipo de ondas.

Característicamente los niños con déficit de atención presentan una relación entre ondas lentas y rápidas a favor de las primeras.

Según su frecuencia se clasifican distintos tipos de ondas Beta:

Ondas Beta Inferiores (Low Beta) o Beta 1

Son consideradas como las ondas de alta actividad cerebral. Su frecuencia oscila entre 12 y 16 Hz. Cada vez que el usuario se enfoca, analiza, hace cálculos o piensa en su ambiente externo, las ondas beta son generadas. En un adulto se encuentran cantidades significativas de ondas beta en comparación con los niños. El exceso de ondas beta en el hemisferio derecho del cerebro se relaciona a estados de ansiedad, preocupaciones y tensión. Las ondas beta que se encuentran en el hemisferio izquierdo del cerebro se consideran saludables.

El aumento de ondas beta puede ser de beneficio en caso de que la persona se sienta deprimida. Las ondas beta en exceso se asocian a desordenes de: ansiedad, insomnio, y trastornos de índole obsesivo-compulsivo (TOC). En los momentos de stress y tensión hay un aumento significativo de ondas beta en el cerebro (Sciottto & Niripil, 2018).

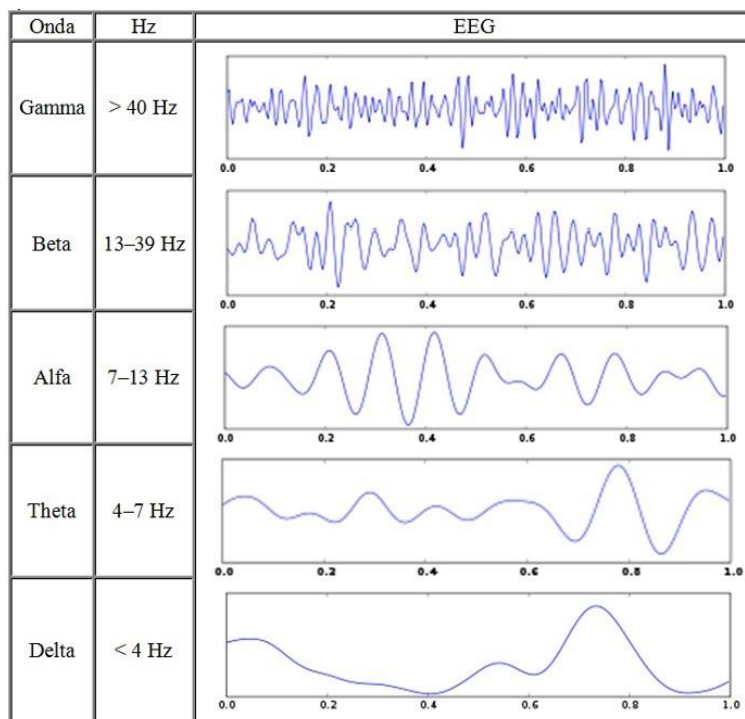


Figura 3.4. Clasificación de ondas cerebrales respecto a su frecuencia de onda: Gamma, alfa, beta, theta y delta (Damasio, 2010).

Ondas Beta Superiores (High Beta) o Beta 2 -3

Las ondas beta superiores se asocian a miedos, pensamientos excesivos, ansiedad, pensamientos rápidos, trastorno obsesivo-compulsivo (TOC), adicciones y estados de desempeño extremo a nivel mental. En ocasiones, las ondas beta superiores son creadas por el cerebro para compensar por la actividad excesiva de ondas theta.

Si la persona está alerta de forma exagerada, nervioso o hipocondríaco, puede estar experimentando alta cantidad de ondas beta (Sciottto & Niripil, 2018).

Ondas Alfa

Alfa representa un estado de escasa actividad cerebral y relajación. Estas ondas son más lentas y de mayor amplitud que las beta. Su frecuencia oscila entre 8 y 14 Hz. Una persona que ha terminado una tarea y se sienta a descansar, se encuentra por lo general en un estado alfa; así como la persona que está dando un paseo, disfrutando del paisaje.

Estas ondas se generan momentos antes de dormirse, sus efectos característicos son: pensamientos tranquilos, relajación agradable, optimismo y un sentimiento de integración de cuerpo y mente. La desincronización alfa es un fenómeno normal que se presenta cuando un

individuo abre sus ojos y se encuentra con una imagen o un texto que atrapa su atención. Precisamente la ausencia de la desincronización alfa es característica en niños que presentan dificultades de aprendizaje de la lectoescritura o de comprensión. Son activadas por el descanso de nuestro sistema óptico (al cerrar los ojos). Cada vez que la persona se relaja o cierra sus ojos, la actividad de ondas alfa aumenta considerablemente. Se encuentran patrones normales de ondas alfa normalmente en personas que tienen balanceados sus hemisferios cerebrales izquierdo y derecho.

Las ondas alfa generalmente se encuentran en su mayoría en la parte trasera del cerebro, aunque también existen en las partes frontales, pero en menor cantidad. Si la persona presenta un exceso de ondas alfa en el hemisferio izquierdo de su cerebro, se puede traducir en que está deprimido. Por otro lado, si se encuentran cantidades excesivas de ondas alfa en la parte frontal del cerebro, esto se relaciona a la depresión y los desórdenes de déficit de atención e hiperactividad.

Una cantidad baja de ondas alfa al momento de levantarse de la cama en la mañana se le conoce como "Bloqueo Alfa". Esto ocasiona la llamada "jaqueca matutina", parecido a un dolor de cabeza. Por otro lado, las ondas alfa nos ayudan a enlazar la mente consciente con el inconsciente (o subconsciente) (Oñate, 2018).

Ondas Theta

Son ondas de mayor amplitud y menor frecuencia (entre 4 y 8 Hz). Se presentan bajo un estado de calma profunda. La persona que está fantaseando (o soñando despierta), se encuentra en este estado. Se dice que es un estado de inspiración de ideas y soluciones creativas. En este estado es como si las tareas realizadas se automatizan, ya no se necesita tener un control atencional y consciente de su ejecución, pudiendo el sujeto alejarse de ellas mentalmente, es decir, que su mente esté en "otro sitio".

Se producen durante los estados de meditación profunda, entrenamiento autógeno, yoga; así como durante el esfuerzo intelectual asociado al aprendizaje académico. Las características de este estado son: mayor capacidad de aprendizaje, memoria plástica, inspiración creativa fantasía e imaginación. Sin embargo, la presencia excesiva de ondas theta en niños, más allá de la edad esperada, o la asimetría en la distribución de la frecuencia theta, es indicador de lento procesamiento de la información declarativa.

Este predominio theta es característico de niños con menor plasticidad cerebral y cierto retraso en su maduración bioeléctrica. Las ondas theta están relacionadas a altos niveles de creatividad, emociones y espontaneidad en las personas; cuando el cerebro produce cantidades excesivas de estas ondas, la persona puede sentirse deprimida, sufrir de "soñar despierto" (daydreaming), puede tener hiperactividad, desórdenes de déficit de atención, sentirse distraído y en muchas ocasiones, ansioso.

También cuando el cerebro produce cantidades excesivas de ondas theta, la persona tiene pensamientos incompletos (particularmente en las personas distraídas), la toma de decisiones es afectada (hace una tarea y no la termina, por ejemplo), la persona sufre de impulsividad, y disminuye el tiempo de reacción y respuesta. Los niños generalmente tienen mucha más actividad de ondas theta en comparación con los adultos.

Estas ondas permiten recuperar la memoria a largo plazo y emociones reprimidas o represivas (Sciottto & Niripil, 2018), (Oñate, 2018).

Ondas Delta

Son las ondas de mayor amplitud y menor frecuencia (entre 1,5 y 4 Hz). Nunca llegan a cero, pues eso significaría la muerte cerebral. Se generan en un estado de “sueño profundo” y son las que dominan en el cerebro de los niños. Cuando es el momento de dormir, las ondas cerebrales van pasando sucesivamente de beta a alfa, theta y finalmente, delta. Durante el sueño se producen ciclos que duran aproximadamente 90 minutos. Las ondas delta con una amplitud alta y rítmica generalmente son encontradas en personas adultas con lesiones cerebrales o desórdenes mentales. Pueden observarse ondas delta en el EEG de los niños con déficit de atención y desordenes de hiperactividad (ADD o ADHD).

Si las ondas delta predominan, se traduce a pérdida de reconocimiento del cuerpo físico. Si la persona está en un estado inconsciente, las ondas delta son las que predominan en el cerebro. Existe un patrón de ondas delta presentes habitualmente en baja proporción en el cerebro normal de adolescentes, son llamadas "Ondas delta de la juventud".

Ondas Gamma

Son ondas cerebrales de alta frecuencia (40 Hz o más). Se asocian a las ondas generadas al momento de resolver problemas, sean estos lógicos o matemáticos, tanto en niños como en adultos. Se pueden encontrar ritmos gamma en cualquier parte del cerebro. Su presencia favorece el procesamiento de la información declarativa a través de las vías cerebrales.

Las personas con desordenes de aprendizaje y deficiencias de memoria, tanto a corto como a largo plazo, pueden tener necesidad de mayor cantidad de ondas gamma en su cerebro. Se ha observado incremento de la potencia de ondas gamma en niños con antecedentes de dificultades en el aprendizaje de la lectoescritura, como en disléxicos severos (Sciottto & Niripil, 2018).

Las ondas gamma son las que más llaman la atención en el ser humano debido a que son las más rápidas y las que mayor actividad mental representan (momentos repentinos de lucidez, brillantez, intuición, y concentración extrema). También pueden realizarse mediciones bajo las mismas condiciones y encontrar a personas situadas en los rangos de 26 Hz a 70 Hz.

Las actividades cognitivas de alto nivel, que son las que más requieren el uso del intelecto, se dan cuando se supera el límite de los 24 Hz ya que es en este punto cuando se detectan las ondas gamma. Las ondas gamma también pueden estar presentes en la fase REM del sueño, aunque son propias de la reflexión cognitiva. De hecho, los últimos estudios realizados, sugieren que la capacidad de resolver problemas se relaciona directamente con el aumento de las frecuencias de 20 Hz a 40 Hz (Oñate R. , 2018), (Aguilar, 2013).

3.3. Adquisición de señales cerebrales

Algunas de las técnicas que se utilizan para registrar la actividad cerebral se describen a continuación.

Electroencefalografía (EEG): registra la actividad eléctrica del cerebro a lo largo del cuero cabelludo. El EEG mide las fluctuaciones de voltaje resultantes de los flujos de corriente iónica dentro de las neuronas del cerebro.

Magnetoencefalografía (MEG): es una técnica funcional de neuroimagen para mapear la actividad cerebral mediante el registro de campos magnéticos producidos por corrientes eléctricas que ocurren naturalmente en el cerebro, utilizando magnetómetros muy sensibles.

Imagen de resonancia magnética funcional o resonancia magnética funcional (fMRI): es un procedimiento de neuroimagen funcional que utiliza tecnología de resonancia magnética que mide la actividad cerebral mediante la detección de cambios asociados en el flujo sanguíneo. Esta técnica se basa en el hecho de que el flujo sanguíneo cerebral y la activación neuronal están acoplados. Cuando se usa un área del cerebro, el flujo sanguíneo a esa región también aumenta.

Espectroscopía de infrarrojo cercano funcional (fNIRS): es el uso de NIRS (espectroscopía de infrarrojo cercano) con el propósito de neuroimagen funcional. Usando fNIR, la actividad cerebral se mide a través de respuestas hemodinámicas asociadas con el comportamiento neuronal. FNIR es un método de imagen no invasivo que implica la cuantificación de la concentración de cromóforo resuelta a partir de la medición de la atenuación de la luz del infrarrojo cercano (NIR), cambios temporales o fásicos (AAMODH, 2016).

3.4. Electroencefalografía (EEG)

El cerebro humano genera ondas eléctricas para realizar cualquiera de sus funciones básicas y elementales, así como para recibir, enviar y procesar información. Estas señales pueden ser adquiridas de manera externa y manipuladas a través de un sistema de lecturas electroencefalográficas. Actualmente existen técnicas e interfaces que permiten adquirir las señales EEG del cerebro humano, mediante la cual se realiza un pre procesamiento de ondas para acoplarlas a un dispositivo digital, y a través de herramientas de filtrado, poder determinar qué tipo de señales y sus características específicas genera el cerebro humano al recibir un estímulo visual (imágenes, pictogramas), y verificar si es posible encontrar las características discriminantes asociadas a un determinado concepto utilizado la lectura de estas señales.

Las primeras grabaciones de actividad eléctrica [electroencefalografía (EEG)] del cuero cabelludo humano se obtuvieron a mediados de la década de 1920. EEG proporciona información funcional, en lugar de estructural, del cerebro y tiene importantes aplicaciones en medicina y ciencias cognitivas. El campo de EEG abarca potenciales evocados y relacionados con eventos (ERP por sus siglas en inglés) y EEG espontáneo.

La electroencefalografía humana (EEG) proporciona una "ventana en la mente" conveniente, pero a menudo opaca, que permite observaciones de procesos eléctricos cerca de la superficie del cerebro (Nunez, 2017) obtenidos a través del cráneo por medio de electrodos situados en varios lugares de la superficie del cuero cabelludo. Esta actividad cerebral es medida como fluctuaciones o diferencias de voltaje, resultado de los flujos de corriente iónicos de las neuronas del cerebro (Oñate R. , 2018). El patrón generado de variación en el voltaje a lo largo del tiempo presenta amplitudes que comúnmente suelen situarse entre -100 y +100 microvoltios, con una frecuencia de 40 Hz o más (Aguilar, 2013). La capa externa del cerebro es la corteza cerebral, que se cree que es en gran parte responsable

de nuestros pensamientos, emociones y comportamiento individuales. Los procesos corticales involucran señales eléctricas que cambian con el tiempo en el rango de 0.01 s.

EEG es la única tecnología ampliamente disponible con suficiente resolución temporal para seguir estos cambios dinámicos que ocurren más rápido que la velocidad del pensamiento. Por otro lado, la resolución espacial del EEG es pobre en relación con la de los métodos modernos de imágenes estructurales del cerebro: tomografía computarizada (CT), tomografía por emisión de positrones (PET) e imágenes de resonancia magnética (MRI o fMRI).

El EEG intracraneal a menudo no está correlacionado o solo se correlaciona débilmente con la cognición y el comportamiento y las "medidas mentales" humanas se obtienen más fácilmente a gran escala de las grabaciones del cuero cabelludo. Las limitaciones técnicas y éticas de la grabación intracraneal humana fuerzan el énfasis en las grabaciones del cuero cabelludo. Afortunadamente, estas estimaciones a gran escala proporcionan medidas importantes de la disfunción cerebral para el trabajo clínico y la cognición o el comportamiento para estudios científicos básicos.

EEG monitorea el estado de conciencia de los pacientes en trabajos clínicos o sujetos experimentales en investigación básica. Las oscilaciones del voltaje del cuero cabelludo proporcionan una parte muy limitada pero importante de la historia del funcionamiento del cerebro. En la Fig. 3.5 se proporciona un resumen del EEG clínico y de investigación. Las flechas indican relaciones comunes entre subcampos.

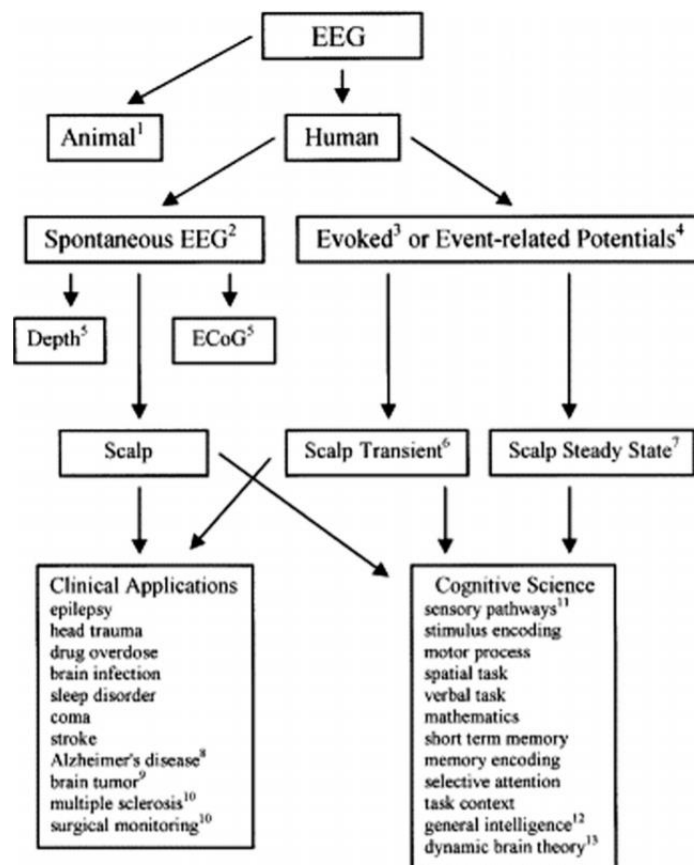


Figura 3.5. Relaciones comunes entre subcampos EEG. Las aplicaciones clínicas están relacionadas principalmente con enfermedades neurológicas. La investigación de EEG es llevada a cabo por neurólogos, neurocientíficos cognitivos, físicos e ingenieros que tienen un interés especial en EEG (Nunez, 2017).

3.4.1. Ventajas e inconvenientes de la Electroencefalografía

Existen otros métodos para el estudio del cerebro, tales como la magnetoencefalografía o la espectroscopia infrarroja; y a pesar de que la electroencefalografía (EEG) pueda parecer un método sensitivamente pobre, presenta múltiples ventajas que la anteponen a otros métodos.

Ventajas:

- Los costos hardware son significativamente más baratos.
- Los sensores EEG pueden ser colocados en mayor número de lugares que otras técnicas, además de permitir movilidad en el sujeto o el equipo de pruebas.
- La Electroencefalografía es silenciosa, lo que posibilita un mejor estudio cuando el sujeto debe experimentar estímulos sensoriales.
- La frecuencia de muestreo es muy superior a las demás, partiendo de 250 Hz y alcanzando los 2000 Hz. En equipos modernos mucho más sofisticados es posible superar los 20,000 Hz incluso.
- La Electroencefalografía también puede detectar procesamiento de información encubierto (que no requiere una respuesta).
- Puede ser implementada con personas que no son capaces de dar respuestas motrices.
- Es extremadamente no invasiva, a diferencia de otras técnicas como la electrocorticografía.

Sin embargo, también cuenta con algunos inconvenientes:

Inconvenientes:

- Baja resolución espacial del cuero cabelludo. Por ejemplo, la imagen por resonancia magnética funcional (fMRI por sus siglas en inglés) muestra directamente áreas activas del cerebro, mientras que EEG requiere de intensa interpretación.
- La actividad neuronal que tiene lugar bajo el cortex se determina con poca precisión.
- En ocasiones, es necesario mayor tiempo de preparación del sujeto, ya que requiere un posicionamiento de los electrodos muy precisos (Aguilar, 2013).

3.4.2. Aplicaciones

El uso básico de la Electroencefalografía es el análisis rutinario encefalográfico, consistente en la grabación de la actividad cerebral los últimos 20-30 minutos y su posterior análisis. Generalmente, con este análisis es posible:

- Distinguir las crisis epilépticas de otros tipos de enfermedades, como crisis psicogénicas no epilépticas, síncope (desmayo), trastornos del movimiento subcortical y variantes de migraña.
- Servir como test de muerte cerebral.
- Pronosticar en pacientes en coma.
- Diferenciar la encefalopatía orgánica de síndromes psiquiátricos como la catatonia.
- Determinar cuándo administrar medicamentos anti-epilépticos.

La EEG puede servir como medio para monitorizar:

- La profundidad de la anestesia.
- De forma indirecta, la perfusión cerebral en la endarterectomía carotídea.
- Efecto amobarbital durante la prueba de Wada.

Otro uso, relativo a unidades de cuidado intensivo y supervisión cerebral:

- Supervisar los ataques no convulsivos epilépticos.
- Supervisar el efecto de sedantes en paciente en coma inducido.
- Supervisar daños cerebrales colaterales en condiciones de hemorragia subaracnoidea (Aguilar, 2013).

3.5. Sistema Internacional 10-20

En el año de 1958 el neurobiólogo Herbert Jasper dirigió la reunión de Sociedades de Electroencefalografía, gracias a la misma se desarrolló el sistema de localización de electrodos. La colocación de los electrodos debe tener una ubicación definida y específica. El esquema que se sigue se llama sistema internacional 10-20, se basa en el procedimiento para ubicar y medir las posiciones a zonas iguales en el cuero cabelludo, utiliza sitios identificables en el cráneo que sirven como punto de referencia. El sistema toma el nombre de 10-20, porque los electrodos están ubicados en espacios del 10 o 20% de la distancia total entre un determinado par de lugares en el cráneo. El sistema 10-20, está diseñado para dar una cobertura adecuada a la totalidad de la cabeza, con la opción para la colocación de electrodos adicionales, dentro de las normas establecidas, empleando la nomenclatura 10-20 (ver Fig. 3.6) (Oñate, 2018).

A la cabeza se la divide en seis zonas: frontopolar (Fp), frontal (F), central (C), parietal (P), occipital (O) y temporal (T). A los electrodos situados en el hemisferio izquierdo se les asigna números impares y a los situados en el hemisferio derecho se les asigna números pares. A los que se encuentran en la línea central se les asigna la letra z minúscula. Los electrodos de referencia suelen colocarse en las orejas (A1, A2) (Martín & Martínez, 2015).

3.6. Potenciales relacionados con eventos

A medida que la información se procesa dentro de las redes neuronales del cerebro humano, la actividad eléctrica que surge de millones de neuronas participantes se suma para formar potenciales de campo que pueden registrarse a través del cuero cabelludo intacto. Los potenciales de campo del cerebro incluyen las oscilaciones de voltaje rítmico del electroencefalograma en curso y los potenciales evocados más breves o relacionados con eventos (ERP) que se activan en asociación con eventos sensoriales, motores y cognitivos. Mientras que los ritmos espontáneos del EEG son monitores sensibles de los estados generales de excitación, conciencia y el ciclo sueño-vigilia, los ERP reflejan patrones más discretos de actividad neuronal que subyacen a procesos perceptivos y cognitivos específicos. Estos patrones de actividad se revelan en el ERP grabado en el cuero cabelludo con un alto grado de precisión temporal (del orden de milisegundos), pero el método generalmente carece de la precisión de la localización anatómica que proporcionan las técnicas de neuroimagen como la tomografía por emisión de positrones o las imágenes por resonancia magnética funcional.

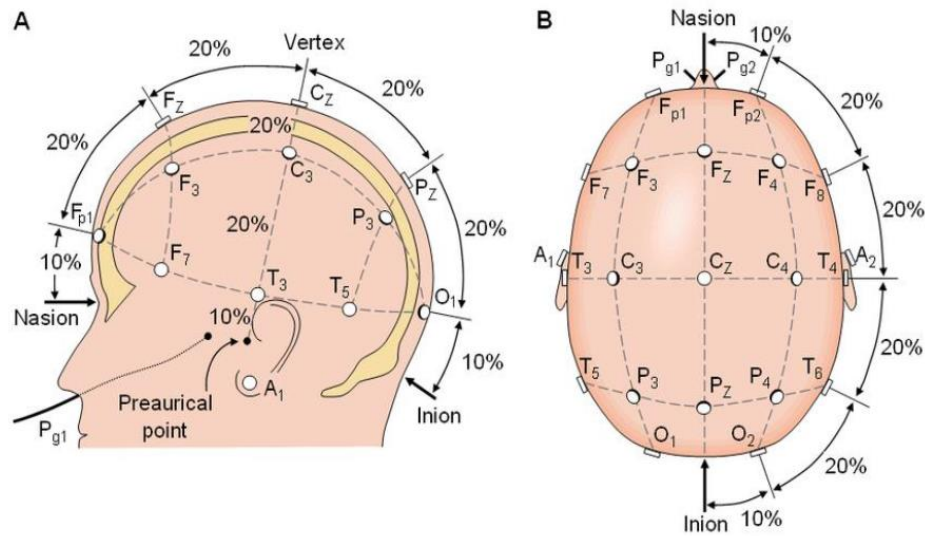


Figura 3.6. Sistema internacional 10-20 de colocación de electrodos (Oñate, 2018).

Sin embargo, los enfoques recientes para modelar los generadores neurales de los ERP han permitido su localización en estructuras cerebrales específicas con mayor precisión. Al mapear la topografía del cuero cabelludo de un ERP y aplicar algoritmos basados en la física de la conducción del volumen eléctrico, las posiciones de los generadores neurales subyacentes se pueden estimar con bastante precisión en relación con la anatomía del cerebro. En varios casos, las regiones anatómicas que supuestamente generan ERP de superficie se han verificado mediante grabaciones de ERP intracraneales, mediante magnetoencefalografía que tiene una mejor resolución espacial o mediante estudios de pacientes con lesiones cerebrales circunscritas. Los estudios combinados de registros de ERP junto con métodos de neuroimagen que revelan patrones de flujo sanguíneo cerebral y metabolismo ofrecen la posibilidad de definir tanto el curso de tiempo preciso como los sustratos anatómicos del procesamiento de información en el cerebro humano.

La metodología de grabación ERP implica el análisis computarizado de las formas de onda de voltaje-tiempo registradas desde diferentes puntos en el cuero cabelludo. En general, es necesario promediar las formas de onda ERP en muchas presentaciones de estímulo para lograr relaciones adecuadas de señal a ruido con respecto al EEG en curso. Los nuevos procedimientos de filtrado espacio-temporal (por ejemplo Filtros wavelet) también pueden ayudar a distinguir las señales débiles de ERP del EEG de fondo.

Los ERP promediados por computadora activados por señales sensoriales consisten en una secuencia de desviaciones (componentes) de voltaje positivo (P) y negativo (N) que tienen retardos de tiempo y formas de onda características. El ERP típico provocado por un breve estímulo auditivo, por ejemplo, consta de 16 o más componentes que aparecen en el intervalo de 1,5 a 1000 ms después del estímulo (Fig. 3.7). Muchos de los componentes posteriores a los 50 ms son muy sensibles a la importancia de un estímulo y a las demandas cognitivas de la tarea. Estos ERP "endógenos" se han estudiado ampliamente como índices de procesos perceptivos y cognitivos (Hillyard, 2017).

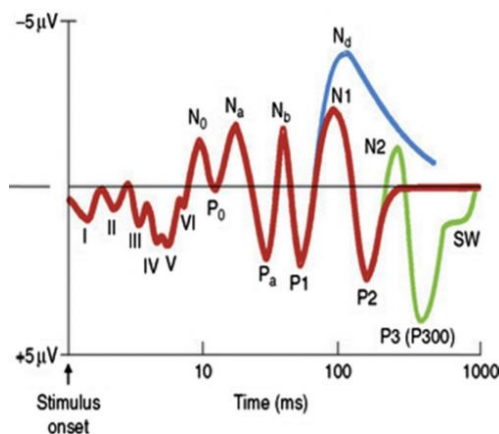


Figura 3.7. Forma de onda del ERP auditivo promediado por computadora a un breve sonido. La visualización del tiempo logarítmico permite la visualización de las respuestas tempranas del tronco encefálico (ondas I – VI), componentes evocados intermedios y componentes endógenos de latencia larga (Nd, N2, P300 y SW, u onda lenta). Los componentes de ERP se etiquetan de acuerdo con su polaridad, P (positiva) o N (negativa), seguidos de un número o letra (P3, Pa, etc.) que especifica la posición ordinal de la onda en una secuencia particular o por un número más grande (P300, etc.) que se refiere a la latencia de la onda en milisegundo (Hillyard, 2017).

Los estímulos visuales y somatosensoriales provocan ERP con formas de onda características y topografías del cuero cabelludo que corresponden al procesamiento sensorial en sus respectivas áreas corticales específicas de modalidad. El ERP visual para un breve destello, por ejemplo, comienza con un componente "C1" a 50–60 ms que parece originarse en la corteza visual (estriada) primaria, seguido de un componente "P1" (a 80–120 ms) que se genera en áreas corticales extraestriadas de las corrientes visuales dorsales y ventrales.

Una estrategia de investigación importante es identificar componentes individuales de ERP como marcadores de procesos perceptivos o cognitivos específicos. Esto se ha logrado en varios casos mediante el estudio sistemático de las variables de estímulo, tarea y comportamiento conductual que influyen en el componente en cuestión. Estos marcadores ERP revelan el momento y el orden de las sucesivas etapas de procesamiento de información y, por lo tanto, pueden revelar la estructura subyacente de las operaciones cognitivas. De esta manera, los ERP producen información sobre las bases neuronales de los procesos cognitivos, incluida la atención selectiva, el reconocimiento de estímulos, la toma de decisiones, el almacenamiento y la recuperación de la memoria y el procesamiento del lenguaje. Como signos objetivos de las operaciones cognitivas, las formas de onda ERP anormales tienen importancia diagnóstica para los trastornos del procesamiento de la información asociados con los síndromes neurológicos, psiquiátricos y del desarrollo.

Las grabaciones de ERP también han ayudado a aclarar los mecanismos de búsqueda visual en tareas en las que los sujetos intentan identificar un estímulo objetivo en una pantalla visual. La detección de dicho objetivo desencadena un componente ERP conocido como N2pc, que se propone para reflejar el enfoque de la atención en la ubicación del objetivo y el filtrado de la información interferente de los elementos distractores cercanos. Cuando hay múltiples elementos de destino en una pantalla, la sincronización de los cambios de amplitud de N2pc brinda soporte a los modelos en serie en lugar de modelos paralelos de búsqueda visual.

La asignación de atención en una pantalla visual también se puede estudiar mediante el potencial evocado visual en estado estacionario (SSVEP), que es la respuesta oscilatoria de la corteza visual a un estímulo repetitivo que oscila entre 6 y 30 Hz. El SSVEP aumenta sustancialmente su amplitud cuando la atención se dirige a la ubicación de un estímulo

parpadeante (Fig. 3.8). Por lo tanto, cuando hay múltiples estímulos en una pantalla, cada uno parpadeando a una velocidad diferente, las amplitudes relativas de los SSVEP con etiquetas de frecuencia pueden indicar hacia dónde se dirige la atención. Los estudios de SSVEP han demostrado que el foco de atención espacial se puede dividir de manera efectiva, de modo que dos regiones separadas de una pantalla visual pueden ser atendidas simultáneamente. Existen trabajos donde analizan este tipo de respuestas para sistemas BCI (Hillyard, 2017). Los estímulos visuales provocan este tipo de respuestas en forma de diferencias de potencial en la zona de la corteza visual del cerebro (área occipital). Estas respuestas pueden ser registradas mediante EEG, y dependen de la integridad funcional de las distintas etapas del tracto visual primario, incluyendo el ojo, la retina, el nervio óptico, las radiaciones ópticas y la corteza visual (Pérez, Minguión, & López, 2018).

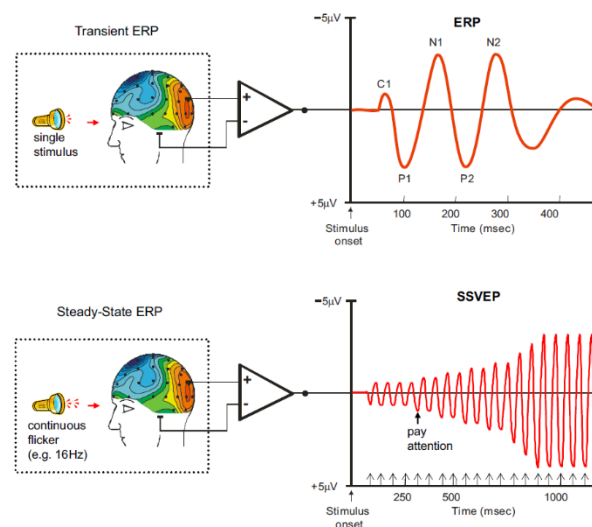


Figura 3.8. Comparación de la forma de onda típica del ERP visual con un estímulo transitorio (arriba) con el potencial evocado visual en estado estacionario (SSVEP) provocado por un estímulo parpadeante repetitivo (abajo). En este ejemplo, el SSVEP es provocado por un estímulo visual que parpadea a 16 Hz, y el SSVEP aparece como una respuesta sinusoidal de la corteza visual a la misma frecuencia. Siguiendo una señal para atender el estímulo parpadeante, se ve que el SSVEP crece en amplitud, y el curso del tiempo de este crecimiento proporciona una estimación de la velocidad a la que se dedica la atención (Hillyard, 2017).

3.6.1. Ondas P3 o P300

Un potencial relacionado con eventos (ERP) es una respuesta cerebral debido a algún estímulo específico percibido por el usuario de BCI. Un ERP típico utilizado para el diseño de BCI es el P300. La onda P3 es un componente endógeno ERP provocado en el curso del proceso de toma de decisiones. La onda P3 se identifica como una desviación positiva aproximadamente 300 ms después del inicio del estímulo. Una muestra de onda P3 se muestra en la Fig. 3.9. Se obtiene cuando se cruza el objetivo con un estímulo objetivo de baja probabilidad entre el estímulo no objetivo de alta probabilidad. Para obtener la onda P3, el método más común y validado es implementar el paradigma impar, en el que los elementos objetivo de baja probabilidad se mezclan con elementos no objetivo de alta probabilidad. Tenga en cuenta que todos los estímulos deben ser aleatorios e impredecibles en el tiempo.

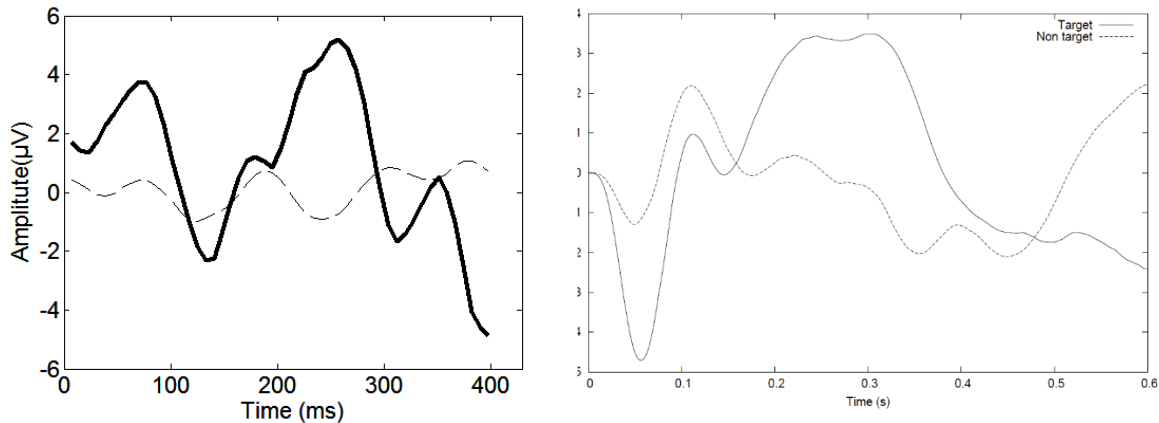


Figura 3.9. La onda P3 (- - no objetivo, - objetivo) (Fabien L. , 2014), (Apdullah, 2017).

Un ejemplo de un ERP P300 promedio después de un estímulo raro y relevante (Target). Podemos observar claramente el aumento de la amplitud unos 300 ms después del estímulo, en comparación con el estímulo no relevante (no objetivo) (Apdullah, 2017).

La onda P3 es una desviación positiva en el voltaje y puede evocarse aproximadamente de 250 a 500 ms. después del inicio del estímulo. Mientras que la latencia de la onda P3 provocada por estímulos visuales es de 350 a 450 ms, la latencia de la onda P3 provocada por estímulos auditivos o táctiles es de 150 a 300 ms. La onda P3 se usa ampliamente en sistemas BCI; control del mouse, control de brazo robótico, detección de objetos visuales. La onda P3 ha demostrado ser una respuesta confiable, para controlar el sistema BCI para personas con trastornos neurológicos.

Los ERP se caracterizan por variaciones temporales específicas con respecto al inicio del estímulo. Como tal, al contrario de BCI basado en la actividad oscilatoria, BCI basado en ERP explota principalmente una información temporal, pero rara vez una espectral. Sin embargo, en cuanto a BCI basado en actividad oscilatoria, basado en ERP también puede beneficiarse mucho al usar la información espacial. En el BCI basado en P300, la información espacial generalmente se explota al enfocarse principalmente en los electrodos ubicados sobre el lóbulo parietal (es decir, extrayendo características solo para estos electrodos), donde se sabe que el P300 se origina. Como ejemplo, en (Fabien L. , 2014) recomiendan usar un conjunto de 8 canales, en las posiciones Fz, Cz, P3, Pz, P4, PO7, Oz, PO8 (ver Fig. 3.10).

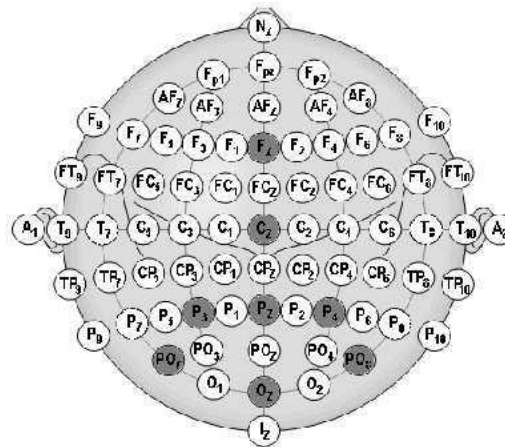


Figura 3.10. Electrodo comúnmente usados para detectar los potenciales P300 (Fabien L. , 2014).

3.7. Interfaz cerebro-computadora (BCI)

Desde finales del siglo XX, las señales eléctricas generadas por el cerebro han sido utilizadas como método de control en las BCI. Las BCI son un grupo de dispositivos que constituyen un canal de comunicación basado en la actividad neuronal y que, por tanto, no precisan del uso de ningún tipo de musculatura (Pérez, Minguillón, & López, 2018).

Es un canal de comunicación que conecta el cerebro con un ordenador u otro dispositivo externo, de esta forma una persona puede comunicarse utilizando el cerebro como única interfaz (Prieto, Pelayo, López, & Romero, 2013).

La tecnología de Interfaz Cerebro-Computadora (BCI por sus siglas en inglés) es un sistema que traduce los reflejos de las actividades cerebrales, las cuales están dadas como estímulos sensoriales ya sea por medios auditivos, visuales, o tareas motoras imaginarias; para ser analizados e interactuar con hardware y software desde la acción real y directa del usuario, y de esta manera llevarlos a aplicaciones tangibles como lo son prótesis, sillas de ruedas, y en general biomecanismos con el objetivo de asistir de manera óptima a los usuarios (Oñate R. , 2018).

Las BCI inicialmente se orientaron a ayudar, reparar o amplificar las funciones sensoriomotoras o cognitivas humanas. Frente a otros sistemas existentes, en el campo de la interacción hombre-máquina, las IHC ofrecen dos características únicas:

- Permiten establecer un canal de comunicación natural con el hombre, dado que traduce las intenciones directamente en órdenes.
El resto de los interfaces no son naturales ya que el pensamiento debe de ser traducido de forma que se adapte al tipo de interfaz, como cuando usamos un teclado, el pensamiento de escribir una letra se traduce en la pulsación de uno de los dedos sobre una tecla.
- Tienen la capacidad de acceso a la información cognitiva y emocional del usuario del sistema.

La tarea principal de una BCI es permitir al sujeto que la utiliza interactuar con el dispositivo gracias a su respuesta electrofisiológica ante un estímulo. Se pueden utilizar

distintos canales para llevar a cabo la estimulación y diversos métodos para registrar la actividad neuronal. Tras la estimulación, se recoge la respuesta electrofisiológica producida en el sujeto y esta es procesada para su posterior análisis. A continuación, se extraen las características más relevantes de la misma y finalmente se realiza una clasificación de la respuesta que resulta en una interacción del sujeto con el dispositivo. Se crea, por tanto, un ciclo cerrado de comunicación que comienza con la estimulación del sujeto y termina con una acción sobre el dispositivo, que viene determinada por su respuesta electrofisiológica. De esta forma, se replica el funcionamiento del sistema nervioso, en el cuál se realiza una acción cuando el córtex motor se comunica con los músculos y fibras del cuerpo a través de impulsos nerviosos como resultado de la respuesta electrofisiológica ante un estímulo, creando un ciclo comunicativo cerrado. Entre las respuestas electrofisiológicas más utilizadas en los sistemas BCI, se encuentran los potenciales evocados visuales (PEV). Los estímulos visuales provocan este tipo de respuestas en forma de diferencias de potencial en la zona de la corteza visual del cerebro (área occipital) (Pérez, Minguillón, & López, 2018).

Podemos decir que principal objetivo de estos sistemas dar un uso a esta tecnología de manera asistida para aquellas personas discapacitadas o en situación de dependencia (Martín & Martínez, 2015).

3.7.1. Elementos de una BCI

Las funciones básicas de una BCI son medir la actividad cerebral, procesarla para obtener características de interés, y, una vez obtenidas interactuar con el entorno de la forma deseada por el usuario (Pérez, Minguillón, & López, 2018).

1. **Adquisición de señal:** Su objetivo es el registro de la actividad cerebral y su adecuación (digitalización) para la siguiente etapa de procesado de la señal. Se trata, pues, de la captura de las ondas cerebrales gracias a los sensores (en este caso, electrodos en el cuero cabelludo o microelectrodos implantados en la superficie del córtex). Aunque no es necesario que los sistemas BCI almacenen la señal registrada, casi todos implementan tal función con el objetivo de permitir posteriormente realizar diferentes análisis o procesados diferentes de señal (variando el algoritmo de procesado).
2. **Procesado de señal:** Este bloque funcional recibe la señal digitalizada generada por la actividad cerebral y la transforma en comandos entendibles por el dispositivo del usuario. Este proceso lo lleva a cabo en tres etapas secuenciales:
 - a) **Cancelación de ruido:** En esta fase se elimina el posible ruido recibido con la señal, fuente de otro tipo de actividad bioeléctrica como pueda ser el movimiento ocular. Esta fase no suelen implementarla muchos de los sistemas BCI; algunos la incluyen en la fase segunda.
 - b) **Obtención de características:** En esta fase se traduce la señal cerebral en un vector de características en correlación con el fenómeno neurológico asociado a la señal. Pueden ser vectores numéricos, donde cada valor se asocie a un nivel de onda.

- c) Traducción de características: En esta fase el vector de características se transforma en una señal de control adecuada al dispositivo. Cuando la señal de control generada es un valor discreto, se habla de clasificación de la característica.
3. **Aplicación:** Este bloque funcional recibe los comandos de control y realiza las acciones correspondientes en el dispositivo a través del controlador. La señal procesada es expandida o transformada a través del interfaz de control y también puede incorporarse una pantalla que proporcione retroalimentación al usuario. Es el módulo encargado de materializar la interacción del usuario con el entorno. Puede ser escribir en una pantalla de ordenador, controlar un brazo robótico, etc
4. **Configuración:** Este último bloque funcional permite definir los parámetros del sistema. El operador que lo lleve a cabo no tiene por qué ser un operador técnico, puede ser el propio usuario del sistema; en el mejor de los casos esta labor la realizarán autómatas informáticos que ajusten el comportamiento del sistema en función de los resultados y la retroalimentación del usuario (Aguilar, 2013); (Pérez, Minguillón, & López, 2018).

3.7.2. Tipos de interfaces BCI

Las interfaces cerebrales se clasifican en torno a tres grandes grupos, los cuales están compuestos de varios tipos de interfaces cada uno. Por un lado, tenemos las Interfaces Cerebrales Invasivas, en las cuales los electrodos han de colocarse directamente en el cerebro, en la materia gris, mediante neurocirugía. Estas interfaces ofrecen la mejor calidad de señal. Por otro lado, están las Interfaces Parcialmente Invasivas, cuyos sensores se implantan aún dentro de la cabeza, pero fuera del cerebro en lugar de la materia gris. Por último, tenemos las Interfaces Cerebrales No Invasivas, que ubican los sensores fuera de la cabeza, sobre la superficie del cuero cabelludo. Acorde a esta clasificación general, se analizan los diferentes tipos de interfaces dentro de cada una de ellas.

Métodos invasivos

Con los sistemas invasivos se introducen sensores en el cerebro capaces de capturar las señales cerebrales. La relación señal-ruido que se obtiene y la intensidad de las señales es mucho mejor que en los sistemas no invasivos, que utilizan dispositivos superficiales para capturar estas señales. El grado de especificidad o resolución espacial y temporal que proporcionan las BCI depende en gran medida de su carácter invasivo. La mayoría de las investigaciones que se realizan en este ámbito están enfocadas a sistemas para personas con discapacidad motora severa o para la recuperación de la visión en personas ciegas. Los sistemas centrados en neuroprótesis permiten mejorar la calidad de vida de personas con parálisis, proporcionándoles dispositivos que les sirven de ayuda, como sillas de ruedas controladas por impulsos cerebrales o brazos robóticos. Se implantan directamente en la materia gris del cerebro mediante cirugía siendo los que obtienen el mejor nivel de señal registrada (Ver Fig. 3.11) (Morán, 2015).

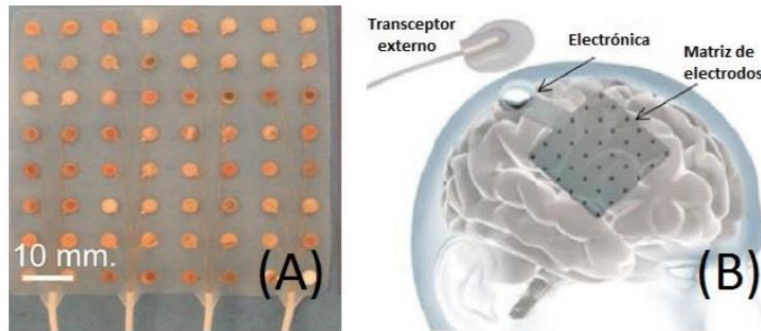


Figura 3.11. (A) Ejemplo de matriz de electrodos utilizada para ECoG y (B) su disposición sobre el cerebro (Schalk & Leuthardt, 2011)

Métodos parcialmente invasivos

Dentro de estas interfaces encontramos los siguientes tipos:

Electrocorticografía: Mide la actividad eléctrica del cerebro de una forma similar a la Electroencefalografía, pero los electrodos se encuentran en un relleno protector situados sobre el córtex (ver Fig. 3.12). Se utilizó en humanos por vez primera en 2004, obteniendo muy buenos resultados: el control es rápido, requiere muy poco entrenamiento y ofrece una buena relación entre calidad de la señal y nivel de invasión en el usuario. La señal puede ser subdural o epidural.

La Electrocorticografía es prometedora debido a su gran resolución espacial, mejor relación señal-ruido que otras interfaces, amplio rango de frecuencia y la menor necesidad de entrenamiento por parte del usuario; al mismo tiempo, la dificultad técnica es menor, hay menos riesgo clínico que las totalmente invasivas y proporcionan una mayor estabilidad.

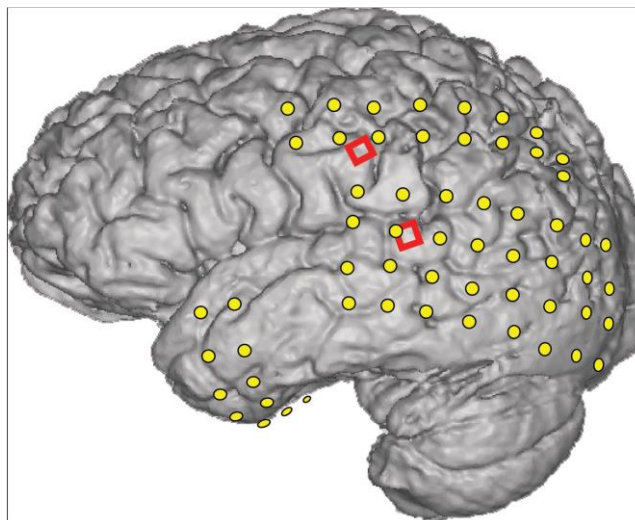


Figura 3.12. Posicionamiento de los electrodos sobre el córtex en la Electrocorticografía.

Imagen reactiva a la luz: Estas interfaces aún son teóricas, es decir, no se han desarrollado. Pretenden implantar un láser dentro del cráneo que pueda ser entrenado sobre una neurona y capture la reflexión del rayo con otro sensor. Cuando la neurona se active, el patrón de luz láser y longitudes de onda variarán, permitiendo a los investigadores supervisar neuronas individuales con menos contacto con el tejido y menor riesgo de producir cicatrices en el tejido.

Métodos no invasivos

Este tipo de métodos limitan el impacto de la adquisición de la señal sobre el usuario realizando una toma de señales desde el exterior de la cabeza del usuario. Uno de los métodos no invasivos más utilizados para la implementación de interfaces cerebro-máquina es la electroencefalografía (EGG), consistente en la colocación de electrodos sobre el cuero cabelludo con ayuda de un gel o líquido conductor (ver Fig. 3.13), lo que no proporciona una resolución espacial muy elevada ya que las medidas que se toman corresponderán a la actividad eléctrica de una zona de la superficie del cerebro relativamente grande (Oñate R. , 2018).



Figura 3.13. El casco Emotiv EPOC+ es un ejemplo de instrumento no invasivo para tomar datos (Emotiv, 2020).

3.8. Sistemas aumentativos y alternativos de comunicación (saac)

Los sistemas aumentativos y alternativos de comunicación (SAAC) son formas de expresión distintas al lenguaje hablado, y su objetivo principal es aumentar (aumentativos) y/o compensar (alternativos) las dificultades de comunicación y lenguaje de muchas personas con discapacidad. Por esta razón, todas las personas que por cualquier causa no han adquirido o han perdido un nivel de habla suficiente para comunicarse de forma satisfactoria, necesitan usar un SAAC.

Las principales causas de la falta del nivel suficiente del habla son las enfermedades neurológicas tales como la esclerosis lateral amiotrófica (ELA), la esclerosis múltiple (EM) o el párkinson, la parálisis cerebral (PC), los trastornos del espectro autista (TEA), la discapacidad intelectual, las distrofias musculares, los traumatismos cráneo-encefálicos, las afasias o las pluridiscapacidades de tipologías diversas, entre otras.

Lo más recomendable es introducir este tipo de comunicación a edades tempranas, tan pronto como se observan dificultades en el desarrollo del lenguaje oral, o poco tiempo después de que cualquier accidente o enfermedad que provoque su deterioro.

La Comunicación Aumentativa y Alternativa (CAA) incluye diversos sistemas de símbolos, tanto gráficos (fotografías, dibujos, pictogramas, palabras o letras) como gestuales (mímica, gestos o signos manuales) y, en el caso de los primeros, requiere también el uso de

otros sistemas de apoyo. Los diversos sistemas de símbolos se adaptan a las necesidades de personas con edades y habilidades motrices, cognitivas y lingüísticas en diferentes niveles (ARASAAC, 2014).

3.9. Los sistemas de símbolos

Los sistemas de símbolos para la CAA se dividen en gestuales y gráficos. En ambos casos existe una graduación desde sistemas muy sencillos, que se adaptan a personas con déficits cognitivos y lingüísticos de diversa consideración, hasta sistemas complejos que permiten niveles avanzados de lenguaje signado (basado en signos manuales) o asistido (basado en signos gráficos).

En el caso de los símbolos gestuales, esta gradación va desde el uso de mímica y gestos de uso común hasta el uso de signos manuales, generalmente en el orden correspondiente al lenguaje hablado; es lo que se denomina lenguaje signado o bimodal. Las lenguas de signos utilizadas por las personas no oyentes no se consideran SAAC, ya que constituyen idiomas que se han desarrollado y se adquieren de forma natural, al igual que ocurre con el lenguaje hablado. El uso de signos manuales requiere disponer de habilidades motrices suficientes, como puede ser el caso de personas con discapacidad intelectual o TEA.

Los símbolos gráficos abarcan desde sistemas muy sencillos basados en dibujos o fotografías, hasta sistemas progresivamente más complejos como los sistemas pictográficos o la ortografía tradicional (letras, palabras y frases). Gracias a los sistemas de apoyo para la comunicación y los diversos recursos para el acceso, los sistemas gráficos pueden ser utilizados por personas con movilidad reducida, incluso en casos de extrema gravedad. Es por esto que, además de ser usados por personas con discapacidad intelectual o TEA, es posible que lo utilicen también personas con discapacidades motoras (PC, ELA, EM, etc.) (ARASAAC, 2014).

3.9.1. Los sistemas pictográficos de comunicación

Los sistemas pictográficos se aplican a personas que no están alfabetizadas a causa de la edad o la discapacidad. Tienen la ventaja de permitir desde un nivel de comunicación muy básico, que se adapta a personas con niveles cognitivos bajos o etapas muy tempranas, hasta un nivel de comunicación avanzado, aunque nunca tan completo y flexible como el que se puede alcanzar con el uso de la lengua escrita. Los sistemas pictográficos más usados en diversos territorios son el sistema SPC (Sistema Pictográfico de Comunicación) (ARASAAC, 2014). Estos fueron creados en 1981 por Roxana Mayer - Johnson. Alrededor de 5000 pictogramas en blanco y negro, y además en imágenes a todo color, fueron recogidos en un libro en el cual, estaba acompañado por la representación de estas imágenes en diferentes idiomas y aspectos ortográficos. Esa gran cantidad de imágenes fueron recogidas en ocho categorías: verbos, descripciones, alimentos, ocio, nombres, cosas misceláneas, personas y lo social. A pesar de que implica categorías bastantes generales que engloban miles de imágenes, es uno de los sistemas pictográficos más utilizados en el mundo gracias a su “alta definición visual” con respecto a las imágenes. Los sistemas pictográficos de comunicación representan todo tipo de conceptos, acciones (individuales o grupales) u objetos en forma de dibujos, utilizando un lenguaje visual que se ve reflejado a través de imágenes muy simples para un mejor entendimiento de lo que se expone ante ellos (Regis & Callejón, 2015).

3.10. Preprocesamiento de las señales EEG

Los análisis de señales EEG pueden realizarse en el dominio del tiempo, dominio de la frecuencia, tiempo-frecuencia y dominio tiempo-escala. En el dominio del tiempo se estudia el cambio de las señales en el tiempo para obtener una función que aisle estos elementos en el tiempo, mientras que en el dominio de la frecuencia se estudia cuántas señales simples componen dichas señales, utilizando sinusoides para aislar componentes de frecuencia de la señal, se puede afirmar que las representaciones en tiempo y en frecuencia contienen información similar pero difieren en las características que son acentuadas en cada dominio y debido a que un método no provee información de frecuencia y el otro solo retorna información de elementos en el tiempo al aplicar ventanas a las señales, uniendo los extremos para analizarlas como ondas continuas, para trabajar con señales no estacionarias se han preferido los métodos tiempo-frecuencia, obteniendo información útil de ambos dominios. Existe una clara necesidad de utilizar este tipo de técnicas, ya que en el análisis de Fourier clásico las señales se consideran infinitas en el tiempo o periódicas, cuando en la realidad su duración puede ser corta y tanto su longitud como su amplitud suelen variar continuamente, una de las formas básicas para aplicarlo es la short-time Fourier Transform (STFT Por sus siglas en inglés), pero con el tiempo se han desarrollado técnicas más sofisticadas como los wavelets, transformada de Hilbert-Huang, etc (Leonardo, 2016).

En los preprocesamientos, los objetivos principales son proporcionar eliminación de ruido y mejora de la señal. Las señales biomédicas siempre tienen componentes de ruido basados en fisiología. Los preprocesamientos son tan cruciales que afectan mucho el modelo de reconocimiento de patrones. Si se produce una implementación incorrecta, el paso restante nunca puede dar como resultado un rendimiento adecuado (Apdullah, 2017).

3.10.1. Artefactos EEG

La actividad del cerebro se registra mediante EEG. La actividad eléctrica registrada que no se origina en el cerebro se conoce como artefacto EEG. Se divide principalmente en artefactos fisiológicos y no fisiológicos. Un artefacto se considera una alteración en una señal cerebral medida. Se consideran señales no deseadas o interferencia en una señal. Los artefactos fisiológicos se generan a partir de regiones distintas al cerebro (es decir, el cuerpo), mientras que los artefactos no fisiológicos surgen del exterior del cuerpo.

Hay muchos tipos diferentes de artefactos que pueden contaminar los datos de EEG, agregando ruido o confundiendo las comparaciones entre las condiciones (ver Figura 3.14). Los artefactos más comunes y problemáticos son típicamente el parpadeo y los movimientos sacádicos del ojo, que se derivan del fuerte potencial corneal-retiniano dentro de cada ojo). Otros artefactos comunes incluyen la actividad muscular, cambios repentinos en el potencial causados por movimientos, cambios graduales en el voltaje causados por los potenciales cutáneos, el electrocardiograma y el bloqueo (saturación) del amplificador o convertidor analógico-digital. Como se presenta en la Fig. 3.14, cada tipo de artefacto tiene una forma de

onda distintiva. Algunos artefactos también tienen distintas distribuciones en el cuero cabelludo (Lopez & J.Luck, 2014).

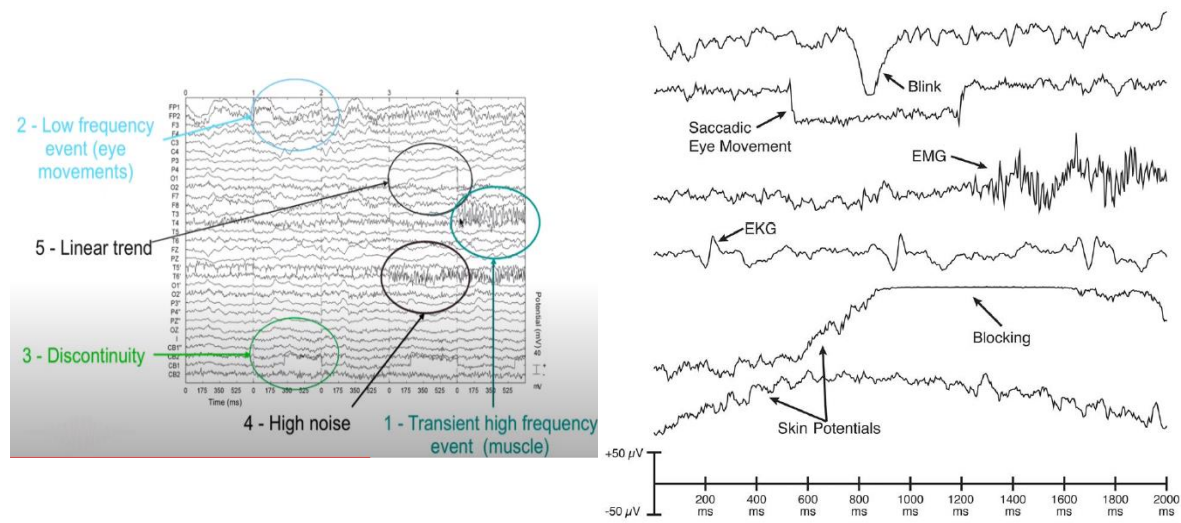


Figura 3.14. Ejemplos de formas de onda de prueba única que muestran artefactos comunes. Cada artefacto tiene un conjunto interactivo de propiedades y, por lo tanto, puede detectarse de forma precisa con los algoritmos reales que se ajustan a estas propiedades (Delorme, EEGLAB, 2018), (Lopez & J.Luck, 2014).

Los artefactos pueden afectar muchas de las características del EEG, como la media, la mediana, la distribución, la desviación estándar y la relación señal / ruido. La adquisición de señales del cuero cabelludo contiene fuentes de ruido que pueden causar problemas al analizar, mostrar y restaurar las señales de EEG. El ruido en la señal produce errores que conducen a una mala interpretación de las señales cerebrales.

De ahí la necesidad de pretratamiento. En los casos más simples, se utilizan métodos de eliminación de artefactos convencionales, como paso bajo, paso de banda o paso alto. Pero estos métodos no siempre son efectivos en el caso en que la frecuencia de la señal y el ruido se superponen. Se han propuesto técnicas alternativas para proporcionar una respuesta a este problema, como el filtrado adaptativo y la transformada wavelet (Bacary, Alice, & Amine, 2020).

Los movimientos de la cabeza pueden introducir una amplia gama de actividad eléctrica no cerebral en el EEG. Típicamente, estos movimientos resultan en contaminación en forma de alguna combinación de músculo (EMG), electrodo, movimiento y artefactos oculares. Estos componentes de las señales de artefactos muestran una amplia gama de características. Los artefactos musculares son predominantemente señales de alta frecuencia, y pueden variar de baja a alta amplitud. El estallido de electrodos suele ir acompañado de picos rápidos de gran amplitud. El movimiento del electrodo tradicionalmente resulta en desviaciones de la línea de base de onda lenta, pero a veces puede producir una oscilación aparente en el EEG. Los artefactos oculares generalmente producen desviaciones de gran amplitud en el EEG (Simon, Stephen, & William, 2010).

Si bien estos componentes de las señales de artefactos exhiben diversas características temporales, de frecuencia y estructurales, son significativamente diferentes de las de la actividad EEG normal. Por lo tanto, la eliminación o separación de estos artefactos reduciría

significativamente la complejidad del sistema y mejoraría las velocidades de cálculo para permitir una implementación en tiempo real del sistema.

Algunas de las formas de minimizar los artefactos son el diseño adecuado de la prueba de EEG, la instrucción adecuada antes de realizar la prueba, EMG de respuesta o agarres de fuerza, rastreador ocular, equipo de localización de electrodos, entorno confortable y buen dispositivo de respuesta.

Existen varios métodos para eliminar los artefactos EEG. Estos se utilizan principalmente para aplicaciones de diagnóstico clínico, investigación e interfaz cerebro-computadora (BCI). Algunos de ellos son análisis de componentes independientes (ICA) y transformaciones de wavelet discretas (DWT). El análisis de componentes independientes como la identificación ciega de segundo orden (SOBI) es el método preciso para corregir el EEG medido.

Filtros Wiener o filtros adaptativos: se utiliza para eliminar artefactos EMG. El filtrado adaptativo se puede utilizar para eliminar artefactos EOG y cualquier ruido de fondo en general. El enfoque más fundamental para la cancelación de ruido en la señal EEG es el filtro Wiener. El filtro Wiener se aplica a un ruido y señales ruidosas. La señal ruidosa es principalmente una combinación de los datos deseados y el ruido. Un filtro Wiener proporciona la salida como la estimación de la señal de ruido que se resta aún más de los datos ruidosos para producir la señal de error.

Análisis de componentes independientes (ICA): se utiliza para eliminar el ruido de la línea eléctrica, así como los artefactos EOG, EMG y ECG. Los potenciales eléctricos del cerebro son las mezclas de algunos componentes subyacentes de la actividad cerebral que pueden analizarse y observarse utilizando ICA. Se extrae información interesante sobre la actividad cerebral a través de ICA.

Otro método utiliza una técnica de eliminación de artefactos. Se ha propuesto que incorpore la transformación wavelet de elevación, en lugar de la transformación wavelet, con la técnica ICA para la eliminación efectiva de artefactos de la señal EEG. Este método proporciona una forma mejor y eficiente de eliminar artefactos que el método tradicional ICA.

En algunos métodos, las señales de EEG se filtraron primero en paso de banda para obtener la banda de frecuencia deseada, seguido de la técnica ICA para eliminar cualquier artefacto no deseado de la señal.

Algoritmo de mínimos cuadrados medios (LMS): el algoritmo de mínimos cuadrados medios es un proceso adaptativo, donde la adaptación se dirige mediante la señal de error entre la señal primaria y la salida del filtro. Este algoritmo se basa en un error cuadrado medio mínimo y, por lo tanto, funciona principalmente con el procesamiento de señales no estacionarias. Cuando hay interferencia en la salida del sistema analizado, este algoritmo obtiene las soluciones óptimas de los problemas de procesamiento de la señal.

Transformada discreta de Fourier (DFT): este algoritmo convierte la señal del dominio del tiempo en el componente del dominio de la frecuencia que establece una correlación entre la representación del dominio del tiempo y la representación del dominio de la frecuencia. Por lo tanto, el análisis de frecuencia de una señal de dominio de tiempo como EEG se realiza

con el algoritmo DFT. La transformada discreta de Fourier puede ser una transformada continua de Fourier para señales conocidas solo en N instantes con una muestra multiplicada por T .

Separación de fuente ciega (BSS): las señales neuronales de los datos de EEG se extraen utilizando este algoritmo. BSS principalmente extrae una serie de vectores de libro de códigos. Estos vectores de libro de códigos representan el contenido espectral subyacente a la señal grabada. Como resultado, se hace posible extraer la actividad rítmica particular subyacente a las grabaciones de señal.

Haar wavelet: esta técnica se utiliza para la detección de artefactos oculares. La aplicación de Haar wavelet a la señal de EEG sin formato ayuda en el desarrollo de un sistema confiable para la detección temprana en la grabación clínica de EEG. La eliminación del ruido de los artefactos oculares de las señales de EEG en función de las técnicas de filtrado disponibles da como resultado la pérdida de datos. Por lo tanto, la wavelet de Haar se utiliza para superar este problema de pérdida de datos (Angana & Sudip, 2019).

Particularmente, para este trabajo de categorización de conceptos, se implementan esquemas de eliminación de artefactos, el cual está basado en el análisis de componentes independientes (ICA) y en el exponente de Hurst. Este método implica calcular los componentes independientes de las señales de EEG y posteriormente, calcular el exponente de Hurst de cada uno de ellos. Por último, se vuelven a combinar los componentes independientes que no fueron eliminados, usando la matriz de mezcla calculada con ICA, para la reconstrucción de las señales de los electrodos pero sin artefactos.

3.10.2. Filtrado de las señales EEG

El filtrado de señal se utiliza para descartar frecuencias no deseadas e irrelevantes de la señal adquirida. Es crucial diseñar correctamente el filtro para mejorar la calidad de la señal y aumentar la relación señal/ruido (SNR por sus siglas en inglés). De lo contrario, los componentes necesarios de las señales reales pueden eliminarse.

Hay tres instrumentos principales de filtrado; filtro de pasa altas (HPF), filtro de pasa bajas (LPF) y filtro pasa bandas (BPF) (por sus siglas en inglés respectivamente). Los filtros se ajustan en hercios (Hz). HPF permite que pasen señales que tienen frecuencias iguales y más altas, mientras que LPF permite que pasen frecuencias iguales y más bajas.

En cuanto a BPF, es una combinación de HBF y LBF. Aunque la onda P3 existe entre 0.1 Hz y 4 Hz), debido a la posibilidad de que exista en frecuencias más altas, se prefieren las frecuencias de corte como más altas que 4 Hz. El filtro de paso de banda de tercer orden butterworth que tiene una frecuencia de corte baja de 0.1 Hz y una frecuencia de corte alta de 20 Hz está diseñado para filtrar datos EEG sin procesar (Apdullah, 2017). La implementación de los filtros tiene distintas aplicaciones, por ejemplo:

- Filtro de paso alto activo: las señales EEG generalmente se mezclan con ruidos de baja frecuencia. Para superar estos ruidos, un filtro de paso alto es muy esencial, ya que elimina la deriva de la línea de base. El filtro de paso alto activo puede diseñarse con una frecuencia de corte de 1 Hz.

- Filtro de paso bajo activo: cuando una señal EEG se altera con ruidos de alta frecuencia, un filtro de paso bajo adecuado la elimina. La frecuencia de corte de este tipo de filtro es de 100 Hz (Angana & Sudip, 2019).

Basado en la investigación del estado del arte, los filtros más implementados para el proceso de categorización de objetos con EEG con el fin de eliminar desviaciones lentas en la señal y ruido de alta frecuencia, son los filtros de respuesta finita (FIR) pasa bandas con frecuencias entre 1 y 100 Hz, con variaciones de uso de las ventanas de Hamming.

En resumen, al diseñar BCI basado en ERP, es importante utilizar la información temporal. Esto se logra principalmente mediante el uso de la amplitud de los puntos de tiempo de EEG preprocesados como características, con filtrado de paso bajo o paso de banda y disminución de resolución como preprocesamiento. Los algoritmos de selección de funciones también pueden resultar útiles. También es importante tener en cuenta la información espacial. Para hacerlo, es útil usar o seleccionar canales relevantes. El uso de algoritmos de filtrado espacial como xDAWN o los filtros espaciales de Fisher también puede ser una solución muy eficiente, particularmente cuando hay pocos datos de entrenamiento disponibles. A continuación, describiremos brevemente algunas herramientas alternativas de procesamiento de señales que son menos utilizadas pero que también pueden resultar útiles en la práctica (Fabien L. , 2014).

3.11. Extracción de características de señales EEG

Con la llegada de nuevos métodos de obtención de imágenes anatómicos y funcionales, ahora es posible recopilar grandes cantidades de datos del cerebro humano vivo. Por lo tanto, se ha vuelto muy importante extraer las características esenciales de los datos para permitir una representación o interpretación más fácil de sus propiedades.

Basado en la investigación de (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007), en la cual menciona que a pesar de algún método específico utilizado para extraer características de las señales EEG, se deben de considerar algunas propiedades críticas con respecto al diseño de la BCI:

- Ruido y los valores atípicos (outliers): las señales EEG tienen una mala relación señal-ruido (SNR);
- Alta dimensión: normalmente se extraen muchas características de varios canales y de varios segmentos de tiempo antes de que se concatenen en un vector de características único.
- Información del tiempo: los patrones de actividad cerebral generalmente se relacionan con las variaciones de tiempo específicos del EEG.
- No estacionarias: las señales EEG presentan una frecuencia variable en el tiempo o su contenido de frecuencia es variable en el tiempo.
- Conjuntos de entrenamiento pequeños: el proceso de entrenamiento es lento y exigente para los sujetos.

Se ha intentado una gran variedad de características para diseñar BCI, como valores de amplitud de señales EEG, potencias de banda (BP), valores de densidad espectral de potencia

(PSD), autorregresivo (AR) y Parámetros adaptativos de autorregresión automática (AAR), características de frecuencia de tiempo y características basadas en modelos inversos (Fabien L. , 2014).

Existen múltiples formas de representación para las señales de EEG, los dos tipos más comunes de características utilizadas para representar las señales de EEG son las características de potencia de banda de frecuencia y las características de punto de tiempo. Las características de potencia de banda representan la potencia (energía) de las señales de EEG para una banda de frecuencia dada en un canal dado, promediada en una ventana de tiempo dada (típicamente 1 segundo para muchos paradigmas BCI). Las características de potencia de la banda se pueden calcular de varias maneras, y se usan ampliamente para los BCI que explotan la actividad oscilatoria, es decir, los cambios en las amplitudes del ritmo EEG.

Las características de punto de tiempo son una concatenación de muestras de EEG de todos los canales. Por lo general, estas características se extraen después de cierto preprocesamiento, especialmente el filtrado de paso de banda o paso bajo y el muestreo descendente. Son las características típicas que se utilizan para clasificar los potenciales relacionados con eventos (ERP), descritas en el apartado 3.6.

Ambos tipos de características se benefician de ser extraídas después del filtrado espacial. El filtrado espacial consiste en combinar las señales originales del sensor, generalmente linealmente, lo que puede dar como resultado una señal con una relación señal / ruido más alta que la de los sensores individuales. El filtrado espacial puede ser independiente de los datos, basado en la consideración física de cómo las señales de EEG viajan a través de la piel y el cráneo, lo que lleva a filtros espaciales como el conocido filtro laplaciano o el filtrado espacial basado en solución inversa. Los filtros espaciales también se pueden obtener de una manera no supervisada y basada en datos con métodos como el análisis de componentes principales (PCA) o el análisis de componentes independientes (ICA). Finalmente, los filtros espaciales se pueden obtener de una manera basada en datos, con aprendizaje supervisado, que actualmente es uno de los enfoques más populares.

Los filtros espaciales supervisados incluyen los conocidos patrones espaciales comunes (CSP), dedicados a las características de potencia de banda y la actividad oscilatoria BCI, y los filtros espaciales como xDAWN o los filtros espaciales de Fisher para la clasificación ERP en función de las características de los puntos de tiempo. Estos filtros han demostrado resultados óptimos en la clasificación de las señales, lo que ha conllevado al desarrollo de múltiples variantes de dichos algoritmos que son más robustos al ruido o las señales no estacionarias, utilizando enfoques de regularización, promedios de datos robustos y / o nuevas medidas de divergencia.

Del mismo modo, se han propuesto extensiones de estos enfoques para optimizar los filtros espectrales y espaciales simultáneamente (por ejemplo, el método CSP (FBCSP) del banco de filtros popular y otros. Finalmente, algunos enfoques han combinado ambos filtros espaciales impulsados físicamente basados en modelos inversos con filtros espaciales controlados por datos.

El filtrado espacial seguido por la potencia de banda o la extracción de características de puntos de tiempo son comúnmente, las características más comunes utilizadas en las BCI

basadas en EEG actuales, es necesario comentar que se han explorado y utilizado otros tipos de características. En primer lugar, un tipo cada vez más utilizado son las características de conectividad. Dichas características miden la correlación o sincronización entre señales de diferentes sensores y / o bandas de frecuencia. Esto se puede medir utilizando características como la coherencia espectral, los valores de bloqueo de fase o las funciones de transferencia dirigida, entre muchas otras.

Finalmente, en lugar de usar vectores de características, la investigación reciente también ha explorado cómo representar las señales de EEG por matrices de covarianza o por tensores (es decir, matrices y matrices multidireccionales, con dos o más dimensiones), y cómo clasificar estas matrices o tensores directamente. En procesos que utilizan descomposiciones de matriz o tensores, las características resultantes son combinaciones lineales de datos, puntos de tiempo o frecuencias de varios sensores (entre otros). Como tales, pueden no tener una interpretación física / fisiológica obvia, pero de todos modos resultan útiles para el diseño de BCI (F, L, A, M, & M, 2018).

Como se mencionó anteriormente, la extracción de características tiene como objetivo representar las señales de EEG sin procesar mediante un número idealmente pequeño de valores relevantes, que describen la información relevante para la tarea contenida en las señales. Sin embargo, los clasificadores pueden aprender de los datos qué clase corresponde a qué características de entrada. Como tal, ¿por qué no usar directamente las señales EEG como entrada al clasificador? Esto se debe a la llamada "maldición de la dimensionalidad", que establece que la cantidad de datos necesarios para describir adecuadamente las diferentes clases aumenta exponencialmente con la dimensionalidad de los vectores de características (Fabien L. , 2014).

Existen muchas herramientas matemáticas para extraer información de las observaciones de EEG. Las distribuciones de frecuencia de tiempo (TFD) caracterizan las señales no estacionarias en un plano de frecuencia de tiempo. También pueden servir como base para la síntesis, codificación y procesamiento de señales. Debido a que las señales EEG tienen propiedades no estacionarias, el análisis de frecuencia de tiempo es una herramienta poderosa para las señales EEG (Abdelhakim & Daoud, 2017).

De manera general, existen 3 fuentes principales de información que se pueden utilizar para extraer características de las señales de EEG:

- **Información espacial:** tales características describirían de dónde (espacialmente) proviene la señal relevante. En la práctica, esto significaría seleccionar canales específicos de EEG o centrarse más en canales específicos que en algún otro. Esto equivale a centrarse en la señal que se origina en áreas específicas del cerebro.
- **Información espectral (frecuencial):** tales características describirían cómo varía la potencia en algunas bandas de frecuencia relevantes. En la práctica, esto significa que las características usarán la potencia en algunas bandas de frecuencia específicas.
- **Información temporal:** tales características describirían cómo la señal relevante varía con el tiempo. En la práctica, esto significa utilizar los valores de las señales de EEG en diferentes puntos de tiempo o en diferentes ventanas de tiempo.

Existen múltiples métodos para calcular las características de potencia de banda a partir de señales EEG. Sin embargo, uno de los más eficientes y utilizados es primero filtrar la señal de EEG de un canal dado en la banda de frecuencia de interés, después cuadrar la señal resultante para calcular la potencia de la señal y finalmente promediarla con el tiempo (por ejemplo, en una ventana de tiempo de 1 s).

Para ERP en general, incluido el P300, las características generalmente explotan la información temporal de las señales, es decir, cómo la amplitud de la señal EEG varía con el tiempo. Esto generalmente se logra utilizando los valores de los puntos de tiempo de EEG preprocesados como características.

Más precisamente, las características para ERP generalmente se extraen mediante 1) paso bajo o paso de banda que filtran las señales (por ejemplo, en 1-12 Hz para el P300), ERP es generalmente ondas lentas, 2) disminución de las señales filtradas, en orden para reducir el número de puntos de tiempo de EEG y, por lo tanto, la dimensionalidad del problema y 3) reunir los valores de los puntos de tiempo de EEG restantes de todos los canales considerados en un vector de características que se utilizará como entrada para un clasificador. Este proceso se ilustra en la Figura 3.15 para extraer características del canal Pz para un experimento BCI basado en P300.

Una vez extraídas las características, se pueden proporcionar a un clasificador que será entrenado para asignarlas a la clase objetivo (presencia de un ERP) o a la clase no objetivo (ausencia de un ERP). Esto a menudo se logra usando clasificadores clásicos como LDA o SVM. En realidad, se necesitan muchos puntos de tiempo para describir ERP, pero solo unas pocas bandas de frecuencia (o solo una) para describir la actividad oscilatoria. Alternativamente, la selección de características o las técnicas de selección de canales también se pueden utilizar para tratar esta alta dimensionalidad.

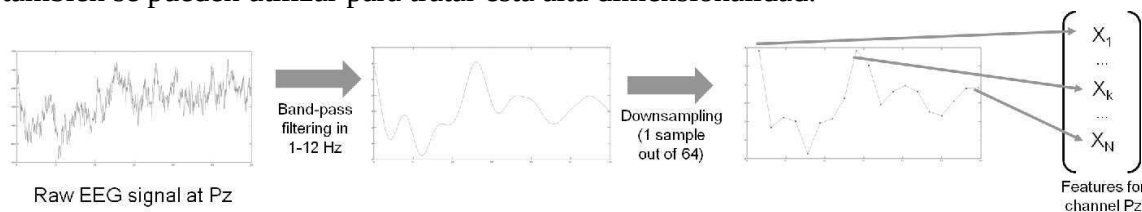


Figura 3.15. Proceso típico para extraer características de un canal de datos EEG para un diseño BCI basado en P300. En esta imagen podemos ver que el P300 se vuelve más visible con los diferentes pasos de procesamiento (Fabien L. , 2014).

Sin embargo, existen otras herramientas de procesamiento de señales y, en particular, otros tipos de características o fuentes de información que pueden explotarse para procesar señales EEG:

Representaciones temporales: las representaciones temporales miden cómo varía la señal con el tiempo. Contrariamente a las características básicas utilizadas para ERP, que simplemente consisten en los puntos de tiempo del EEG a lo largo del tiempo, se han desarrollado algunas medidas para caracterizar y cuantificar esas variaciones. Las características correspondientes incluyen parámetros de Hjorth o parámetros de dominio de tiempo (TDP). Los resultados de investigaciones recientes incluso han sugerido que TDP podría ser más eficiente que las características de potencia de banda estándar de oro.

Medidas de conectividad: miden hasta qué punto la señal de dos canales está correlacionada, sincronizada o incluso si una señal puede ser la causa del otro. En otras palabras, las características de conectividad miden cómo se relacionan las señales de dos canales. Esto es particularmente útil para BCI ya que se sabe que, en el cerebro, hay muchas comunicaciones de larga distancia entre áreas separadas. Como tal, las características de conectividad se utilizan cada vez más para BCI y parecen ser un complemento muy valioso para las características tradicionales. Las características de conectividad incluyen coherencia, valores de bloqueo de fase o función de transferencia dirigida (DFT).

Medidas de complejidad: naturalmente miden cuán compleja puede ser la señal EEG, es decir, miden su regularidad o cuán predecible puede ser. También se ha demostrado que proporciona información sobre el estado mental del usuario y también proporciona información complementaria a las características clásicas, como las características de potencia de banda. Las características de esta categoría utilizadas en BCI incluyen entropía aproximada, complejidad predictiva o longitud de forma de onda.

Medidas inspiradas en la teoría del caos: otra categoría de características que se ha explorado son las medidas relacionadas con el caos, que evalúan cuán caótica puede ser la señal de EEG o qué propiedades caóticas puede tener. También se ha demostrado que esto extrae información relevante. Entre los ejemplos de características correspondientes se incluyen la dimensión fractal o los acumuladores multicractales.

En algunas investigaciones, el objetivo principal es realizar una exploración más completa y sistemática de una gama más amplia de características de EEG (Xiang, et al., 2018).

Teniendo en cuenta que las señales de EEG no son estacionarias y el contenido espectral varía temporal y espacialmente, es necesario utilizar un enfoque de solución múltiple. La transformación Wavelet descompone una señal usando un conjunto de funciones llamadas ondas, que son versiones escaladas y desplazadas de otra función llamada onda madre (Germán, Pressel, Iván, & Leonardo, 2017).

El análisis del dominio del tiempo y la frecuencia extrae principalmente las diversas características de la señal. Estas características se definen como parámetros que proporcionan información sobre las características de la señal particular.

Las características del dominio del tiempo son muy rápidas y fáciles de implementar. No requieren ninguna transformación y, por lo tanto, la complejidad computacional es menor en comparación con las características del dominio de frecuencia. Las características del dominio de frecuencia son propiedades estadísticas de la densidad espectral de potencia de las señales.

Existen varias técnicas de extracción de características.

Transformación de Hilbert-Huang: la transformación de Hilbert-Huang descompone una señal en funciones de modo intrínseco. Principalmente obtiene datos de frecuencia instantánea de señales no estacionarias y no lineales. La transformación de Hilbert-Huang analiza la frecuencia instantánea de las señales de EEG y le da un significado físico a la señal.

Análisis de componentes principales (PCA): PCA es el método de análisis de bioseñales. Principalmente detecta la combinación lineal de un conjunto de variables que tiene la máxima

varianza y finalmente elimina su efecto. Es una herramienta dominante para analizar los datos del EEG y encontrar patrones en ellos. Se puede obtener una tarea diferente a partir de la señal EEG de un sujeto utilizando PCA con red neuronal.

Bases discriminantes locales (LDB): las bases discriminantes locales (LDB) se basan en la descomposición de paquetes wavelet completamente diezmados (WPD). Solo es adecuado para datos estacionarios. Las características discriminantes no son consistentes en la aplicación de la versión desplazada de la misma señal que se debe al WPD diezmado. En su mayoría estiman cada coeficiente individualmente. El LDB depende principalmente del paradigma de mejor base. Utiliza un ajuste de árbol que selecciona una base del diccionario que resalta la discrepancia entre las clases.

Modelo autorregresivo (AR): el modelo AR es la representación de un proceso aleatorio que funciona en ciertos procesos de naturaleza variable en el tiempo. Aquí, la variable de salida depende linealmente de sus propios valores anteriores. El modelado AR es un algoritmo eficiente para la estimación de parámetros. Los modelos autorregresivos se utilizan para el análisis de parámetros espectrales (SPA) del EEG.

Transformación de wavelet discreta (DWT): DWT es una implementación de la transformada de wavelet utilizando un conjunto discreto de escalas de wavelet. En su mayoría obedece algunas reglas predefinidas durante la traducción. En consecuencia, el conjunto de wavelets forma una base ortonormal que puede usarse para descomponer la señal. El coeficiente DWT contiene información temporal de la señal analizada.

Transformación del paquete wavelet: en la descomposición del paquete wavelet, la señal de tiempo discreto pasa a través de varios filtros. Originalmente se conoce como estructura óptima de árbol de sub-banda. Esta transformación ofrece una representación óptima de la señal EEG, cuya estructura proporciona la flexibilidad para las características. Para investigar las características variables en el tiempo de las señales EEG multicanal, la transformación de paquetes wavelet es una herramienta útil. Permite un análisis de subbandas sin la restricción de una descomposición diádica.

Entropía de muestra: la entropía de muestra es una característica modificada de la entropía aproximada que se utiliza para detectar la complejidad de las señales de series temporales fisiológicas y, por lo tanto, diagnosticar estados enfermos. Es la independencia de la longitud de los datos y una implementación relativamente libre de problemas. La entropía de muestra es más adaptativa a la detección en tiempo real y tiene una mejor capacidad para acceder al nivel de conciencia de los pacientes durante la cirugía. El único inconveniente de la entropía de muestra es el valor de entropía calculado que varía de 0 a 3 de manera diversa (Angana & Sudip, 2019).

En la tabla 3.2 se desglosa una descripción de las ventajas y desventajas de algunos métodos de extracción de características que utilizan señal EEG.

Tabla 3.2. Ventajas y desventajas de métodos de extracción de características (Bacary, Alice, & Amine, 2020).

	Ventajas	Desventajas
PCA (análisis de componentes principales)	Reducción de dimensiones sin perder datos.	Incapaz de procesar un conjunto complicado de datos
ICA (análisis de componentes independientes)	Computacionalmente eficiente, alto rendimiento para datos de gran tamaño	Requiere más cálculo para la descomposición
WT (Transformada Wavelet)	Adecuado para señales no estacionarias, capaz de analizar señales en el dominio del tiempo y la frecuencia	Falta de metodología específica para aplicar al ruido generalizado
AR (Auto regresivo)	Requiere una corta duración de los datos, reduce los problemas de pérdida de espectros y brinda una mejor resolución de frecuencia	No aplicable a señales estacionarias.

El procesamiento de señales EEG es una tarea muy difícil, debido al ruido, la no estacionariedad, la complejidad de las señales, así como a la cantidad limitada de datos de entrenamiento disponibles. Como tal, las herramientas existentes aún no son perfectas, y muchos desafíos de investigación aún están abiertos. En particular, es necesario explorar y diseñar características de EEG que sean 1) más informativas, para alcanzar mejores rendimientos, 2) resistentes al ruido y a los artefactos, para usar el BCI fuera de los laboratorios, potencialmente con usuarios en movimiento, 3) invariante, para tratar con la no estacionariedad y la transferencia de sesión a sesión y 4) universal, para diseñar BCI independiente del sujeto, es decir, BCI que puede funcionar para cualquier usuario, sin la necesidad de una calibración individual. Como hemos visto, algunas herramientas existentes pueden abordar parcialmente, o al menos, mitigar tales problemas. Sin embargo, hasta ahora no existe una herramienta de procesamiento de señal EEG que tenga todas estas propiedades simultáneamente y que sea perfectamente robusta, invariante y universal. Por lo tanto, todavía hay trabajos de investigación interesantes por delante (Fabien L. , 2014).

3.12. Clasificación de las señales EEG

Se retoma la definición mencionada en (Alejandro, 2016), la clasificación cubre cualquier contexto en el que alguna decisión o pronóstico es hecho sobre la base de información histórica disponible. Esta base de información disponible D se define de la siguiente forma:

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} = (X, Y)$$

donde los valores $x_1 \in X$ son típicamente vectores multi-dimensionales de la forma: $x_i = \{z_1, z_2, \dots, z_n\}$ cuyos elementos pueden tomar valores reales o discretos. Estos componentes se denominan atributos (o características). El objetivo es inferir la función (o relación) desconocida $y = f(x)$.

$$f(x): X \rightarrow Y.$$

donde los valores de Y están contenidos en un conjunto finito de clases $C = \{C_1, \dots, C_k\}$ que caracterizan los datos dados, en el caso del aprendizaje supervisado. Los modelos aprendidos de los datos de entrenamiento son, entonces, evaluados con un conjunto de prueba distinto para determinar si los modelos pueden ser generalizados a nuevos casos.

Para elegir el clasificador más apropiado para un conjunto dado de características, se deben conocer las propiedades de los clasificadores disponibles. Esta sección proporciona una taxonomía de clasificadores. También se ocupa de dos problemas de clasificación especialmente relevantes para la investigación de BCI, a saber, la maldición de la dimensionalidad y la compensación de Bias-Variance.

Taxonomía de los clasificadores

Varias definiciones se usan comúnmente para describir los diferentes tipos de clasificadores disponibles:

Generativo-discriminativo: Los clasificadores generativos (también conocidos como informativos), por ejemplo, Bayes cuadrático, aprenden los modelos de clase. Para clasificar un vector de características, los clasificadores generativos calculan la probabilidad de cada clase y eligen la más probable. Las discriminatorias, por ejemplo, Máquinas de vectores de soporte, solo aprenden la forma de discriminar las clases o la membresía de la clase para clasificar un vector de características directamente.

Dinámica estática: Los clasificadores estáticos, por ejemplo, los perceptrones multicapa, no pueden tener en cuenta la información temporal durante la clasificación, ya que clasifican un único vector de características. Por el contrario, los clasificadores dinámicos, por ejemplo, el modelo oculto de Markov, pueden clasificar una secuencia de vectores de características y, por lo tanto, capturar la dinámica temporal.

Estable-inestable: Los clasificadores estables, por ejemplo, el Análisis discriminante lineal, tienen una baja complejidad (o capacidad). Se dice que son estables ya que pequeñas variaciones en el conjunto de entrenamiento no afectan considerablemente su rendimiento. Por el contrario, los clasificadores inestables, por ejemplo, el perceptrón de múltiples capas, tienen una alta complejidad. En cuanto a ellos, pequeñas variaciones del conjunto de entrenamiento pueden conducir a cambios importantes en el rendimiento.

Regularizado: La regularización consiste en controlar cuidadosamente la complejidad de un clasificador para evitar el sobreentrenamiento. Un clasificador regularizado tiene buenos resultados de generalización y es más robusto con respecto a los valores atípicos (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).

Técnicas de clasificación de señales aplicadas a señales EEG

En los últimos años, se han llevado a cabo varios estudios de clasificación EEG. La mayoría de estos utilizaron diferentes técnicas de clasificación y compararon su desempeño en términos de precisión de clasificación. A continuación se presenta un listado de los clasificadores más populares y ampliamente utilizados para fines de investigación

Máquina de vectores de soporte (SVM): el método para la técnica de clasificación de objetos, que se basa en la teoría del aprendizaje estadístico, se conoce como máquina de vectores de soporte. Los clasificadores SVM clasifican los datos al encontrar el mejor hiperplano que separa todos los puntos de datos de una clase de los de la otra que aplica el concepto de planos de decisión. El plano de decisión separa un conjunto de características que tienen varias membresías de clase que definen los límites de decisión. El clasificador SVM encuentra el hiperplano que maximiza el margen de separación entre clases (Angana & Sudip, 2019).

Al transformar los datos en un espacio dimensional superior, los problemas de clasificación complejos se pueden convertir en problemas más simples que pueden utilizar funciones discriminantes lineales. Como el SVM se basa solo en aquellos patrones de entrenamiento que están cerca de la superficie de decisión, los tiempos de clasificación son bastante pequeños, lo que permite una implementación sencilla en un sistema en tiempo real. Además, como los vectores de soporte son un subconjunto de los datos de entrenamiento y contienen toda la información necesaria para definir el clasificador, los SVM son altamente insensibles a la dimensionalidad del espacio de características. Esto hace que el SVM sea ideal en el caso de un sistema con una gran cantidad de características extraídas de las señales de EEG.

Los clasificadores SVM de kernel gaussianos han demostrado su eficacia en varios otros sistemas de detección de eventos en EEG. En cada uno de los sistemas clasificadores, se aplica un kernel gaussiano SVM a los datos de características extraídos para cada una de las longitudes de ventana examinadas (Simon, Stephen, & William, 2011).

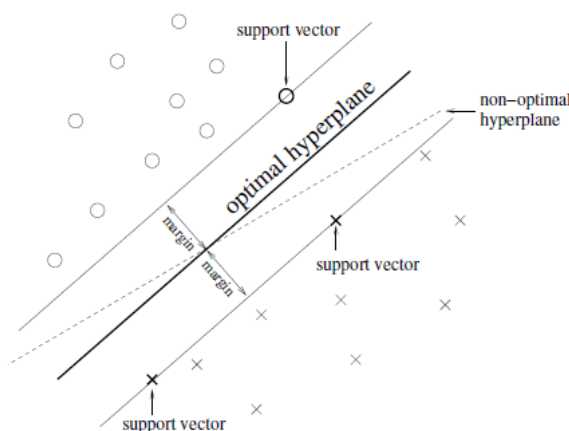


Figura 3.16. El SVM encuentra el hiperplano óptimo para generalización. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007)

K-vecino más cercano (k-NN): k-NN es un clasificador ampliamente utilizado, basado en el aprendizaje. Este método se implementa considerando la similitud de las pruebas y los datos de capacitación. En este proceso de clasificación, una muestra se clasifica a través de un favor mayoritario de sus vecinos. Cuando se proporciona una muestra desconocida, busca en el espacio de características las k muestras de entrenamiento que están más cerca de la muestra desconocida.

La medida de distancia euclidiana simple es el método más utilizado en el vecino más cercano. Considere patrones que tienen n características, $a = [a_1, a_2 \dots a_n]$ y $b = [b_1, b_2 \dots b_n]$, la distancia medida es:

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (3.1)$$

Para etiquetar los datos de la prueba, con mayor frecuencia k las muestras más cercanas se asignan por mayoría de votos (ver Fig. 3.17).

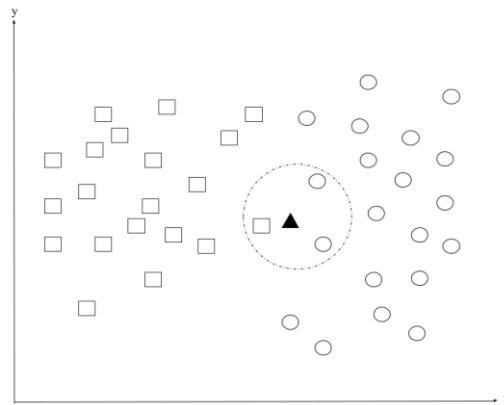


Figura 3.20. El algoritmo k vecino más cercano comienza en los datos de prueba y aumenta la región circular hasta que cubre el número K de datos de entrenamiento. Los datos de prueba se etiquetan por mayoría de votos. Fuente: (Apdullah, 2017).

Análisis discriminante lineal (LDA): LDA es ampliamente utilizado en el campo de la estadística y el aprendizaje automático. Principalmente proporciona más separabilidad de clase. Dibuja una región de decisión entre las clases dadas y busca los ejes del proyecto. En los ejes del proyecto, los puntos de datos de las diferentes clases están lejos unos de otros, mientras que los puntos de datos de la misma clase deben estar cerca uno del otro. LDA crea una combinación lineal de datos que detecta las mayores diferencias de medias entre las clases deseadas (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007).

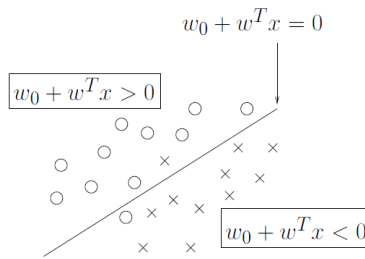


Figura 3.24. Un hiperplano que separa dos clases: los círculos y las cruces. Fuente: (Fabien, Marco, Anatole, Fabrice, & Bruno, 2007)

LDA asume una distribución normal de los datos con la misma matriz de covarianza para ambas clases. El plano de separación es obtenido al buscar la proyección que maximice la distancia entre las dos clases medias y minimice la varianza interclase.

Red neuronal de perceptrón multicapa (MLPNN): el perceptrón multicapa incluye una capa de entrada, un mínimo de una capa oculta y una capa de salida. Cada capa tiene al menos una neurona. Realiza un mapeo no lineal entre el espacio de entrada y el espacio de salida y envía los pesos asignados a las capas posteriores. Finalmente, determina la salida y la compara con la salida deseada. Identifica la señal de error que ajusta los pesos de conexión correspondientemente.

Árboles de decisión (DT): el árbol de decisión es un método de aprendizaje supervisado no paramétrico que se utiliza ampliamente como clasificador. Proporciona una estructura de árbol tipo modelo que se puede utilizar para detectar la etiqueta de clase de una variable objetivo mediante el aprendizaje de reglas de decisión simples inferidas de los datos de entrenamiento etiquetados en clase. El proceso de clasificación de las muestras de prueba se realiza organizándolas en función de los valores de las características. Cada nodo en un árbol de decisión denota una característica, y cada rama se genera comparando el valor de prueba con un valor umbral del nodo. La clasificación se realiza trazando una ruta desde el nodo raíz a un nodo hoja que no tiene borde saliente.

Red neuronal de propagación hacia atrás: la propagación hacia atrás minimiza el error cuadrado total de la salida que calcula el resto. La red neuronal responde con precisión a los patrones de entrada que se utilizan para el entrenamiento. Esto se conoce como memorización. La red también debe tener una entrada razonable que sea similar a las muestras utilizadas para la capacitación. Esto se llama generalización. Las tres etapas del entrenamiento son las siguientes:

- (i) Avance del patrón de entrada.
- (ii) Cálculo y propagación inversa del error asociado.
- (iii) Ajustes de los pesos.

Clasificación difusa: la difusa es una técnica utilizada de forma dominante en el reconocimiento y clasificación de objetos. Los clasificadores difusos son más adecuados para datos reales donde los límites entre subgrupos no están bien definidos. En su mayoría, este

clasificador está integrado con otros sistemas de clasificación como los clasificadores neuro-difusos, los árboles de decisión difusos y el algoritmo vecino difuso k-más cercano para un mejor rendimiento. El análisis de los EEG del sueño se puede hacer usando clasificadores difusos (Angana & Sudip, 2019).

Medidas de desempeño

Se requiere evaluar el desempeño de las BCI. Con respecto al clasificador solo, la medida de rendimiento más básica es la precisión de la clasificación. Esto es válido solo si las clases están equilibradas, es decir, con el mismo número de muestras por clase y si el clasificador es imparcial, es decir, tiene el mismo rendimiento para cada clase.

Si no se cumplen estas condiciones, la métrica de Kappa o la matriz de confusión son medidas de rendimiento más informativas. El par de sensibilidad-especificidad, o precisión, se puede calcular a partir de la matriz de confusión. Cuando la clasificación depende de un parámetro continuo (por ejemplo, un umbral), a menudo se utilizan la curva de características de funcionamiento del receptor (ROC) y el área bajo la curva (AUC).

El rendimiento del clasificador generalmente se calcula fuera de línea en datos pregrabados, utilizando una estrategia de retención: algunos conjuntos de datos se reservan para ser utilizados para la evaluación y no forman parte del conjunto de datos de capacitación. Sin embargo, algunos autores también informan medidas de validación cruzada estimadas en datos de entrenamiento, que pueden sobrevalorar el rendimiento (F, L, A, M, & M, 2018).

Las métricas basadas en la época se utilizan para evaluar el rendimiento del clasificador. En problemas de decisión binaria por ejemplo (artefacto versus no artefacto), la decisión tomada por el clasificador puede representarse en una estructura conocida como tabla de contingencia. Esta tabla de contingencia tiene cuatro categorías: los verdaderos positivos (TP) son épocas correctamente etiquetadas como artefactos; los falsos positivos (FP) se refieren a épocas etiquetadas incorrectamente como artefactos; los verdaderos negativos (TN) corresponden a épocas sin artefactos correctamente etiquetados y los falsos negativos (FN) se refieren a épocas incorrectamente etiquetadas como no artefactos. La precisión de cada clase se evalúa utilizando Sensibilidad y especificidad. La sensibilidad se define como:

$$\text{Sensibilidad} = \text{TP}/(\text{TP} + \text{TN})$$

La especificidad se define como:

$$\text{Especificidad} = \text{TN}/(\text{TN} + \text{FP})$$

Las curvas características del operador receptor (ROC) muestran cómo la sensibilidad varía con la especificidad. El área bajo la curva ROC es una forma efectiva de comparar el rendimiento de dos sistemas diferentes. Una discriminación aleatoria dará un área de 0.5 debajo de la curva, mientras que la discriminación perfecta dará un área de unidad debajo de la curva ROC. El área bajo la curva ROC (área ROC) se calcula utilizando la integración numérica trapezoidal. (Simon, Stephen, & William, 2011)

CAPÍTULO 4

4. Estado del arte

Los primeros perfiles de usuarios de una BCI eran personas con trastornos motores graves, tales como la esclerosis lateral amiotrófica (ELA) o el síndrome del enclaustramiento.

Las mejoras en los sistemas de medición de EEG permiten el desarrollo de nuevas formas de acceso, pensando en cubrir las necesidades de los pacientes con discapacidades cognitivas y motoras, el desafío de obtener los mayores beneficios de un EEG portátil es el enfoque de múltiples trabajos. Las pruebas realizadas con aplicaciones piloto cubren áreas terapéuticas como lenguaje, déficit de atención, discapacidad motora y educación temprana. Sus resultados establecen una perspectiva alentadora para continuar con el desarrollo de este tipo de aplicaciones. Además, de que los trabajos se adentran en la investigación del comportamiento individual con el análisis de las respuestas de estímulos visuales afectivos (Garzón J. R., 2013).

4.1. Historia y desarrollo del campo

Las primeras investigaciones sobre las BCI se iniciaron en la década de 1970 en la Universidad de California Los Ángeles (UCLA) con asistencia de la National Science Foundation, seguido por un contrato de DARPA. Los artículos que se publicaron a raíz de esta investigación marcaron por primera vez lo que es una interfaz cerebro-computadora.

Tras años de experimentación animal, las primeras neuro-protésis implantados en seres humanos aparecieron a mediados de la década de 1990.

Actualmente se han realizado grandes progresos, por ejemplo, un artículo publicado por Hochberg de la Universidad de Brown, publicado en la revista Nature (2012). En este trabajo presentó una BCI con la que dos personas con tetraplejia podían controlar un brazo robótico y realizar varias tareas, por medio de su actividad cerebral (Prieto, Pelayo, López, & Romero, 2013).

Se han demostrado diferencias en la actividad cerebral relacionada con la categoría con varios métodos de neuroimagen en sujetos sanos, para seres vivos versus objetos hechos por el hombre, y para varias categorías de objetos específicos, como caras, partes del cuerpo,

animales, frutas / verduras, edificios, herramientas y muebles (Gerlach, 2007), (Martin A. , 2007), (Binder, Desai, Graves, & Conant, 2009).

Las investigaciones y retos actuales se centran en las modalidades anteriormente mencionadas:

- Nuevos procedimientos y técnicas sensoriales que proporcionen una mejor resolución temporal y espacial de las medidas de la actividad.
- Con respecto al procesamiento de las señales, el reto es mejorar los procesos de filtrado y de aprendizaje automático de señal.
 - Los aspectos que dificultan esta parte son debidos a que cada individuo tiene una actividad cerebral distinta y además esta actividad cerebral cambia con el tiempo.
- Nuevas aplicaciones con interés asistencial o comercial.

Dentro de la variedad de propuestas a nivel general que tienen su principio de funcionamiento en la tecnología EEG, se encontraron algunas destacadas por la relación al momento de optimizar funciones relacionadas con las emociones en las personas, la mayoría surgen con base al estudio en la interacción de los individuos en diferentes entornos, entregando datos y estadísticas cuyas aplicaciones no han generado impactos significativos relacionados con el beneficio directo en la sociedad; esto generó la aparición de ciertas aplicaciones tales como aquella que realizó un estudio de predicción en los resultados en un examen hecho por estudiantes a partir de sus emociones; también un programa de mensajería instantánea que incorpora en sus mensajes iconos y palabras de acuerdo al estado emocional de la persona que se comunica, o un estudio de las sensaciones en las personas que interaccionan en una ciudad concurrida (Garzón J. R., 2013).

La identificación de los procesos neurales que subyacen a las representaciones semánticas es un desafío clave en la neurociencia cognitiva. Se han propuesto diferentes hipótesis sobre cómo las representaciones de conceptos particulares establecen un sistema de conocimiento conceptual. La aceptación general es que las propiedades de los objetos compartidos se reflejan en la organización del sistema semántico y que el sistema se generaliza a través de conceptos que pertenecen a una categoría particular (como animales, herramientas o edificios).

La noción de especificidad de categoría en la organización del conocimiento del objeto surgió en la década de 1980, cuando Warrington y sus colegas informaron por primera vez sobre pacientes con discapacidad selectiva para una categoría semántica en comparación con otras categorías semánticas (Simanova, Gerven, Oostenveld, & Hagoort, 2010), (Gerlach, 2007), (Binder, Desai, Graves, & Conant, 2009). Desde esas investigaciones iniciales, una gran cantidad de estudios han confirmado el fenómeno de los déficits semánticos específicos de la categoría. Se ha informado que los pacientes tienen impedimentos para todo tipo de conocimiento sobre una categoría particular, como, por ejemplo, los seres vivos (Patterson, Nestor, & Rogers, 2007), (Mahon & Caramazza, 2009).

Se han demostrado diferencias en la actividad cerebral relacionada con la categoría con varios métodos de neuroimagen en sujetos sanos, para seres vivos versus objetos hechos por el hombre, y para varias categorías de objetos específicos, como caras, partes del cuerpo,

animales, frutas / verduras, edificios, herramientas y muebles (Binder, Desai, Graves, & Conant, 2009)

La activación diferencial, que sugiere una organización funcional específica, se ha demostrado en el procesamiento de las modalidades de estímulo visual y verbal. Para algunos tipos de objetos, la organización funcional por categoría semántica se ha demostrado dentro de una modalidad dada, por ejemplo, categoría-especificidad en la vía visual para caras (Chand, Lamichhane, & Dhamala, 2016) o para entidades vivas versus no vivas (Kozunov, et al., 2018). También se ha demostrado que los objetos y sus atributos sensoriales o funcionales (como las acciones asociadas a herramientas) activan las mismas regiones neuronales (Kiefer, Sim, Herrnberger, J, & Hoenig, 2008), lo que sugiere que estas regiones están implícitamente involucradas en la representación conceptual.

Lograr una imagen más clara de las distinciones categoriales en el cerebro es esencial para nuestra comprensión del léxico conceptual, pero investigaciones mucho más precisas tanto en distinciones categoriales como en otros aspectos de la representación conceptual (Maguire, Brier, & Ferree, 2010) serán necesarios para que esta evidencia contribuya a la investigación léxica. Aunque la semántica claramente desempeña un papel central en la capacidad del lenguaje humano, ya que la transferencia de significado es el objetivo de una comunicación con mucho propósito, nuestra comprensión de su instanciación y localización funcional en el cerebro está lejos de ser completa. Una razón para esto es el carácter complejo y multimodal del significado: las representaciones conceptuales están arraigadas conjuntamente en el lenguaje, los sistemas visuales y de otro tipo (Murphy, et al., 2011), y resultados de imágenes que sugieren que los centros clásicamente "visuales", particularmente el giro fusiforme izquierdo, son sensibles a la semántica del lenguaje independientemente de sus regularidades visuales (Martin A. , 2007).

Este tipo de sistemas pueden tener aplicaciones en la vida real muy concretas, por ejemplo, como apoyo a personas que perdieron el habla, personas que no escuchan, personas que han tenido un derrame cerebral y se necesita verificar si los conceptos aún se encuentran presentes en su mente (Sciotto, 2018).

A continuación se presentan trabajos destacados y relacionados con en el área, en los cuales se expresan los métodos implementados para los procesos de categorización semántica, los avances existentes hasta hoy en día, con el objetivo de analizar la teoría subyacente a la metodología del tema en cuestión.

4.2. La representación de conceptos de objeto en el cerebro

Concepto de objeto: “Representaciones de memoria de una clase o categoría de objetos. Necesario para numerosas funciones cognitivas, incluida la identificación de un objeto como miembro de una categoría específica y hacer inferencias sobre las propiedades del objeto” (Martin A. , 2007).

Las bases neuronales del reconocimiento de objetos y del conocimiento semántico se han estudiado ampliamente, pero la elevada dimensionalidad del espacio objetual hace que tenga un alto grado de dificultad desarrollar teorías globales sobre cómo el cerebro organiza el conocimiento de los objetos. Para ayudar a entender cómo el cerebro nos permite reconocer,

categorizar y representar objetos y categorías de objetos, existe un creciente interés en utilizar bases de datos de imágenes a gran escala para experimentos de neuroimagen. Existen estudios que se han encargado de recolectar tales señales, en (Grootswagers, Zhou, Robinson, Hebart, & Carlson, 2022) se presenta un conjunto de datos “THINGS-EEG” que contiene respuestas de electroencefalografía humana de 50 sujetos a 1,854 conceptos de objeto y 22,248 imágenes del conjunto de estímulos THINGS, una base de datos de imágenes de alta calidad y curada manualmente que fue diseñada específicamente para estudiar la visión humana. El conjunto de datos THINGS-EEG proporciona registros de neuroimagen a una colección sistemática de objetos y conceptos y, por lo tanto, puede apoyar una amplia gama de investigaciones para entender el procesamiento de objetos visuales en el cerebro humano.

La evidencia de la neuroimagen funcional del cerebro humano indica que la información sobre las propiedades sobresalientes de un objeto, como su aspecto, cómo se mueve y cómo se usa, se almacena en los sistemas sensoriales y motores activos cuando se obtuvo esa información. Como resultado, los conceptos de objetos que pertenecen a diferentes categorías, como animales y herramientas, se representan en redes neuronales basadas en propiedades sensoriales y motoras parcialmente distintas. Esto sugiere que los conceptos de objeto no se representan explícitamente, sino que surgen de la actividad ponderada dentro de las regiones del cerebro basadas en propiedades. Sin embargo, algunas regiones basadas en propiedades parecen mostrar una organización categórica, proporcionando así evidencia consistente con formulaciones específicas de dominio basadas en categorías también.

Durante más de cien años, el pensamiento sobre la representación de los conceptos de objeto en el cerebro ha estado dominado por los modelos de propiedades sensoriomotoras (que datan de 1890 a informes recientes de muchos investigadores). La idea central es que el conocimiento del objeto está organizado por características sensoriales (Forma, movimiento, color) y propiedades motoras asociadas con el uso del objeto (y en algunos modelos, otras propiedades funcionales / mediadas verbalmente, como dónde se encuentra típicamente un objeto, su significado social, etc.) (Martin A. , 2007).

Los registros de la actividad cerebral de las técnicas de neuroimagen y electrofisiológicas se han establecido durante mucho tiempo como fuentes cruciales de evidencia en el estudio del conocimiento conceptual. Los estudios de imagen (por ejemplo (Chao, Weisberg, & Martin, 2002)) han proporcionado una amplia evidencia que respalda las hipótesis relativas a la localización del conocimiento conceptual derivado del trabajo anterior con pacientes que presentan déficits de categoría semántica (Mahon & Caramazza, 2009). Ahora, los métodos computacionalmente intensivos se están convirtiendo cada vez más en el método de elección para analizar dichos datos (Shinkareva, et al., 2008). La baja relación señal / ruido de las grabaciones neuronales significa que muchas regularidades permanecen ocultas y solo se pueden detectar a través del análisis de patrones a gran escala.

La mayoría de los estudios examinan solo distinciones en campos semánticos muy alejados, por ejemplo comparando conceptos abstractos y concretos, verbos y sustantivos, o tipos naturales y artefactuales (Calvo, et al., 2018).

La evidencia convergente de la neurofisiología del mono, la neuropsicología y las imágenes cerebrales funcionales ha establecido que el reconocimiento de objetos depende críticamente de la corriente de procesamiento occipitotemporal ventral (ver (Grill-Spector &

Malach, 2004)). Además, los estudios funcionales de imágenes cerebrales del reconocimiento de objetos han proporcionado pruebas convincentes de que la corteza occipitotemporal no es un sistema homogéneo de procesamiento de objetos, sino que tiene una estructura de grano fino que parece estar relacionada con la categoría de objeto. Las categorías más estudiadas han sido rostros humanos, casas, animales y herramientas (Chand, Lamichhane, & Dhamala, 2016). La comparación directa de una categoría de objeto con otra ha revelado distintos grupos de actividad (por ejemplo, el área fusiforme de las caras, (FFA por sus siglas en inglés); la región del cerebro llamada en inglés Parahippocampal Place Area (PPA) (Fig. 4.1). Además, las técnicas de análisis de patrones han identificado distintos patrones de actividad relacionados con la categoría de objetos que discriminan entre un número relativamente grande de categorías de objetos (Behroozi, et al., 2015), (Cichy, Radoslaw, Pantazis, & Dimitrios, 2017), (Grootswagers, et al., 2018). Estos patrones relacionados con la categoría de objetos se extienden sobre una gran extensión de corteza occipitotemporal, son estables tanto dentro como entre los sujetos, y pueden identificarse incluso cuando los sujetos ven libremente escenas complejas.

(Martin A. , 2007) Proporcionó evidencia de que los grupos de actividad relacionados con la categoría en las cortezas occipitotemporales asociadas con la visualización de imágenes de objetos también se ven cuando los sujetos participan en una tarea de procesamiento conceptual verbal. El procesamiento perceptual se evaluó mediante la visualización pasiva y la comparación diferida de muestras con imágenes de animales, herramientas, caras y casas. El procesamiento conceptual se evaluó mediante el uso de nombres de imágenes silenciosas y una tarea de verificación de propiedad que prueba el conocimiento de animales y herramientas denotadas por sus nombres escritos.

Los principales hallazgos fueron que los animales que perciben (así como las caras) mostraron una mayor actividad bilateral en la región más lateral del giro fusiforme, adicional de que las caras y los animales activan selectivamente la región lateral del giro fusiforme incluyendo FFA, y (Grill-Spector & Malach, 2004) por evidencia de que esta actividad está significativamente correlacionada con la identificación de caras y pájaros, pero no con otros objetos). En contraste, las herramientas mostraron una mayor actividad bilateral en la región medial del giro fusiforme.

Por lo tanto, aunque cada concepto de objeto debe tener una representación neural distinta, parece existir una amplia distinción entre agentes animados y objetos manipulables hechos por el hombre en las porciones lateral y medial del giro fusiforme, respectivamente. La reciente demostración de que los cuerpos humanos están representados junto a la representación de caras en la porción lateral de la circunvolución fusiforme es consistente con esta visión (Cichy, Radoslaw, Pantazis, & Dimitrios, 2017).

Los seres humanos son capaces de reconocer visualmente un gran número de objetos diferentes e interactuar con ellos de forma significativa, a pesar de los cambios drásticos en la proyección de la retina, la iluminación o el ángulo de visión, y de que los objetos estén situados en entornos visuales desordenados. El reconocimiento de objetos y el conocimiento semántico, nuestra capacidad para dar sentido a los objetos que nos rodean, han sido objeto de una gran cantidad de investigaciones de neurociencia cognitiva. Sin embargo, las

investigaciones previas de neuroimagen en este campo se han basado a menudo en una selección manual de un pequeño conjunto de imágenes.

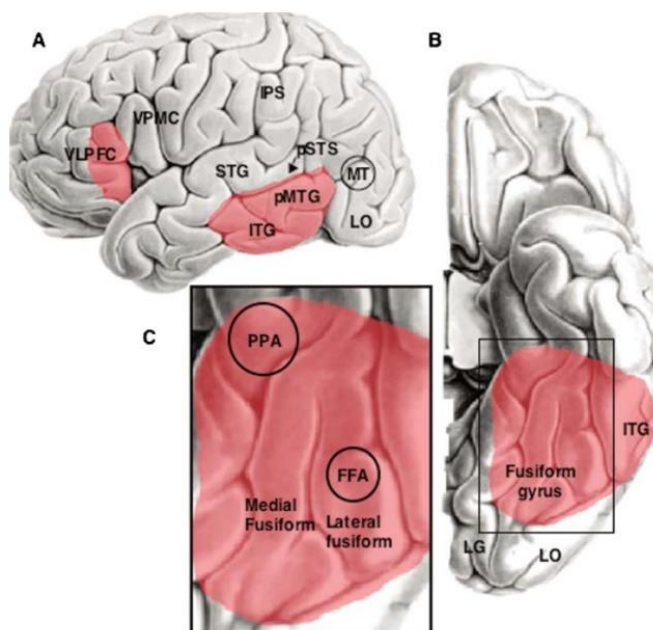


Figura 4.1. Vista lateral esquemática del hemisferio izquierdo (A) y vista ventral del lóbulo frontal y temporal derecho (B). El giro fusiforme se muestra con mayor detalle en (C). Las regiones en rojo muestran la ubicación aproximada de áreas típicamente involucradas en tareas de procesamiento conceptual, especialmente con objetos concretos. ITG, giro temporal inferior; LG, circunvolución lingual. Ver texto para otras abreviaturas. Fuente: (Martin A. , 2007)

Una gran cantidad de evidencia funcional de neuroimagen ha implicado los lóbulos temporales, particularmente la región posterior del lóbulo temporal izquierdo, como un sitio crítico para las representaciones almacenadas, especialmente sobre objetos concretos.

Estudios recientes han proporcionado apoyo adicional para este punto de vista al demostrar que las regiones de la corteza temporal posterior izquierda que se sabe que están activas durante el procesamiento conceptual de imágenes y palabras (giro fusiforme y giro temporal inferior y medio) también estaban activas durante la comprensión auditiva de las oraciones (Hinojosa, et al., 2019).

Otro enfoque reciente para investigar la neuroanatomía funcional del procesamiento conceptual ha sido utilizar tareas de repetición de estímulo. Está bien establecido que la experiencia previa con un estímulo da como resultado un procesamiento más eficiente (primado de repetición) y una respuesta hemodinámica reducida, típicamente conocida como supresión de repetición, pero también como adaptación, primado neural y atenuación de repetición, cuando se encuentra ese estímulo más adelante (Hinojosa, et al., 2019). En esta revisión literaria se presentan resultados de los estudios de fMRI y PET, donde indican que el procesamiento de las características emocionales de las palabras implica regiones cerebrales involucradas en el procesamiento de información afectiva y semántica. La Fig. 4.2 muestra las regiones del cerebro que se ha demostrado que se activan más consistentemente en investigaciones previas.

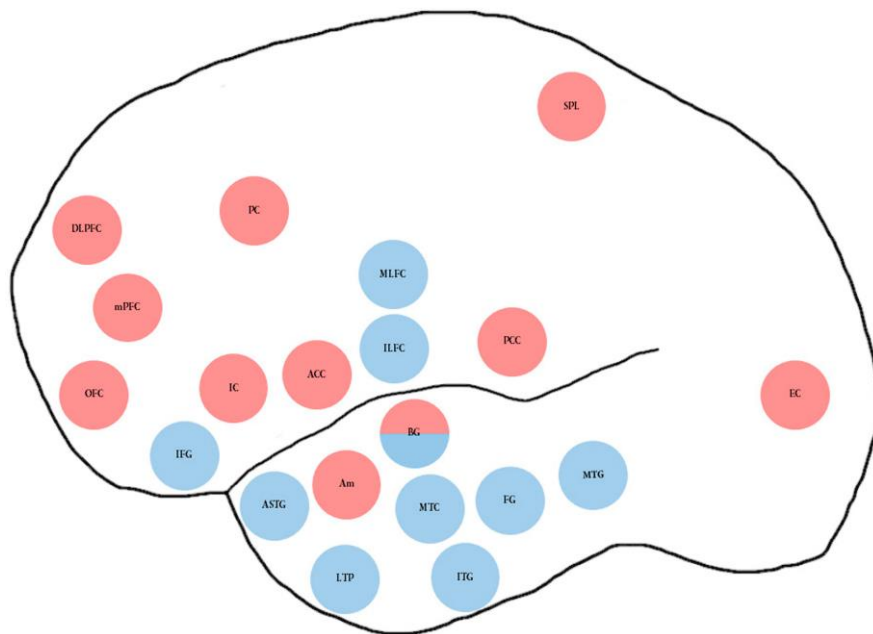


Figura 4.2. Las regiones cerebrales se activan más consistentemente durante el procesamiento de las características emocionales en los estudios de procesamiento de palabras y oraciones. Los círculos rojos denotan áreas que se han asociado principalmente con el procesamiento afectivo, mientras que los círculos azules indican regiones que sirven principalmente a procesos relacionados con el lenguaje. Abreviaturas: ACC, corteza cingulada anterior; PCC, corteza cingulada posterior; OFC, corteza prefrontal orbital; DLPFC, corteza prefrontal dorsolateral; mPFC, corteza prefrontal medial; IC, corteza insular; SPL, lóbulo parietal superior; EC, corteza extra-estriada; Am, amígdala; BG, ganglios basales; PC, corteza premotora; MTC, corteza temporal media; IFG, giro frontal inferior; FG, giro fusiforme; MTG, giro temporal medio; LTP, polo temporal izquierdo; ITG, giro temporal inferior; ILFC, corteza frontal inferior izquierda; MLFC, corteza frontal media izquierda; ASTG, giro temporal superior anterior (Hinojosa, et al., 2019).

Se han observado patrones de actividad relacionada con la categoría de objetos en las regiones ventral y lateral de la corteza temporal posterior utilizando una variedad de estímulos (imágenes, nombres escritos, sonidos asociados a objetos, nombres escuchados). En la corteza temporal ventral, se ha encontrado una mayor actividad en la región lateral de la circunvolución fusiforme (incluida la FFA) mediante la asignación de nombres, la categorización de nivel básico o las tareas de decisión semántica con imágenes de animales y / o sus nombres escritos también en respuesta a los nombres hablados (Grill-Spector & Malach, 2004)).

Las ubicaciones de las activaciones relacionadas con la categoría informadas en estos estudios muestran un notable grado de consistencia entre sujetos a pesar de las marcadas diferencias en la tarea de procesamiento, el formato de estímulo (verbal, no verbal) y la modalidad de presentación (visual, auditiva). Sin embargo, estos factores son claramente relevantes. De hecho, los estudios de Price y sus colegas muestran que la actividad neuronal relacionada con la categoría en la corteza temporal posterior puede ser modulada por una serie de estímulos y variables relacionadas con la tarea (Mitchell, Heatherton, & Macrae, 2002).

La creciente evidencia sugiere que el "reconocimiento de objetos centrales", es decir, la capacidad de reconocer rápidamente objetos a pesar de una variación sustancial de apariencia, se resuelve en el cerebro a través de una cascada de cálculos reflexivos, en gran

parte anticipados, que culminan en una poderosa representación neuronal en la corteza temporal inferior (ver Fig. 4.3). Sin embargo, el algoritmo que produce esta solución sigue siendo poco conocido (James, Davide, & Nicole, 2012). En este artículo se revisa la evidencia que va desde neuronas individuales y poblaciones neuronales hasta modelos de comportamiento y computacionales. Se proponemos que la comprensión de este algoritmo requiere el uso de datos neuronales y psicofísicos para examinar muchos modelos computacionales, cada uno basado en bloques de construcción de pequeñas subredes canónicas con un objetivo funcional común.

En (Olague, et al., 2019) se define el reconocimiento de objetos como una tarea fundamental para los seres humanos y todos los seres vivos dotados del sentido de la vista ya que permite la interacción del organismo con el entorno circundante y su comprensión. Aquí se describe un sistema virtual compuesto por una vía dorsal artificial o corriente "dónde" y una vía ventral artificial o corriente "qué", ambas se fusionan para recrear una corteza visual artificial (ver Fig. 4.4). Se proporciona evidencia de que se podrían lograr altas tasas de reconocimiento en problemas de categorización de objetos bien conocidos mediante esta corteza artificial.

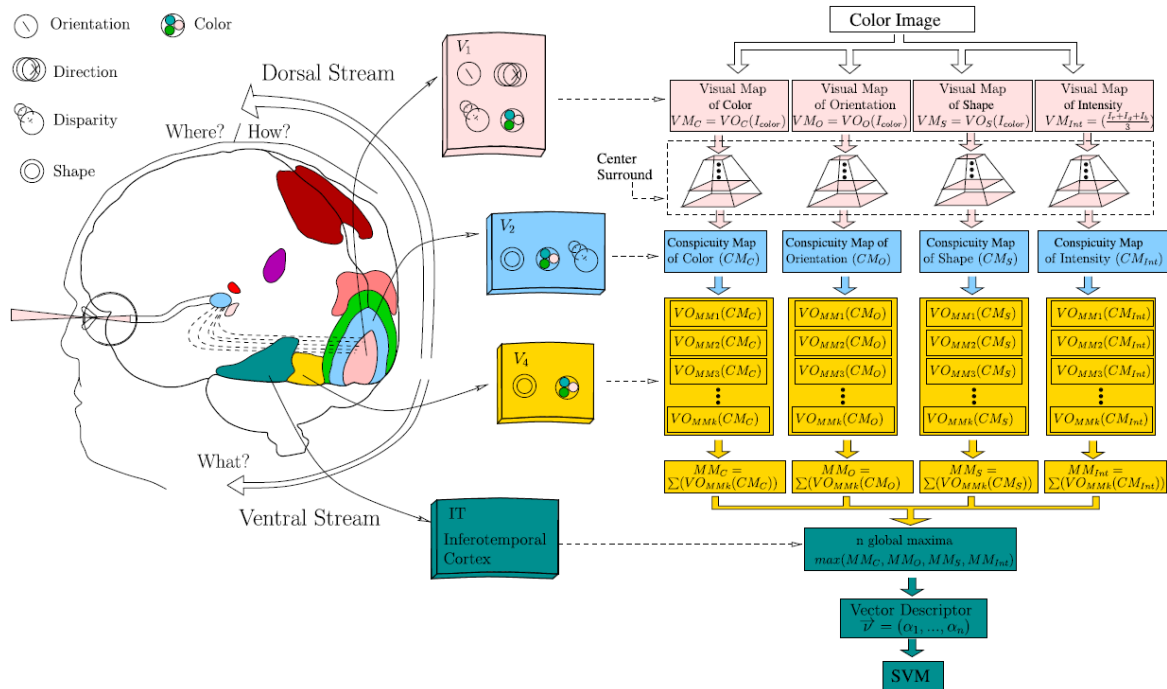


Figura 4.3. (A) Ubicación del área cortical de la corriente ventral en el cerebro del mono macaco y flujo de información visual desde la retina. (B) Cada área se traza de manera que su tamaño sea proporcional a su área de superficie cortical. El número total aproximado de neuronas (ambos hemisferios) se muestra en la esquina de cada área (M = millones). La dimensionalidad aproximada de cada representación (número de neuronas de proyección) se muestra encima de cada área, según las densidades neuronales, la fracción neuronal de capa 2/3 y la porción (color) dedicado al procesamiento de los 10 grados centrales del campo visual. La latencia de respuesta media aproximada se enumera a la derecha (James, Davide, & Nicole, 2012).

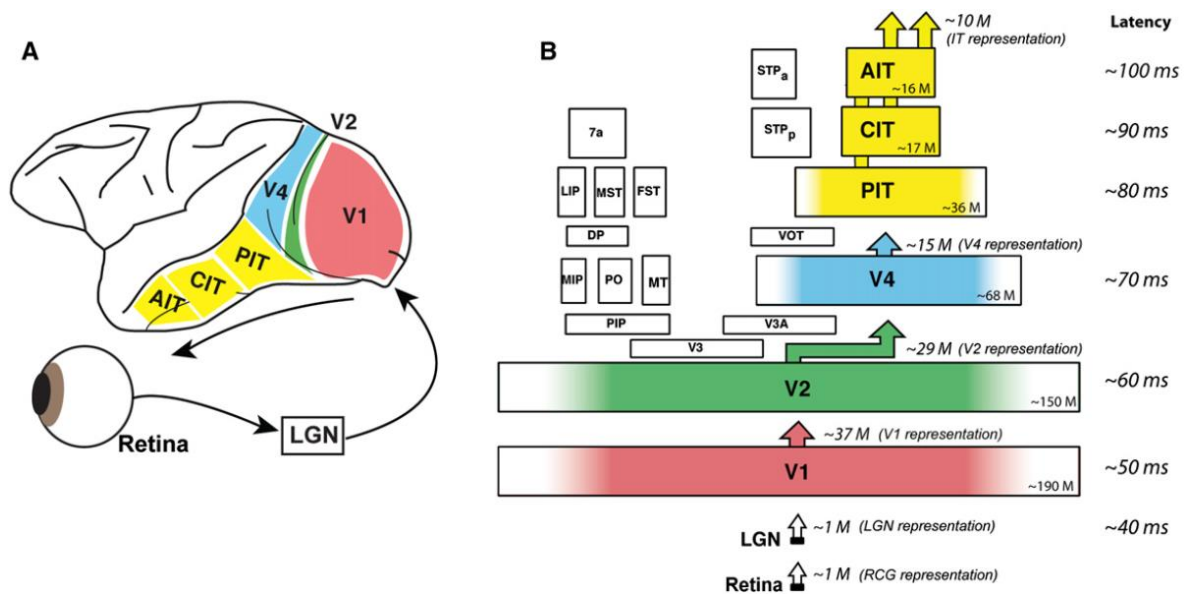


Figura 4.4. Modelo conceptual de la corteza visual artificial. La imagen en color se descompone en cuatro dimensiones (color, orientación, forma e intensidad). Luego, una estructura jerárquica se encarga de resolver el problema de clasificación de objetos a través de un paradigma dirigido por funciones (Olague, et al., 2019).

La decodificación de la representación visual se refiere a la tarea de descifrar lo que un sujeto está viendo o visualizando al observar el estado del cerebro a través de neuroimágenes. En los últimos años, existe un interés creciente en abordar esta tarea mediante el uso de enfoques de aprendizaje automático. En (Papadimitriou, Passalis, & Tefas, 2019) se proporciona una evaluación extensa que sirve como base para la decodificación de representaciones visuales, mediante la exploración de una amplia gama de configuraciones de modelos, representaciones de características y configuraciones de evaluación. De esta manera, este trabajo sienta las bases para desarrollar tuberías de decodificación más sofisticadas y precisas. Se evalúan métodos de decodificación, donde sugieren que las redes neuronales proporcionan, en promedio, el mejor rendimiento, mientras que la elección de la métrica de similitud más adecuada para el proceso de decodificación de clases depende principalmente de la tarea en cuestión. Finalmente, este trabajo otorga información de alto nivel sobre la función del cerebro, a través de varias observaciones interesantes, por ejemplo, regiones del cerebro que son dominantes para tareas específicas y respaldan afirmaciones relacionadas sobre la posible correspondencia de la jerarquía cortical con representaciones visuales profundas.

En general, el sistema visual humano puede reconocer y clasificar un objeto según su categoría con facilidad. Ambas tareas consideran que el conjunto de atributos o características extraídas de las imágenes son lo suficientemente generales como para clasificar el objeto como parte de la clase mientras se mantienen en la memoria los elementos que sirven para identificar ese objeto en particular dentro de una escena dada (Olague, et al., 2019).

(Mohedano, et al., 2014) Propone un sistema que tiene como objetivo detectar y segmentar un objeto a partir de una imagen utilizando las señales cerebrales medidas del usuario en el momento de observar una región específica. La idea es transformar las respuestas de EEG

medidas en un mapa que proporcione una estimación de la probabilidad de que una región en particular vista por el usuario contenga el objeto objetivo, y luego usar este mapa para generar un algoritmo de segmentación. La construcción de este mapa se basa en la clasificación de señales de EEG, ya que se sabe que las respuestas eléctricas del cerebro difieren cuando el usuario detecta un objetivo o un estímulo raro en un escenario mediante presentación visual rápida en serie (RSVP). Este estudio muestra que el BCI y el análisis EEG simple son útiles para localizar los límites de los objetos en las imágenes.

En (Asly, Gilde, Moctezuma, & Molinas, 2019) se analiza la posibilidad de activar dispositivos observando colores primarios, lo que se cree que es especialmente útil para las personas con problemas de control motor restringido. Tal BCI requiere una detección confiable de las características relacionadas con el color capturada en los datos del electroencefalograma (EEG). Este artículo presenta métodos analíticos y empíricos de análisis de señales para analizar señales de EEG, motivados por la búsqueda de características directamente relacionadas con la percepción del color en el cerebro humano. Los métodos utilizados son la transformada de Fourier (FT) y la transformada de Fourier de tiempo corto (STFT). La descomposición en modo empírico (EMD) se utiliza para extraer la información utilizada para la extracción de características. Las precisiones de clasificación se prueban utilizando los algoritmos de aprendizaje automático: bosque aleatorio (RF), máquina de vectores de soporte (SVM), vecinos de knearest (kNN), árbol de decisión (DT) y Bayes ingenuo (NB).

El éxito del aprendizaje profundo en la visión por computadora ha aumentado en gran medida la necesidad de conjuntos de datos de imágenes etiquetados. El trabajo realizado por (Parekh, Subramanian, Roy, & C, 2018) propone un sistema de etiquetado de imágenes basado en EEG. Si bien los humanos pueden reconocer objetos en 20-200 milisegundos, la necesidad de etiquetar imágenes manualmente da como resultado un rendimiento de anotaciones bajo. Este sistema emplea señales cerebrales capturadas a través de un dispositivo EEG de consumidor para lograr una tasa de etiquetado de hasta 10 imágenes por segundo. Se aprovecha la firma del potencial relacionado con eventos (ERP) del P300 para identificar imágenes de destino durante una tarea de presentación visual en serie rápida (RSVP). Además, se lleva a cabo la eliminación de valores atípicos sin supervisión para lograr una puntuación F1 de 0,88 en el conjunto de prueba.

Reconocimiento de objetos en representaciones mentales

Uno de los objetivos fundamentales de la investigación del reconocimiento de objetos es comprender cómo una representación cognitiva producida a partir de la salida de información sensorial filtrada y transformada facilita el comportamiento eficiente del espectador. Las imágenes mentales visuales, por tanto, representan un flujo de información opuesto al de los fenómenos de percepción visual; uno que requiere una desviación de las perspectivas tradicionales de abajo hacia arriba para que se entienda completamente.

Dado que las imágenes mentales se parecen mucho a los procesos perceptivos tanto en las regiones corticales como en las cualidades visuales subjetivas, es razonable cuestionar si las imágenes mentales facilitan la cognición de una manera similar a la de la visión perceptiva: a través de la detección y el reconocimiento de características distintivas. La categorización

del contenido característico de las imágenes mentales tiene el potencial de ser una vía inversa mediante la cual identificar los componentes de un estímulo visual que son más críticos para la creación y recuperación de una representación visual. En (Roldan, 2017) se realiza una revisión para examinar la probabilidad de que la información representada en las imágenes mentales visuales refleje características distintivas del objeto que se cree que facilitan la categorización y el reconocimiento eficiente del objeto durante la visualización perceptiva. Se revisan los métodos prominentes de comportamiento y neuroimagen que se han utilizado para acceder con éxito a las representaciones visuales mentales; se evalúa el potencial de cada método para identificar características de diagnóstico en función de la eficiencia y especificidad que se ha demostrado que logran; y se sugieren posibles direcciones e implicaciones para explorar características distintivas a través de imágenes mentales en investigaciones futuras.

4.3. Identificación de categorías de objetos a partir de EEG relacionado con eventos

En primer lugar, la electroencefalografía (EEG) tiene una capacidad bien documentada para caracterizar ciertos estados cerebrales, en particular el procesamiento de diferentes categorías semánticas. En segundo lugar, la alta resolución temporal del EEG permite una caracterización precisa de la recuperación del concepto en términos de los patrones electrofisiológicos que hacen posible la decodificación. En tercer lugar, el desarrollo de algoritmos de decodificación semántica basados en EEG es interesante desde la perspectiva de las aplicaciones, ya que la resolución temporal de EEG permite la decodificación en tiempo real. Cuando sea posible decodificar información conceptual de EEG, un sistema de interfaz cerebro-computadora que transforma los conceptos léxicos en una salida escrita o hablada podría convertirse en una realidad (Simanova, Gerven, Oostenveld, & Hagoort, 2010). En este estudio, se aplica el enfoque bayesiano para identificar la actividad neuronal relacionada con los conceptos a partir de potenciales cerebrales relacionados con eventos (ERP). Se presentan estímulos de dos categorías semánticas: animales y herramientas, y se entrena a un clasificador para discriminar estas categorías. Estiman el rendimiento de la clasificación e interpretamos los mapas de importancia obtenidos a nivel de sujeto único para tres modalidades de estímulo: auditivo (nombre hablado de un objeto), visual (dibujo de un objeto) y ortográfico (nombre escrito de un objeto), los resultados se aprecian en la Fig. 4.5.

La nueva tecnología EEG permite el desarrollo de nuevas formas de acceso, pensando en cubrir las necesidades de los pacientes con discapacidades cognitivas y motoras, el desafío de obtener los mayores beneficios de un EEG portátil es el enfoque de múltiples trabajos (Sciottto & Niripil, 2018)

Las investigaciones en esta área han logrado una interpretación de las ondas cerebrales que a simple vista podrían parecer totalmente aleatorias. Para una correcta lectura de las señales producidas por el cerebro es necesario tener un dispositivo robusto que permita la lectura de las ondas que se requieren medir (Gómez, García, & Roque, 2016), y posteriormente se busca tratar las señales electroencefalográficas (EEG) crudas y aplicar distintos métodos matemáticos, que permiten obtener información que haga posible entender que suceden con las ondas cerebrales ante distintos estímulos externos, accionamientos cognitivos,

enfermedades neurológicas, concentración, influencia de sustancias, etc. con el objetivo de lograr una comprensión más detallada de la actividad cerebral.

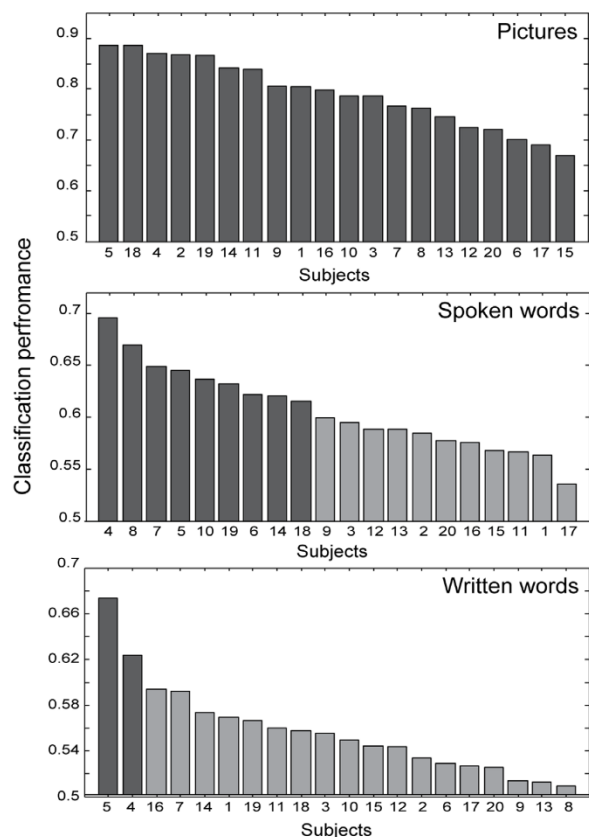


Figura 4.5. Clasificación del desempeño para las tres modalidades. Las barras oscuras indican resultados significativos (Simanova, Gerven, Oostenveld, & Hagoort, 2010).

Múltiples trabajos en la literatura reportan análisis de señales EEG para la categorización de objetos en el cerebro humano. La metodología de estos análisis comprende desde la adquisición de las señales, el procesamiento, como los tipos de filtros que se implementan para la mejora y enriquecimiento de las señales, el estímulo presentado ante los usuarios con el fin de obtener resultados más concretos, el proceso de extracción de características específicas que permite facilitar la interpretación de aspectos y rasgos únicos y discriminantes de cada señal referenciada a un estímulo visual, y finalmente los métodos de clasificación de estas señales, para poder concluir con una categorización con un nivel de precisión aceptable.

Existen múltiples técnicas de análisis de patrones que permiten la decodificación de información conceptual. En (Simanova, Gerven, Oostenveld, & Hagoort, 2010) se investiga la posibilidad de identificar representaciones conceptuales de EEG relacionadas con eventos basadas en la presentación de un objeto en diferentes modalidades: su nombre hablado, su representación visual y su nombre escrito. Se implementa la regresión logística bayesiana con un Laplace multivariado antes de la clasificación, para identificar la actividad neuronal relacionada con los conceptos a partir de potenciales cerebrales relacionados con eventos

(ERP). Se obtuvieron las más altas precisiones (89% de ensayos clasificados correctamente) al clasificar los dibujos de objetos (su representación visual).

Lograr una imagen más clara de las distinciones categoriales en el cerebro es esencial para nuestra comprensión del léxico conceptual, pero se requieren investigaciones mucho más detalladas para que esta evidencia contribuya a la investigación léxica. En (Murphy, et al., 2011). Se presenta una colección de técnicas avanzadas de extracción de datos que permite descifrar la categoría de conceptos individuales a partir de ensayos individuales de datos EEG. Los tiempos de lectura de las palabras en una oración dependen de la cantidad de información que transmiten las palabras, que puede estimarse mediante modelos de lenguaje probabilístico. Existen trabajos que investigan si los ERP también se predicen mediante medidas de información. Una comparación entre las medidas de información y los ERP revelaron una correlación confiable entre la amplitud N400 y una palabra sorpresa (Martin A. , 2007). Los modelos de lenguaje que no utilizan la estructura sintáctica se ajustaban mejor a los datos que una gramática de estructura de frase, que no tenía en cuenta la variación única en la amplitud de N400. Estos hallazgos sugieren que diferentes medidas de información cuantifican procesos cognitivamente diferentes y que los lectores no hacen uso de la estructura jerárquica de una oración para generar expectativas sobre la próxima palabra.

(Martin A. , 2007) Fundamenta su investigación en estudiar la relación semántica entre pares de sustantivos de objetos concretos como “caballo-oveja” y “columpio-melón” y cómo esta relación de actividad se refleja en las señales de EEG. Los autores realizan un análisis enfocado en algoritmos de extracción de características. Entrenan diferentes clasificadores para asociar un conjunto de señales a una respuesta humana previamente aprendida, perteneciente a dos clases: semánticamente relacionadas o no semánticamente relacionadas.

En investigaciones previas se asume que la percepción de la acción y el reconocimiento de objetos se mantienen mediante flujos de procesamiento funcional y neuroanatómicamente distintos en el cerebro. Sin embargo, la evidencia reciente desafía esta visión clásica y sugiere una interacción entre ambos flujos de procesamiento visual. Si bien los estudios anteriores mostraron una influencia de la percepción del objeto en las tareas relacionadas con la acción, en (Grootswagers, et al., 2018) se verifica si las representaciones de la acción facilitan el reconocimiento de objetos visuales.

En (Mohammad, Mitra, & Kavous, 2012), se seleccionan doce categorías diferentes como estímulos visuales y se presentaron a los sujetos en una tarea controlada y se realiza un análisis de cálculos diferentes de ERP donde el usuario distingue si el estímulo que se le presenta es un “objetivo” (categoría detectada)/”no objetivo” u “objetivo”/”descanso”, y los resultados proporcionan información útil sobre los canales y la parte de las señales que se ven afectadas por diferentes categorías de objetos en términos de señales cerebrales.

En la investigación (Chand, Lamichhane, & Dhamala, 2016) se pretende desenredar las actividades a nivel de nodo y de red en milisegundos de escala de tiempo de percepción y toma de decisiones. Se usan imágenes claras y ruidosas de caras y casas para la tarea de categorización de imágenes, y registros de EEG combinadas con técnicas de reconstrucción de fuente para estudiar cuándo y cómo se organiza la actividad oscilatoria dentro del área fusiforme de la cara (FFA), el área del lugar para-hipocampo (PPA) y la corteza prefrontal

dorsolateral (DLPFC). Se revela la dinámica de dos redes oscilatorias: beta (13-30 Hz) y gamma (30-100 Hz).

Cuando se estudian las representaciones cerebrales, los estímulos generalmente se presentan de forma aislada. En (Grootswagers, et al., 2018) se estudia la dinámica de la visión humana usando una combinación de tasas rápidas de presentación de estímulos, electroencefalografía y análisis de decodificación multivariada. Usando una tasa de presentación de 5 imágenes por segundo, se obtiene la estructura representativa de una gran cantidad de estímulos, y se presenta la organización categórica abstracta emergente de esta estructura. Además, es posible separar la dinámica temporal del procesamiento perceptual de los efectos de selección de objetivos de nivel superior.

El análisis de patrones multivariados de los datos de magnetoencefalografía (MEG) y EEG puede revelar la dinámica neural rápida subyacente a la cognición. Sin embargo, MEG y EEG tienen diferencias sistemáticas en el muestreo de la actividad neuronal. Esto plantea la pregunta de hasta qué punto tales diferencias de medición sesgan constantemente los resultados del análisis multivariado aplicado a los patrones de activación MEG y EEG. En (Cichy, Radoslaw, Pantazis, & Dimitrios, 2017) los autores realizan un estudio MEG / EEG concurrente mientras los participantes ven imágenes de objetos cotidianos. Se aplica un análisis de clasificación multivariado a los datos de MEG y EEG, y se comparan los cursos de tiempo resultantes entre sí, y a los datos de fMRI para una evaluación independiente en el espacio.

Un componente particularmente relevante para el procesamiento semántico es el N400. Investigaciones posteriores han demostrado que los componentes N400 se generan siempre que los eventos de estímulo inducen un procesamiento semántico o conceptual. Como tal, muchos investigadores han utilizado el componente N400 de las ondas cerebrales como una variable dependiente en experimentos psicolingüísticos. En (Wu & Coulson, 2007) se investiga cómo el habla y el gesto afectan los procesos de interpretación en tiempo real, y aborda los procesos cognitivos y neuronales que median la integración discurso-gesto.

En (Hanson, Matsuka, & Haxby, 2004) el autor evalúa la contribución de las características de nivel medio a la decodificación de categoría conceptual usando EEG y un nuevo paradigma de decodificación periódica rápida. Utiliza un conjunto de estímulo que consta de objetos intactos de las categorías animadas (por ejemplo, peces) e inanimadas (por ejemplo, silla) y versiones codificadas de los mismos objetos que eran. Sin embargo, la decodificación de animación para los objetos codificados solo fue posible a la velocidad de presentación periódica más lenta.

En (Isik, et al., 2013) se emplea un análisis de decodificación de magnetoencefalografía para medir la dinámica del desarrollo de información visual invariante de tamaño y posición en la corriente visual ventral. Con este método es posible leer la identidad de los objetos a partir de los 60 ms.

Utilizando métodos de decodificación de estímulo multivariante, sorprendentemente, es posible demostrar cómo los objetos son discriminados del patrón de fase de las señales de EEG a lo largo del tiempo en la banda de baja frecuencia (1-4 Hz), pero no del poder de las señales cerebrales oscilatorias en la misma banda de frecuencia. Por el contrario, la precisión de la discriminación de la potencia de las señales de EEG es significativamente mayor que

el rendimiento de la fase de la señal de EEG en la banda de alta frecuencia (20-30 Hz). En (Mohammad, Mitra, & Kavous, 2012) las señales de EEG junto con una técnica de reconocimiento de patrones multivariados se usaron para investigar la posibilidad de identificar la representación conceptual basada en la presentación de 12 categorías semánticas de objetos (5 ejemplares por categoría).

Los recientes desarrollos en visión por ordenador han producido conjuntos de imágenes muy grandes para el entrenamiento de la inteligencia artificial, pero las imágenes individuales de estos conjuntos están mínimamente curadas y, por lo tanto, a menudo no son adecuadas para la investigación en psicología y neurociencia. Para superar estos problemas, los trabajos recientes han creado grandes conjuntos de imágenes curadas que están diseñadas para estudiar las bases cognitivas y neuronales de la visión humana

Uno de ellos es THINGS (MN, et al., 2019), que es un conjunto de imágenes que contiene 1.854 conceptos de objetos que representan un amplio conjunto de conceptos nombrables utilizados en la lengua inglesa. Normalmente, los experimentos clásicos de visión de objetos presentan alrededor de una imagen por segundo, por ejemplo, 9-11. Recoger un ensayo para cada imagen en THINGS llevaría más de siete horas, lo que es inviable de conseguir con una sola sesión y requeriría un diseño complejo que implicara múltiples sesiones de escaneo. Sin embargo, recientemente (Grootswagers, Zhou, Robinson, Hebart, & Carlson, 2022) demostró que es posible descubrir información detallada sobre los estímulos visuales presentados en flujos de presentación visual rápida en serie (RSVP) utilizando la electroencefalografía (ver Fig. 4.6). Aquí se presentó un conjunto de datos de respuestas electroencefalográficas humanas a 50 personas, en total, se presentaron 26.248 imágenes visuales en una sesión de EEG de una hora. Para ello se utilizó un paradigma de presentación visual rápida en serie. Los resultados de la validación técnica indicaron que el conjunto de datos contiene respuestas neuronales detalladas a las imágenes, lo que demuestra que el conjunto de datos es un recurso de alta calidad para futuras investigaciones sobre la neurobiología del reconocimiento visual de objetos.

El cerebro humano logra el reconocimiento de objetos visuales mediante múltiples etapas de transformaciones no lineales que operan a escala de milisegundos. Para predecir y explicar estas rápidas transformaciones, los neurocientíficos computacionales emplean técnicas de modelado de aprendizaje automático. Sin embargo, los modelos más avanzados requieren cantidades masivas de datos para entrenarse adecuadamente, y a día de hoy se carece de vastos conjuntos de datos cerebrales que muestren ampliamente la dinámica temporal del reconocimiento de objetos visuales. En (Gifford, Dwivedi, Roig, & Cichy, 2022) se recolecta un amplio y rico conjunto de datos de respuestas EEG de alta resolución temporal a imágenes de objetos sobre un fondo natural. Este conjunto de datos incluye 10 participantes, cada uno con 82,160 ensayos que abarcan 16,740 condiciones de imagen, en la Fig. 4.7 se presenta la estructura del experimento realizado. A través de un modelo computacional establecimos la calidad de este conjunto de datos de cinco maneras. Se entrenan modelos de codificación lineal que sintetizaron con éxito las respuestas del EEG a imágenes arbitrarias. Se demuestra que tanto el elevado número de condiciones como las repeticiones de ensayos del conjunto de datos de EEG contribuyen a la precisión de predicción de los modelos entrenados. Este

artículo publica un conjunto de datos como una herramienta para fomentar la investigación en la neurociencia visual y la visión por ordenador.

RSVP design

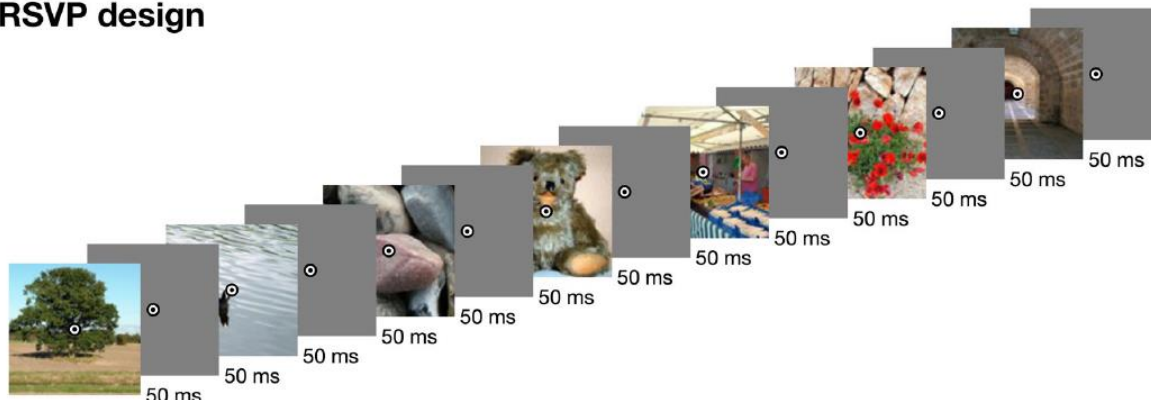


Figura 4.6. Diseño de presentación visual en serie rápida. Se presenta la secuencia de los objetivos visuales que observarán las personas para el proceso de identificación de conceptos (Grootswagers, Zhou, Robinson, Hebart, & Carlson, 2022).

En (Marini, Breeding, & Snow, 2019) se presenta una colección de datos EEG registrados de 24 observadores (14 mujeres, 10 hombres, edad media: 25,4) mientras observaban estímulos presentados individualmente compuestos por 96 objetos del mundo real, y 96 imágenes de los mismos artículos impresas en alta resolución. El EEG ha sido preprocesado, segmentado en épocas no solapadas, y se ha realizado un análisis de componentes independientes (ICA) para rechazar los artefactos. Además, se han completado pasos de preprocesamiento suplementarios para facilitar el análisis de los potenciales relacionados con eventos (ERP). Estos resultados demuestran que los objetos del mundo real desencadenan respuestas cerebrales más fuertes y sostenidas relacionadas con la acción que las imágenes. Estos resultados demuestran que los objetos del mundo real desencadenan respuestas cerebrales más fuertes y sostenidas relacionadas con la acción que las imágenes.

En el estudio descrito en (Calvo, et al., 2018) se estudiar la relación semántica entre pares de sustantivos de objetos concretos como “caballo-oveja” y “columpio-melón” y cómo esta relación de actividad se refleja en las señales de EEG. Los experimentos realizados fueron diseñados con el objetivo de encontrar evidencia de relación semántica entre dos palabras, en señales electroencefalográficas (EEG) tomadas con un dispositivo EEG comercial (EMOTIV Epoc).

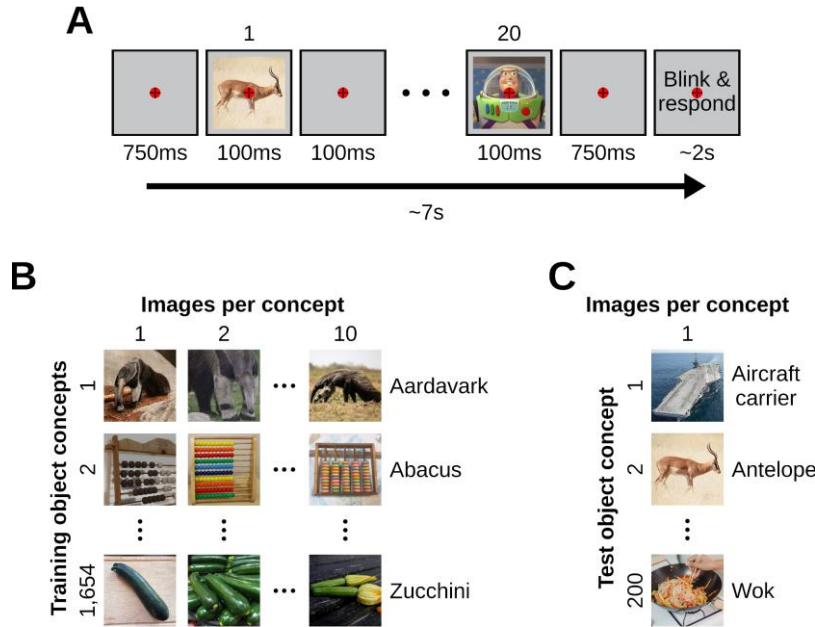


Figura 4.7. Paradigma experimental e imágenes de los estímulos. Se presentan a los participantes imágenes de objetos sobre un fondo natural utilizando un paradigma RSVP.

Los principales avances conceptuales ofrecidos por estos hallazgos son que existe un patrón neuronal identificable asociado con la percepción y la contemplación de objetos individuales, y que, dependiendo del tipo de estímulo, parte de este patrón se comparte entre varios participantes. Este patrón neural se caracteriza por una distribución de activación a través de muchas regiones corticales, que involucra ubicaciones que codifican varias propiedades de objeto. El análisis realizado proporciona información sobre objetos representados visualmente.

4.3. Identificación de categorías de objetos a partir de fMRI

Estudios previos que utilizaron un enfoque de aprendizaje automático han podido identificar los estados cognitivos asociados con la visualización de una categoría de objeto, como las casas (Simanova, Gerven, Oostenveld, & Hagoort, 2010), (Patterson, Nestor, & Rogers, 2007), (Mahon & Caramazza, 2009), (Gerlach, 2007), (Martin A. , 2007), (Binder, Desai, Graves, & Conant, 2009), (Kanwisher & Yovel, 2006), (Chand, Lamichhane, & Dhamala, 2016). La característica central de este enfoque (en comparación con un enfoque similar al mapeo paramétrico estadístico convencional) es su identificación de un patrón multivariado de vóxeles y sus niveles de activación característicos que identifican colectivamente la respuesta neural a un estímulo.

Estos métodos de aprendizaje automático tienen el potencial de ser particularmente útiles para descubrir cómo se representa la información semántica sobre los objetos en la corteza cerebral porque pueden determinar la distribución topográfica de la activación y distinguir el contenido de información en varias partes de la corteza.

El análisis multivariado de patrones es una técnica que permite la decodificación de información conceptual, como la categoría semántica de un objeto percibido a partir de datos de neuroimagen. Se han informado resultados impresionantes de clasificación de un solo ensayo en estudios que utilizaron fMRI

En (Shinkareva, et al., 2008) analizan estudios anteriores y concluyen que se ha logrado identificar el estado cognitivo correspondiente a la percepción de un conjunto de categorías representadas, como las herramientas, mediante el análisis del patrón que acompaña a la actividad cerebral, medido a través de imagen por resonancia magnética funcional (fMRI por sus siglas en inglés). Esta investigación se centró en identificar el estado cognitivo asociado con la visualización en 4 segundos de un dibujo de línea individual (1 de 10 objetos familiares, 5 herramientas y 5 viviendas, como un martillo o un castillo).

Aquí se demuestra la capacidad para:

1. Identificar de manera confiable cuál de los 10 dibujos estaba viendo un participante, según los patrones de activación neuronal del cerebro completo característicos de ese participante, excluyendo las áreas visuales.
2. Identificar la categoría del objeto con una precisión aún mayor, según la activación de ese participante (ver Fig. 4.8).
3. Identificar por primera vez, tanto los objetos individuales como la categoría del objeto que el participante estaba viendo, basándose solo en los patrones de activación de otros participantes. Los vóxels importantes para la identificación de la categoría se ubicaron de manera similar entre los participantes y se distribuyeron en todo el córtex, enfocados en áreas de percepción temporal ventral, pero también incluyeron áreas de asociación frontal (y algo lateral izquierda). Estos hallazgos indican la presencia de estados neuronales estables, distribuidos, comunales e identificables correspondientes a los conceptos del objeto.

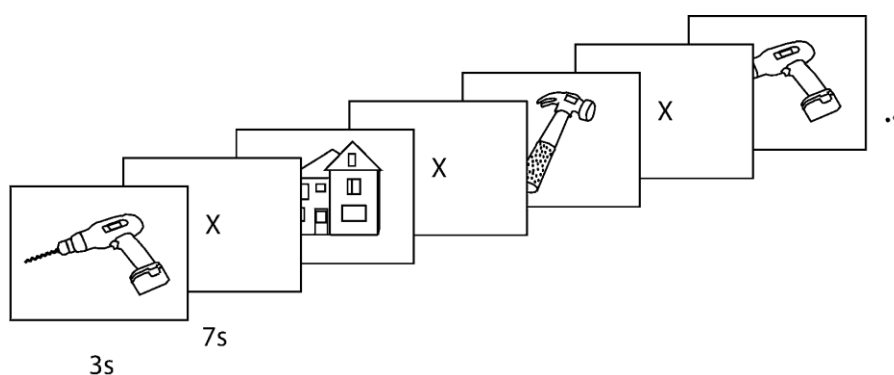


Figura 4.8. Representación esquemática del tiempo de presentación de cada concepto (Shinkareva, et al., 2008).

En (Noppeney, Price, Penny, & Friston, 2006) los patrones neuronales asociados con objetos individuales, así como con categorías de objetos se identificaron utilizando un algoritmo de aprendizaje automático aplicado a la activación distribuida por toda la corteza. Este estudio también investigó el grado en que los objetos y las categorías se representan de forma similar neuronal en diferentes personas.

Se logró entrenar a los clasificadores para identificar cuál de los diez ejemplos de objetos y dos categorías de objetos estaban viendo un participante. Se descubrió un patrón neuronal común entre los participantes, y se usa para entrenar a un clasificador para identificar la categoría de objeto correcta y el ejemplo de objeto de los datos de fMRI de nuevos participantes que no participaron en el entrenamiento del clasificador.

En (Fernandino, et al., 2015) se realiza una investigación basada en el cuestionamiento de si es posible que un modelo de codificación basado en cinco atributos semánticos directamente relacionados con la experiencia motora sensorial puede predecir con éxito los patrones de activación cerebral provocados por conjuntos de palabras. Los resultados muestran por primera vez que los patrones de activación distribuidos, asociados con un concepto combinan información sobre diferentes atributos sensoriales motores según su respectiva relevancia. Indica que un concepto léxico no se representa de manera idéntica en diferentes cerebros, sino que es un reflejo de la vida única de cada participante.

En (Moseley, Rachel, Pulvermüller, & Friedemann., 2014) se realiza un análisis de categorías léxicas en el cerebro, mediante procesamiento digital de imágenes para la detección de regiones del cerebro que se activan cuando el usuario lee sustantivos o verbos abstractos y concretos, con el objetivo de clasificarlos en categorías semánticas.

Los modelos neurocognitivos estándar de procesamiento del lenguaje han tendido a obviar la necesidad de incorporar procesos emocionales, mientras que las teorías de neurociencia afectiva generalmente se han preocupado por la forma en que las personas comunican sus emociones y, a menudo, simplemente no han abordado los problemas lingüísticos. (Hinojosa, et al., 2019) Resume la evidencia de los estudios de imágenes cerebrales temporales y espaciales que han investigado los efectos emocionales sobre los aspectos léxicos, semánticos y morfosintácticos del lenguaje durante la comprensión de palabras y oraciones individuales. La evidencia revisada sugiere que la emoción está representada en el cerebro como un conjunto de características semánticas en una red sensorial, motora, del lenguaje y afectiva distribuida. Además, la emoción interactúa con una serie de características léxicas, semánticas y sintácticas en diferentes regiones y tiempos cerebrales.

La recuperación de información conceptual de imágenes de acción provoca una mayor activación que de imágenes de objetos de forma bilateral en áreas de movimiento humano, como en el área media temporal (MT) y área media superior temporal (MST) y regiones temporales cercanas. Por el contrario, la recuperación de información conceptual de las palabras de acción provoca una mayor activación en las circunvoluciones temporales media izquierda y superior, anterior y dorsal al MT / MST.

4.4. Un enfoque de aprendizaje profundo de extremo a extremo para la clasificación de señales EEG para BCI

Dentro de los enfoques dominantes para las tareas de clasificación EEG se encuentra el algoritmo de patrones espaciales comunes (CSP) y variaciones del mismo, como patrones espacio-espectrales comunes (CSSP), banco de filtros de CSP (FBCSP), o CSP de complejo de transformación sin correlación fuerte (SUTCCSP). Estos algoritmos extraen características de las señales de EEG, que luego se clasifican mediante algoritmos de

aprendizaje automático supervisados, como máquinas de vectores de soporte (SVM). Para mejorar los resultados, a menudo se aplica un procesamiento previo especializado, como el filtrado lineal, la eliminación de artefactos y el rechazo de prueba (Hauke, Jakob S., Helle K., & Sadasivan, 2018).

Si bien los métodos de CSP se centran en gran medida en la selección de frecuencia, recientemente se demostró que tener en cuenta la variabilidad temporal mediante la selección de tiempo-frecuencia puede ser beneficioso para la precisión y puede ayudar a reducir la cantidad necesaria de canales de EEG (Yang, Chevallier, Wiart, & Bloch, 2017). Esto requiere nuevos métodos que detecten características dependientes del tiempo en las señales de EEG.

Los métodos de aprendizaje profundo (DL) están conquistando muchos dominios al superar a los métodos convencionales en una multitud de problemas. Recientemente, DL también se aplicó a la clasificación EEG. (Kumar, Sharma, Mamun, & Tsunoda, 2016) Sugirió reemplazar clasificadores de uso común como SVM por perceptrones multicapa (MLP) mientras se mantienen los mecanismos especializados de extracción de características. En (Bashivan, Rish, Yeasin, & Codella, 2016) utilizaron CNN para clasificar las señales de EEG a través de mapas de topografía espectral generados a partir de grabaciones de transformada de Fourier de corta duración (STFT). Y en (Tabar & Halici, 2017) utilizaron mapas de tiempo-frecuencia de STFT como entrada a una CNN con autocodificadores apilados (SAE) que alcanzaron una precisión muy alta en comparación con los métodos de referencia.

Los autores (Hauke, Jakob S., Helle K., & Sadasivan, 2018) desarrollan e implementan un enfoque de aprendizaje profundo (DL) para un sistema BCI basado en imágenes motoras (MI) basado en electroencefalograma (EEG) que podría usarse para mejorar las estrategias actuales de rehabilitación de accidentes cerebrovasculares. El modelo DL usa capas de red neuronal convolucional (CNN) para aprender características generalizadas y reducción de dimensiones, mientras que una capa convencional totalmente conectada (FC) se usa para la clasificación. Juntos construyen un modelo unificado de extremo a extremo que se puede aplicar a señales de EEG sin procesar. Este modelo propuesto previamente se aplicó a un nuevo conjunto de datos para validar su robustez frente a variaciones de datos. Además, se amplió mediante una adaptación específica del tema. Por último, un análisis de los filtros aprendidos proporciona información sobre cómo dicho modelo deriva una decisión de clasificación.

Si bien los enfoques aplicados mencionados anteriormente involucran preprocesamiento, como la extracción de características o STFT para mapeo de frecuencia de tiempo, en (Shen, Lu, & Jia, 2017) se propuso un enfoque de DL de un extremo a otro que utiliza CNN y memorias a corto plazo (LSTM) para clasificar los datos de EEG sin procesar. (Tang, Li, & Sun, 2017) Desarrollaron una arquitectura CNN con capas convolucionales unidimensionales para clasificar datos EEG sin procesar. Un modelo similar fue desarrollado por (Schirrmeister R. T., y otros, 2017) con etapas de entrada de CNN para filtros espaciales y temporales separados que arrojan resultados excepcionales a pesar de una arquitectura muy simple y poco profunda.

Los nuevos resultados sugieren fuertes límites a la viabilidad de clasificar la actividad del cerebro humano evocada a partir de estímulos de imagen, medida a través de EEG. Múltiples trabajos como (Greene & Hansen, 2018), (Fares, Zhong, & Jiang, 2018) intentan replicar estos experimentos con ensayos aleatorios en un conjunto de datos relativamente grandes (de 30 a 40 categorías de objetos) implementando CNN. (Ahmed, Wilbur, R, M, & Siskind, 2020) Replica dichos experimentos pero con un conjunto de datos mucho más grande (20x), de 1,000 presentaciones de estímulo de cada una de las cuarenta clases, todas de un solo sujeto. Hasta donde se sabe, este es el mayor esfuerzo de recopilación de datos de EEG de un solo sujeto y está en los límites de la viabilidad. En este artículo se recopilan grabaciones de EEG de 40,000 presentaciones de estímulo a un solo sujeto. Entrenan 5 clasificadores en el conjunto de datos completo de 1,000 presentaciones de estímulo de cada una de las 40 clases, utilizando una validación cruzada de 5 veces y se demuestra que solo el SVM y el 1D CNN, arrojan una precisión estadísticamente significativa por encima del azar.

Existen diversas complejidades dentro del procesamiento de señales de EEG. Entre ellas, se puede hacer referencia a: la inevitable presencia de ruido e interferencia en los sensores, la baja resolución espacial de los auriculares BCI, la posible presencia de valores atípicos en las mediciones, la dificultad de recopilar suficientes datos de ensayos para entrenamiento (Fabien L. , Signal processing approaches to minimize or suppress calibration time in oscillatory activity based Brain Computer Interfaces, 2015), la necesidad de determinar el número adecuado de características en esos métodos que se aplican a la reducción de dimensionalidad y la no estacionariedad de las señales EEG.

En las interfaces cerebro-computadora (BCI), los modelos típicos de las observaciones de EEG generalmente conducen a una pobre estimación de las matrices de covarianza de prueba, dada la alta no estacionariedad de las fuentes de EEG. Los autores (Olias, M. Clemente, Sarmiento V, & Cruces, 2019) proponen la aplicación de dos técnicas que mejoran significativamente la precisión de estas estimaciones y se pueden combinar con una amplia gama de métodos BCI (MI-BCI) de imágenes motoras. Las pruebas experimentales revelan que una combinación de las técnicas propuestas con los algoritmos de última generación para la clasificación de imágenes motoras proporciona una mejora significativa en los resultados de clasificación.

A pesar de que se proponen muchos modelos de aprendizaje profundo para la comprensión del contenido o el reconocimiento de patrones de las actividades cerebrales a través de EEG, la clasificación de objetos basada en esta metodología todavía exige esfuerzos para mejorar con respecto a su precisión, generalización e interpretación, sin embargo, algunos caracteres del cerebro humano aún no han sido explorados en su totalidad. En el artículo que presentan (Fares, Zhong, & Jiang, 2018) proponen un marco de aprendizaje profundo bidireccional a nivel de región para la clasificación de objetos basada en EEG. Se basan en la lateralización hemisférica del cerebro humano, proponen extraer información adicional a nivel regional para fortalecer y enfatizar las diferencias entre dos hemisferios (ver Fig. 4.9). La memoria bidireccional a largo y corto plazo se utiliza para capturar las correlaciones dinámicas ocultas tanto del pasado como del futuro con el estado actual en la secuencia de EEG. Se llevan a cabo extensos experimentos y los resultados demuestran la eficacia del marco propuesto.

Comparado con el trabajo existente, este marco logra un desempeño sobresaliente en la tarea de clasificación de objetos basada en EEG.

Una euforia común en la comunidad es que grandes conjuntos de datos han permitido que los métodos de aprendizaje profundo resuelvan de cierta manera todo el proceso. Sin embargo, parece haber alcanzado su límite con la clasificación de objetos a partir de registros de EEG. Ni cantidades grandes de datos, en los límites de la viabilidad, ni las arquitecturas estándar de aprendizaje profundo de redes completamente conectadas (MLP), redes neuronales convolucionales (CNN) o redes neuronales recurrentes (LSTM), o incluso métodos de aprendizaje automático más tradicionales como los clasificadores de vecino más cercano (k-NN) o las máquinas de vectores de soporte (SVM), parecen adecuados para la tarea. En los múltiples trabajos presentados exponen sus datos utilizados, métodos y está latente la tarea a la comunidad como un desafío para ir más allá de los grandes conjuntos de datos y el aprendizaje profundo, hacia la verdadera comprensión de la percepción visual humana.

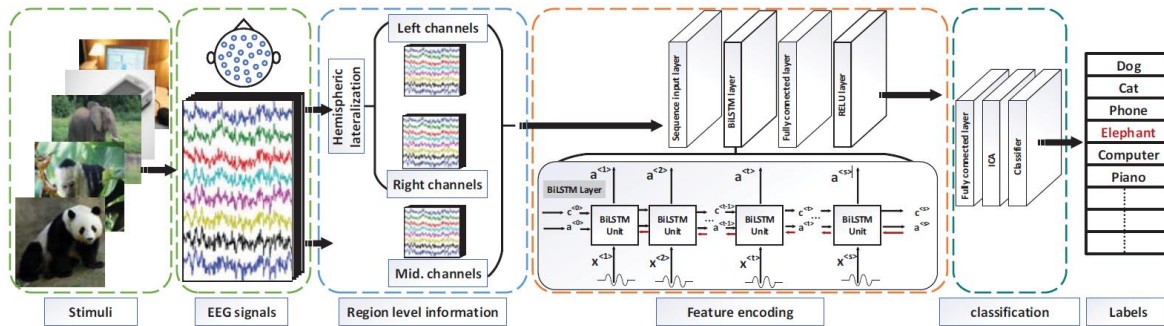


Figura 4.9. Marco de aprendizaje profundo bidireccional a nivel de región para la clasificación de objetos visuales. El enfoque consta de tres etapas, incluida la etapa de extracción de información a nivel de región, la etapa de codificación de características y la etapa de clasificación.

Basado en el análisis del estado del arte, a continuación se presenta en las tablas 4.1 y 4.2 la distribución de los métodos más implementados en la literatura para la categorización de objetos.

Tabla 4.1. Principales métodos de análisis, filtros y procesos de extracción de características en la literatura para el proceso de categorización de objetos.

	Clasificación semántica	Método	Filtros	Extracción
(Fares, Zhong, & Jiang, 2018)	40 categorías (Objetos y animales)	EEG	Rechaza bandas (49-51 Hz) y un filtro Butterworth de paso de banda de segundo orden (14-71 Hz).	ICA
(Ahmed, Wilbur, R. M, & Siskind, 2020)	40 categorías (Objetos y animales)	EEG	Pasa bandas de 1 – 40 Hz	
(Greene & Hansen, 2018)	30 Categorías	EEG	Pasa bandas de 0.1 – 45 Hz	
(Parekh, Subramanian, Roy, & C, 2018)	Objetos de la vida cotidiana	EEG	Pasa bandas de 0.1 – 45 Hz	ICA
(Asly, Gilde, Moctezuma, & Molinas, 2019)	3 colores (RGB)	EEG	Pasa bandas de 0.1 – 30 Hz	FFT, STFT, EMD Empirical mode decomposition (EMD)
(Mohedano, et al., 2014)	Objetivo/ No objetivo	EEG	Pasa bandas de 0.1 – 70 Hz	ICA
(Shinkareva, et al., 2008)	Herramientas y construcciones	fMRI	Pasa altas de 190 s	PCA
(Grill-Spector & Malach, 2004)	Animales y herramientas	EEG	Pasa bandas 1–30 Hz	
(Murphy, et al., 2011)	Mamíferos y herramientas	EEG	Filtro de pasa bandas de 1–120 Hz para eliminar desviaciones lentas en la señal y ruido de alta frecuencia, y luego muestreado a 300 Hz	CSP
(Helbig, Graf, & Kiefer, 2006)	Sustantivos y verbos	fMRI	Pasa altas de 128 s	
(Calvo, Paredes, Figueroa, & Nazuno, 2018)	Sustantivos	EEG	El filtro pasa bandas. Filtro de respuesta finita (FIR) con una frecuencia de corte más baja de 20 Hz y una frecuencia de corte alta de 1 Hz.	LPC, PCA, ICA, SEGN FDTW y CSP
(Simanova, Gerven, Oostenveld, & Hagoort, 2010)	Palabras escritas (acciones y objetos)	fMRI		
(Chand, Lamichhane, & Dhamala, 2016)	Rostros y casas	EEG	Transformaciones rápidas de Fourier (FFT), pasa bandas de 1-100 Hz y filtro de ruido digital a 60 Hz	

(Mohammad, Mitra, & Kavous, 2012)	Categorías animadas e inanimadas	EEG	FIR con ventana Hamming con 0.1 Hz-100 Hz, y reducido a 250 Hz.
(Behroozi, et al., 2015)	Animado vs inanimado, caras y cuerpos, cuerpos humanos vs cuerpos no humanos, rostros humanos vs no humanos		
(Grootswagers, et al., 2018)	Imágenes animadas e inanimadas	EGG	FIR con ventana Hamming 0.1 Hz-100 Hz
(Spiridon & Kanwisher, 2002)	12 categorías de diferentes objetos (animales, flores, partes del cuerpo, etc.)	EEG	FIR con ventana Hamming 0.1–150 Hz

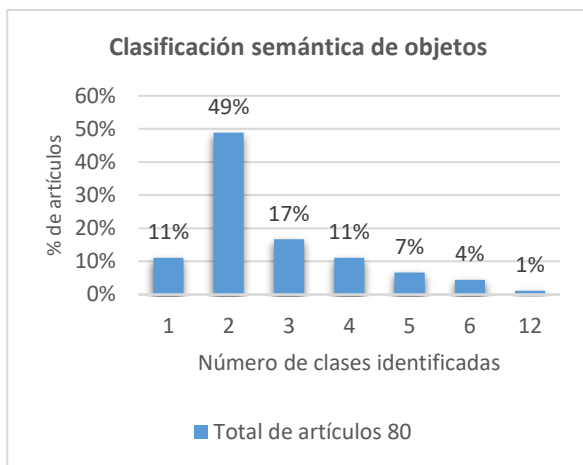
Tabla 4.2. Clasificadores mayormente implementados en la literatura, basados en el mecanismo neural medido y la región cerebral activada.

	Clasificador	Región cerebral estudiada	Mecanismo neural	Componentes analizados
(Fares, Zhong, & Jiang, 2018)	CNN, Clasificador SoftMax y clasificador de máquina de vectores de soporte multiclase (SVM).	Lateralización hemisférica, hemisferio izquierdo, hemisferio derecho y parte media		
(Ahmed, Wilbur, R, M, & Siskind, 2020)	kNN, SVM, una red neuronal de dos capas completamente conectada (MLP) y una CNN unidimensional		ERP	
(Greene & Hansen, 2018)	CNN	Occipital y temporal	ERP	P200
(Parekh, Subramanian, Roy, & C, 2018)	CNN		ERP	P300 Presentación visual rápida en serie (RSVP)
(Asly, Gilde, Moctezuma, & Molinas, 2019)	bosque aleatorio (RF), máquina de vectores de soporte (SVM), vecinos de knearest (kNN), árbol de decisión (DT) y Bayes ingenuo (NB)			
(Mohedano, et al., 2014)	SVM		ERP	Presentación visual rápida en serie (RSVP)
(Shinkareva, et al., 2008)	Clasificador de varianza agrupado gaussiano-Naive Bayes (GNB)	Hemisferio izquierdo, corteza premotora ventral y corteza parietal posterior, giro parahippocampal derecho, el lóbulo parietal superior (SPL)		
(Grill-Spector & Malach, 2004)		Parieto-occipital occipitotemporales	central, ERP	Ondas NI-P2, N400

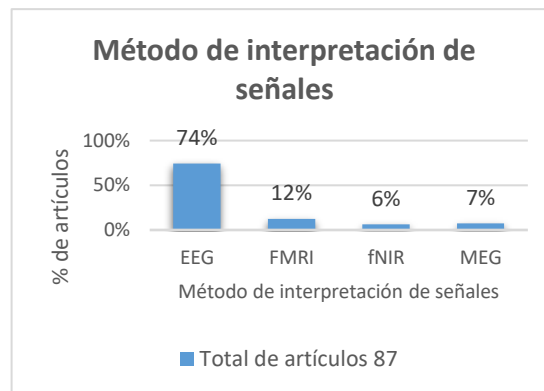
(Murphy, et al., 2011)	SVM		ERP	Ondas NI-P2
(Helbig, Graf, & Kiefer, 2006)		Corteza frontotemporal, corteza occipital inferior, corteza precetral, (IFG)		
(Calvo, Paredes, Figueroa, & Nazuno, 2018)	Árbol de decisión (DT), Naive Bayes (NB), Decisión (RL), Redes neuronales artificiales (ANN), Vecinos k-más cercanos (kNN) y Máquinas de vectores de soporte.		ERP	P300, N400
(Simanova, Gerven, Oostenveld, & Hagoort, 2010)		Corteza occipitotemporal lateral, área temporal media (MT) y área temporal superior medial (MST), corteza frontal izquierda, corteza occipitotemporal ventral, parieto-occipital central.		
(Chand, Lamichhane, & Dhamala, 2016)			ERP	
(Mohammad, Mitra, & Kavous, 2012)	LDA con el marco de análisis de similitud de representación (RSA).			
(Behroozi, et al., 2015)	SVM con el marco de análisis de similitud de representación (RSA)	Área visual primaria V1 y corteza temporal inferior (IT).	ERP	
(Grootswagers, et al., 2018)	LDA		ERP	N300, N400
(Spiridon & Kanwisher, 2002)	Clasificación Naive-bayesiana (NBC)			

Se realizó una clasificación de artículos leídos para fines de esta investigación, en la cual se grafica los principales métodos de interpretación de las señales cerebrales para el proceso de categorización de objetos (ver Gráfica 4.1) y se analiza que la electroencefalografía es de los métodos más utilizados para este tipo de procesos, debido a las características que esta tecnología representa, previamente descritas.

De igual manera se realiza una clasificación de los artículos en cuanto a la cantidad de clases semánticas que se clasifican en los diversos trabajos, donde la mayoría de casos de investigación realizan una clasificación binaria (ver Gráfica 4.2) de elementos completamente distantes, semánticamente hablando. En la tabla 4.3 se observa el tipo de clasificación semántica que más se implementa, se puede observar que las clases son completamente ajenas, es decir, que no están estrechamente relacionadas en cuanto a características físicas o visuales.



Gráfica 4.2. Número de clases identificadas en la literatura.

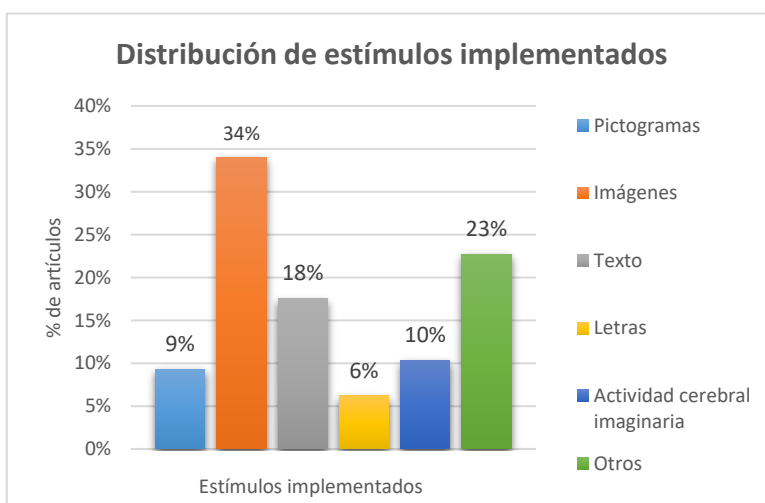


Gráfica 4.1. Métodos de interpretación de señales cerebrales.

Tabla 4.3. Campos semánticos más clasificados en la literatura.

Campos semánticos clasificados comúnmente en la literatura	
Herramientas y edificios	Animales, flores, partes del cuerpo, etc.
Animales y herramientas	Controlar un autómata (avance y paro)
Sustantivos y verbos	Asistencia remota (encendido y apagado de aparatos)
Caras y herramientas	Elementos vivos vs no vivos
Caras y casas	Letras
Categorías animadas e inanimadas	Selección de un objeto

Finalmente se presenta en la gráfica 4.3 una distribución de los principales estímulos presentados ante los usuarios para la clasificación de objetos, donde la mayoría presenta imágenes como estímulos de las categorías semánticas a clasificar.



Gráfica 4.3. Estímulos más utilizados en la literatura para el proceso de categorización de objetos.

Estímulos implementados	%
Otros	23%
Detección de emociones	6%
Luces parpadeantes	4%
Nivel de concentración	3%
Gestos faciales	10%

CAPÍTULO 5

5. Diseño metodológico

En este capítulo se habla de la metodología que será implementada para la solución de la problemática, y se describen las etapas de la investigación.

La investigación actual se centra en la comprensión de la codificación y procesamiento de la información en el sistema cerebral, cuantificando cómo este proceso se ve alterado ante la presencia de conceptos, y cómo se puede obtener esa información a través de la interacción con dispositivos artificiales incluidos BCI.

La investigación se divide en dos secciones, la primera es la parte de entrenamiento del sistema y la otra es el proceso de identificación automática de conceptos.

5.1. Entrenamiento

La etapa de entrenamiento ayudará al sistema a tener una correcta clasificación de las señales recibidas por el cerebro humano. Se comenzará de lo particular a lo general, se crea la herramienta adecuada para realizar el análisis de ondas cerebrales con una persona, en la que este usuario estará expuesto a ciertos conceptos expresados a través de estímulos, y observar que tipo de respuesta cerebral genera.

Se abordará más allá de saber que el cerebro generó una señal y esta se asocie al movimiento de un mecanismo o traducido a una acción, sino que es profundizar en el área, determinar si existe realmente un mecanismo de funcionamiento en el cerebro que este asociado o tenga estímulo ante los conceptos y que tan general es dentro de la población. A continuación se presentan las etapas del sistema.

1. **Selección de la población:** Se determinará la población objetivo sobre los cuales se van a trabajar para la estimación de los rasgos, a través del apoyo de expertos en el área para conocer cuál es el protocolo para realizar una selección adecuada de la población y de ejecución del experimento y obtener resultados satisfactorios.

2. **Creación de la base de datos:** la base de datos de conceptos, estará conformada por pictogramas universales, previamente recopilados, con imágenes muy generales, de modo que sean entendibles para cualquier usuario.
3. **Presentación de pictogramas a los usuarios:** con la selección previa de pictogramas adecuados, serán presentados al usuario para el proceso de entrenamiento, en esta etapa se entrenará al usuario para el uso adecuado del sistema, así como controlar la mayoría de variables que se presentan en los usuarios al utilizar en una BCI.
4. **Adquisición de ondas EEG:** en esta etapa se lleva a cabo la adquisición de las ondas cerebrales generadas por el usuario al percibir un concepto, esta información se puede obtener a través de la interacción con dispositivos artificiales incluidos BCI, en este caso particular una diadema no invasiva de la compañía EMOTIV (Emotiv, 2020). De igual forma se requiere la colaboración de expertos para conocer la manera correcta de colocar el dispositivo de adquisición de señales en la cabeza del usuario, y obtener señales EEG en ciertos lugares mediante electrodos de superficie.
5. **Preprocesamiento de la señal:** en esta etapa se realizará la limpieza de las señales con el objetivo de obtener datos con información lo más pura posible, a través del filtrado de la señal y la eliminación de ruido sobre esta misma. El procesamiento de señal es la labor más compleja en cualquier aplicación de EEG y su objetivo es decodificar el proceso neurofisiológico que refleja la intención del usuario. Las señales EEG en bruto son de una magnitud muy pequeña, por lo que pueden ser contaminadas por ruido e interferencias con facilidad. Es por ello por lo que el primer paso en el procesamiento de señal es limpiar la señal de ruido e interferencias. Para ello, se aplicarán filtros digitales que eliminan las bandas con frecuencias que se traduzcan como ruido para evitar que entorpezcan el procesamiento posterior. En caso de que la eliminación del ruido elimine información primordial de la señal, se necesitará balancear la señal y el ruido para obtener un buen preprocesado, dependiendo de los ritmos cerebrales que se obtendrán con la percepción de los pictogramas. Se estudiará la señal y se observará si requiere o no de algoritmos para eliminar de la señal el ruido producido por la actividad muscular del usuario. En esta etapa se valorará si es necesario crear una representación significativa de la señal original, reduciendo el tamaño de los datos de entrada sin perder información, o si se puede trabajar directamente con toda la información que otorgan las señales.
6. **Extracción de características:** En esta etapa se procede a la extracción de las características, donde se pretende discriminar la información más relevante y lograr la máxima eficiencia entre la relación señal-ruido. Seguidamente, se aplican métodos de selección de características que escogen las más significativas dentro del conjunto extraído, que codifican la intención del usuario. Posteriormente, se realizará una caracterización para el reconocimiento de patrones en la señal, a través de métodos matemáticos que ayuden a identificarla, existen ya varios tipos de filtrados matemáticos que se han vuelto básicamente un estándar en el estudio de señales estocásticas como las

obtenidas del EEG. Las interfaces más recientes utilizan varios algoritmos de procesamiento como: la transformada wavelet, la transformada de Hilbert, Hilbert-Huang, la transformada rápida de Fourier, etc. (Gómez, García, & Roque, 2016), mediante este tipo de métodos matemáticos se analizará la señal en intervalos de tiempo y observar qué información específica y destacada puede ser de utilidad.

7. **Generar base de datos de características:** una vez que se tenga el conjunto de características de cada señal y su etiquetado correspondiente, se genera una base de datos con esta información, con el fin de almacenar los valores y características discriminantes de las señales generadas correspondientes a cada estímulo presentado.

En la Fig. 5.1 se presenta un diagrama a bloques de la metodología propuesta.

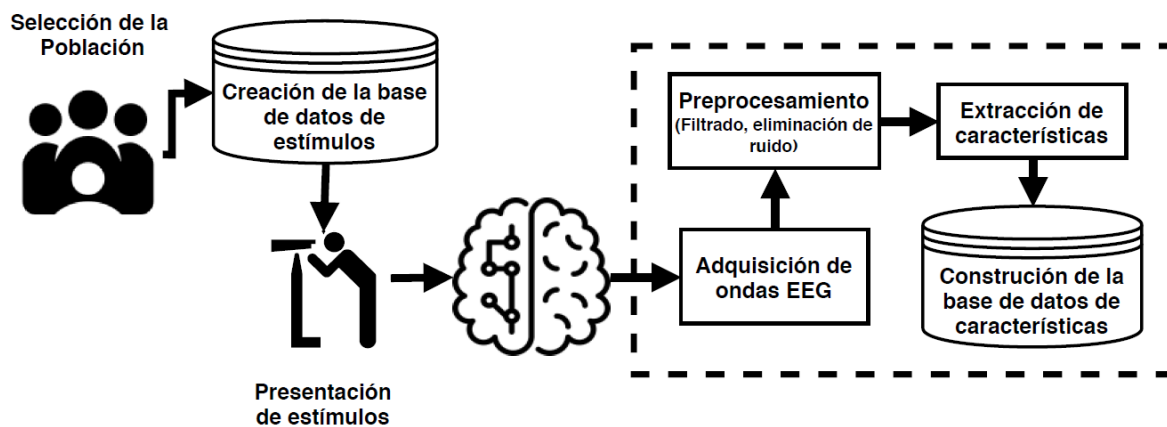


Figura 5.1. Diagrama a bloques de la metodología propuesta para la solución del problema de entrenamiento del sistema.

5.2. Identificación automática de conceptos

En esta vertiente del trabajo de investigación, el usuario hará uso del sistema de adquisición de señales y tiene que pensar en conceptos específicos, acciones o incluso que el usuario trate de generar oraciones fundamentales que incluyan acciones específicas y entendibles. El sistema recibe la onda EEG generada por el cerebro, atraviesa el proceso de filtrado de la señal, eliminación de ruido, extracción de características y esas características serán comparadas con una base de datos generada en el proceso de entrenamiento, con el objetivo de que el sistema identifique la acción específica que el usuario pensó y genere una respuesta de manera automática ante ese pensamiento, es decir, el sistema asociará la imagen correspondiente a tal acción pensada. A continuación se presentan las etapas del sistema.

1. **Generar el concepto:** en esta etapa el usuario tiene que pensar en una acción específica, ya sea un verbo, un sustantivo o incluso una frase con una combinación de ambos, con el fin de que el cerebro genere una señal correspondiente a este pensamiento.
2. **Adquisición de las señales.**

3. Preprocesamiento de la señal.

4. Extracción de características.

5. Clasificación: Una vez seleccionadas las características discriminantes en la señal, es necesario utilizar métodos de decisión que permitan diferenciar dichas características, algoritmos vinculados que permiten tomar un conjunto de datos pertenecientes a una muestra y darles alguna etiqueta que permita ordenarlo, darle alguna categoría o reconocerlo individualmente. En esta etapa se logrará que los algoritmos de clasificación traduzcan el conjunto de características seleccionadas en un comando concreto, relacionado con la intención el usuario. El sistema será capaz de comparar las características extraídas de la señal generada por el pensamiento del usuario, con la información almacenada en la base de datos que se generó en el proceso de entrenamiento, verificar si existe información similar con alguna señal y asociarla a esta. Con las técnicas de clasificación se consigue obtener un modelo con la finalidad de predecir futuros comportamientos ante nuevos casos y suponer que mientras más datos de entrenamiento disponibles, la confiabilidad del modelo obtenido es mayor.

6. Validación: En esta etapa se verifica el porcentaje de precisión del sistema, analizar si efectivamente al aplicar el proceso a distintas personas se obtienen porcentajes altos de coincidencia ante un mismo concepto, verificar que efectivamente lo que el usuario ve a través de los pictogramas se puede asociar a una onda EEG en específico con sus correspondientes características discriminantes.

7. Presentar resultados: en esta etapa el sistema sabrá si lo que el usuario pensó es un elemento dentro de la base de datos de estímulos, si es así presentará un resultado multimedia correspondiente a la acción pensada, o de lo contrario indicará que la acción no se encuentra en la base de datos o no la reconoce.

En la Fig. 5.2 se presenta el diagrama a bloques de la metodología requerida para la identificación automática de conceptos.

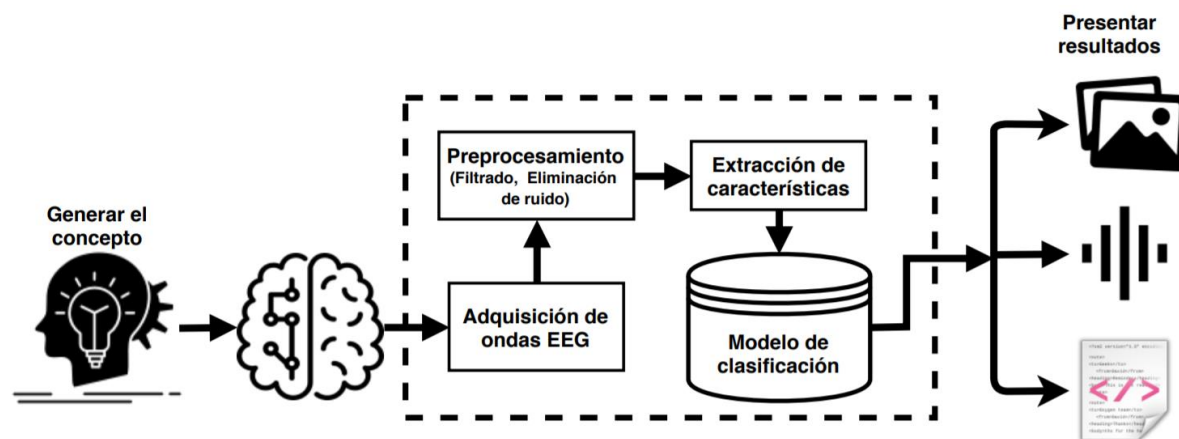


Figura 5.2. Diagrama a bloques de la metodología propuesta para la creación de un sistema de identificación automática de conceptos.

5.3. Población

Como punto de partida para la selección de la población, se analizó el estado del arte, en el cual se encuentran proyectos relacionados con la categorización de objetos a través de señales EEG, en los cuales implementan paquetes de cNORM, que son paquetes para el entorno R para computación estadística que tiene como objetivo generar normas de prueba continuas en psicometría y biometría (Folstein & Monfaredm, 2019).

Con fundamento en estos análisis, se seleccionaron 12 voluntarios, con una edades de 20 a 35 años. Participantes con visión normal o corregida a normal.

La infraestructura con la que se dispone para llevar a cabo esta investigación se divide en 2 elementos, hardware y software, a continuación se presentan las características de cada elemento que se involucrará en este trabajo.

5.4. Hardware

El hardware disponible actualmente para la adquisición de señales EEG son 2 diademas de la compañía EMOTIV, en el modelo EMOTIV Epoc, la cual es una diadema no invasiva que incluye 14 sensores de lecturas EEG, distribuidos es una estructura de electrodos 10-20 (Emotiv, 2020). 1 diadema EMOTIV Insight que incluye 5 sensores de lecturas EEG. Con estos dispositivos se realizarán múltiples pruebas para observar la respuesta generada por los usuarios al recibir los estímulos, y conocer qué tan precisos son los datos a procesar. Posteriormente, cuando se logre identificar la región del cerebro que realiza el procesamiento de conceptos, se evaluará la pertinencia del uso de esta infraestructura, o en su defecto, implementar un dispositivo más robusto que permita la optimización de representar señales limpias y robustas en cualquier momento y en cualquier lugar, aunado a esto, su menor complejidad para usarse. En la tabla 5.1 se presentan las características principales de las diademas disponibles.

Para el procesamiento de las señales se dispone de una computadora con procesador Intel Core i7, con memoria RAM de 16 GB.

5.5. Software

EMOTIV ofrece dispositivos de interfaz cerebro-computadora que pueden combinarse con su software de interfaz cerebro-computadora llamado EmotivBCI. EmotivBCI se puede utilizar directamente para implementar una BCI en una computadora. También puede emparejarse con la plataforma gratuita de código abierto NodeRed, que interconecta las salidas BCI a muchos dispositivos de hardware externos compatibles para implementar el control mental directo sobre equipos del mundo real. EmotivBCI y NodeRed permiten a los desarrolladores crear aplicaciones BCI, y se encuentran de manera gratuita, lo que servirá de apoyo para este tema de investigación.

Para el procesamiento de señales se utilizarán softwares como Python, Java, MATLAB con los cuales se analizarán y filtrarán las señales. En la tabla 5.2 se presentan algunas de las interfaces de análisis y procesamiento de señales EEG implementadas en la literatura.

Tabla 5.1. Características principales de los dispositivos de adquisición de señales EEG de la empresa EMOTIV.

Características	EMOTIV Insight	Emotiv Epoc
Número de sensores	5	14
Nombre de los canales (Sistema 10-20)		AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2
Tasa de muestreo	128 muestras por segundo, por canal	128 o 256 muestras por segundo, por canal (configurado por el usuario)
Resolución		16 bits(14 efectivos) 1 LSB=1.95 uV.
Rango de Radio Frecuencia	Bluetooth® 4.0 (Low Energy)	Bluetooth® 4.0 (Low Energy)
Tiempo de ejecución de la batería	4 horas	12 horas
Frecuencia neurológica	0.5-43 Hz	0.16-43 Hz
Valores inmediatos derivados	Expresiones faciales, estados emocionales, nivel de atención y concentración	Expresiones faciales, estados emocionales, nivel de atención y concentración
Precio	300 USD	800 USD

Tabla 5.2. Lista alfabética de las contribuciones de software destacados en la literatura. Las características principales de todos los paquetes de software se detallan en términos de análisis de series de tiempo (desde el preprocesamiento básico de datos sin procesar hasta métricas de sistemas dinámicos), análisis de tiempo-frecuencia y descomposición, modelado directo MEG / EEG y / o modelado de fuente (ajuste de dipolo, formación de haces / filtros espaciales e imágenes de origen), medidas y modelos de conectividad funcionales o efectivos, análisis estadístico e inferencia, interfaz gráfica de usuario versus bibliotecas de secuencias de comandos, interoperabilidad con otros paquetes de software / la posibilidad de que los usuarios desarrollen complementos, el entorno informático y el sistema operativo (s) (OS) requerido, y características definitivas que caracterizan cada paquete de software Fuente: adaptado de (Sylvain, Karl, & Oostenveld, 2011).

Paquetes de software destacados	Análisis de series temporales	Análisis de frecuencia de tiempo	Modelado hacia adelante	Modelado de fuente	Conectividad funcional o efectiva	Análisis estadístico	Interfaz gráfica del usuario	Interoperabilidad (otros paquetes de software / complementos de usuario)	Entorno / SO	Características únicas
Paquetes todo en uno										
BrainStorm http://neuroimage.usc.edu/brainstorm	X	X	X	X		X	X	Complementos de usuario	Matlab / múltiple	Énfasis en la interacción del usuario a través de GUI o scripts
CARTOOL http://brainmapping.unige.ch/cartool	X	X	X	X		X	X		Windows	Evaluación cuantitativa de topografías EEG; segmentación de microestado
EEGLAB, SIFT, NFT, BCILAB & ERICA http://scn.ucsd.edu/wiki/EEGLAB	X		X	X	X	X	X	EEGLAB, FieldTrip; complementos de usuario	Matlab / múltiple	Gran comunidad de usuarios; kit de herramientas diversas: desde el preprocesamiento de datos hasta análisis en tiempo real
ELAN http://elan.lyon.inserm.fr	X	X				X	X	Matlab, SPM, FieldTrip, NUTMEG, etc.	C / múltiple compilado (máquina virtual)	Amplia gama de herramientas de preprocesamiento y análisis de señales para datos electrofisiológicos.
MEGS http://www.emegs.org/	X	X				X	X	SPM8; Complementos de usuario	Matlab / múltiple	Preprocesamiento de datos basado en GUI
FieldTrip http://fieldtrip.fcdonders.nl/	X	X	X	X	X	X		BESA, EEGLAB, LORETA, SPM8, y así sucesivamente; plugins de usuario	Matlab / múltiple	Desarrollo basado en la comunidad; documentación extensa; características en tiempo real

MNE http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofMNE.php	X		X	X			X	Matlab	C/C++ (Linux/Mac OS)	Gráficos 3D de alta calidad; computacionalmente eficiente
NUTMEG http://nutmeg.berkeley.edu		X	X	X	X	X	X	FieldTrip, SPM, OpenMEEG, ELAN, etc.	Matlab / multiple; Python	Múltiples técnicas avanzadas de generación de imágenes.
sLORETA http://www.uzh.ch/keyinst/loreta.htm	X		X	X	X		X		Windows	Gran comunidad de usuarios
SPM8 http://www.fil.ion.ucl.ac.uk/spm/	X	X	X	X	X	X	X	FieldTrip, BioSig; complementos de usuario	Matlab / múltiple	Herramientas de SPM y comunidad; modelado causal dinámico
PyEEG http://pyeeg.sourceforge.net/	X								Python, SciPy/múltiple	Dimensiones fractales; medidas de entropía y complejidad

CAPÍTULO 6

6. Resultados

Para que el ser humano pueda transmitir todo tipo de información o sus necesidades, requiere de la comunicación, la cual se presenta a través de signos para dar y recibir tales mensajes. Son signos cualquier palabra o elemento portador de significado. Los signos se organizan en lenguajes, que pueden ser verbales o no verbales. El lenguaje no verbal: no se utiliza la palabra para comunicar, sino otros medios: dibujos, iconos, números, fotos, sonidos. Algunos de estos lenguajes se han creado para hacer llegar el mensaje a un receptor universal de la forma más rápida y clara posible. El lenguaje verbal: a través de las palabras, que es la forma de expresión más completa para comunicarnos que existe.

En esta sección se presentan los resultados obtenidos de la metodología propuesta para la resolución de la identificación de objetos a través de las lecturas EEG, se explica paso a paso el procesamiento implementado para la manipulación de las señales y qué técnicas son aplicadas para su correcta clasificación.

6.1. La lengua española

De acuerdo con el principal diccionario de la lengua española, el Diccionario de la Real Academia Española, el español tiene alrededor de 100,000 palabras. Sin embargo, la realidad es mucho más complicada, ya que esto no incluye las diferentes formas de una misma palabra, una de las características más distintivas de la lengua española.

Al abordar la cuestión de cuántas palabras forman parte del vocabulario de la lengua española, es importante entender que, al igual que ocurre en el resto de idiomas, el español contiene un gran número de palabras arcaicas o inusuales, que no están presentes en absoluto en el vocabulario de un hablante medio. Incluso aquellos que hablan español con fluidez solo utilizan un porcentaje relativamente pequeño en comparación con el total de palabras del vocabulario español. Por lo tanto, una pregunta más lógica sería: ¿cuántas palabras utilizan de verdad los hablantes de español?

Hay muchas fuentes que afirman que, como media, alguien que hable español como lengua materna tiene un vocabulario activo en torno a las 10,000 o 15,000 palabras. No obstante,

esto solo incluye las palabras que utilizan al hablar, ya que el vocabulario pasivo (las palabras que reconocen y entienden, pero no usan) seguramente ascienda a aproximadamente unas 25,000 palabras.

Asimismo, es importante dejar claro que este no es el número de palabras que una persona ha de saber para hablar español en su vida diaria. Como es lógico, el vocabulario de los hablantes nativos es mucho más amplio que el de aquellos que han aprendido español como una lengua extranjera y lo hablan con fluidez.

De acuerdo con el principio de Pareto el 80% de los efectos provienen del 20% de las causas. Si se aplica esta teoría al uso del lenguaje y el vocabulario del español, es posible deducir que un pequeño porcentaje de palabras es responsable de una gran parte de las conversaciones. De hecho, la cifra sería mucho menor del 20%.

Algunos expertos en lingüística afirman que con manejar entre el 5 y el 7% de las palabras que conoce un hablante nativo se puede ser capaz de entender del 90 al 95% de las palabras que se utilizan en un texto estándar. Esta cifra del 95% es muy significativa, ya que es el umbral a partir del cual suele ser posible comprender el mensaje perfectamente si a nuestro dominio del idioma le sumamos el conocimiento del contexto. Si aplicamos esto a la lengua española, podemos concluir que un vocabulario pasivo de 1,250 palabras y un vocabulario activo de solo 750 palabras podría, en teoría, ser suficiente para hacer frente a la mayoría de situaciones del día a día (ver tabla 6.1) (Lingoda, 2019).

Tabla 6.1. Uso del vocabulario español

El diccionario de la RAE	Vocabulario pasivo	Vocabulario activo	5 y el 7% de las palabras que conoce un hablante nativo
88,000 palabras	25,000 palabras	10,000-15,000 palabras	Vocabulario pasivo de 1250 palabras y un vocabulario activo de solo 750 palabras

Sin embargo, 720 palabras siguen siendo una cantidad amplia para trabajar con estímulos, ya que es una base de datos con un número alto de estímulos, y se requiere reducir aún más, por ejemplo, como primera instancia se trabajará con pictogramas ya que traspasan las barreras del idioma, pero tener 720 pictogramas y presentárselo a los usuarios es un trabajo que requiere de mucho tiempo, sin mencionar que es agotador e incómodo para la población seleccionada, ya que implica concentración, estados de relajación y el entrenamiento para los usuarios es tardado, ahora hacerlo con 720 pictogramas diferentes es un trabajo sumamente complicado.

Al analizar la literatura disponible sobre el lenguaje humano, se encontró un lenguaje artificial, construido por la lingüista Sonja Lang que a continuación se presenta.

6.2. Lenguaje Toki Pona

Toki Pona es un lenguaje construido de oligoaislamiento, creado por la lingüista y traductora canadiense Sonja Lang como un lenguaje filosófico con el propósito de simplificar los

pensamientos y la comunicación. Se publicó por primera vez en línea en 2001 como borrador, y luego en forma completa en el libro Toki Pona: El lenguaje del bien en 2014. Una pequeña comunidad de hablantes se desarrolló a principios de la década de 2000. La comunidad creció después del lanzamiento del libro oficial y ha seguido creciendo desde su publicación. Si bien la actividad se realiza principalmente en línea en salas de chat, en las redes sociales y en otros grupos, ha habido algunas reuniones organizadas en persona durante la década de 2010.

La característica subyacente de Toki Pona es el minimalismo. Se centra en conceptos universales simples, haciendo uso de muy poco para expresar al máximo. El idioma tiene 120–125 palabras raíz y 14 fonemas que son fáciles de pronunciar en diferentes idiomas. Aunque inicialmente no se pensó como un idioma auxiliar internacional, puede funcionar como uno. Inspirado en la filosofía taoísta, el lenguaje está diseñado para ayudar a los usuarios a concentrarse en cosas básicas y promover el pensamiento positivo, de acuerdo con la hipótesis de Sapir-Whorf. A pesar del pequeño vocabulario, los hablantes pueden entenderse y comunicarse entre sí, basándose principalmente en el contexto y las combinaciones de varias palabras para expresar significados más específicos (Lang, 2014)

El diccionario de Toki pona sólo dispone de 123 palabras para comunicarse, algunas pueden ser a la vez verbos, sustantivos o adjetivos, dependerá de la posición que tome la palabra en la frase. Al ser un lenguaje tan compacto, algunas de las palabras incluidas contienen términos ambiguos, con múltiples significados, en la tabla 6.2 se presentan ejemplos de las palabras que contiene este diccionario (Lang, 2014).

Al tener conocimiento de este lenguaje, la base de datos de estímulos se redujo considerablemente, como se mencionó anteriormente, en primera instancia se está trabajando con pictogramas, estos fueron obtenidos de la base de datos orientada al proceso de comunicación humana en portal de ARASAAC, quien ofrece recursos gráficos y materiales para facilitar la comunicación de aquellas personas con algún tipo de dificultad en esta área. El catálogo de pictogramas contiene 16,801 pictogramas a color y 15,026 en blanco y negro, con tamaño de 500x500 píxeles cada uno y con formato png (ARASAAC., 2016).

El autor (Olivares, 2019) realiza una clasificación previa de los pictogramas de este portal, en la cual implementa un Etiquetado POS (Part Of Speech por sus siglas en inglés) del pictograma. También es importante etiquetar cada pictograma respecto a la parte de oración que corresponde (sustantivo, verbo, adjetivo, etc), en la Fig. 6.1 a) se aprecia la clasificación. De igual forma, en este trabajo se realiza una búsqueda de palabras ambiguas: De todos los pictogramas obtenidos del corpus utilizado, se realizó una búsqueda profunda y detallada de todos aquellos pictogramas que pudieran ser ambiguos, se creó una tabla de igual forma con un índice para cada palabra, así como su etiqueta POS, y las palabras que pueden diferenciar las palabras ambiguas (ver Fig. 6.1 b)).

Se separaron los pictogramas de conceptos y palabras individuales: Un pictograma puede representar una sola palabra (por ejemplo: casa, carro, libro, etc) ó una serie de palabras (por ejemplo: jugar al golf, cama de gato) se realizó un filtrado de palabras y conceptos dividiéndolo en dos tablas y se eliminaron los pictogramas repetidos (ver Fig. 6.2). Se localizaron 2817 conceptos.

Categorías	Cantidad de pictogramas por categoría
Adjetivo	1780
Adverbio	254
Artículo	12
Conjunción	36
Determinante	31
Número	226
Preposición	44
Pronombre	137
Sustantivo	8759
Verbo	2456

a)

Palabra	Cantidad de definición por palabra	POS
Apuntar	2	Verbo
Banco	3	Sustantivo
Carta	1	Sustantivo
Copa	2	Sustantivo
Gato	2	Sustantivo
Lima	2	Sustantivo
Lista	2	Adjetivo, Sustantivo
Pila	2	Sustantivo
Planta	2	Sustantivo
Radio	4	Sustantivo
Ruta	2	Sustantivo

b)

Figura 6.1. Subclases de taxonomía. b) Palabras ambiguas contenidas en el corpus de pictogramas

Tipo	Cantidad
Normal	12496
Concepto	4056
Ambiguas	38
Total	16590

a)

Tipo	Cantidad
Normal	7481
Concepto	2817
Ambiguas	11
Total	10309

b)

Figura 6.2. a) Clasificación del total de pictogramas en el corpus. b) Clasificación de pictogramas sin repetir localizados en el corpus (OLIVARES, 2019).

Tabla 6.2. Estructura del diccionario Toki Pona con pictogramas obtenidos del portal de ARASAAC correspondientes al significado

kon: del chino mandarín 空氣 /kōngqì	- sust; viento, aire, olor, alma - adj; aéreo, gaseoso	 
kule: del francés acadiano couleur	- sust; color, pintura - adj; colorido - verbo tr; colorear, pintar	 
kulupu: del tongolés kulupu; del inglés group	- sust; comunidad, sociedad - adj; comunitario, público, social	 
kute: del francés acadiano écouter	verbo tr; escuchar, oír	
lape: del holandés slapen	- sust; sueño - verbo intr; dormir, soñar - adj; durmiente	

Se realizó un formulario con el objetivo de saber qué tan genéricos son los pictogramas (ver Fig. 6.3), este formulario se envió a más de 160 personas con las instrucciones de escribir lo primero que pensarán al momento de observar el pictograma (ver Fig. 6.4), con un total de 173 pictogramas, y de esta manera obtener respuestas de diferentes personas, con una población de un rango de edades de 12 a 33 años.

Identificación del significado de pictogramas

Este formulario se utilizará para determinar el grado de consistencia de una expresión pictográfica y el significado esperado.

***Obligatorio**

Nombre del evaluador (si requiere anonimato no responda esta pregunta)

Tu respuesta

Edad *

Tu respuesta

Este formulario se utilizará para determinar el grado de consistencia de una expresión pictográfica y el significado esperado.

A continuación se solicita que se llene el siguiente cuestionario escribiendo lo primero que llegue a su mente cuando observa cada una de las imágenes.

Figura 6.3. Formulario realizado y enviado en línea a un público de más de 160 personas, para obtener respuestas de lo que significa para ellos cada pictograma.

Las respuestas recibidas fueron en muchos casos satisfactorias, en el sentido de que la mayoría de la población pensó en el mismo concepto al observar el pictograma, incluso utilizaron en ocasiones sinónimos, que al realizar una correcta clasificación de estas respuestas, generan un alto porcentaje de similitud al significado original del pictograma (ver Gráfica 6.1 y 6.2). Pero como era de esperarse, existen pictogramas muy ambiguos en los cuales las personas escribieron significados totalmente ajenos al significado, e incluso muchos significados distintos, casos en los cuales un mismo significado no llega ni al 20% de similitud (ver gráfica 6.3), es decir, que estos pictogramas no son adecuados para el experimento y para ser utilizados como estímulo de conceptos.

Precisamente este fue el objetivo de la encuesta, lograr identificar y seleccionar los pictogramas más concretos y de fácil entendimiento para la población. Al final de la encuesta se realizó una pregunta con el propósito de conocer si este formulario había presentado dificultad para responderlo, en la gráfica 6.4 se aprecia el grado de dificultad con el cuál la población respondió, en donde aproximadamente el 70% dijo que fue un cuestionario de baja dificultad para responder.

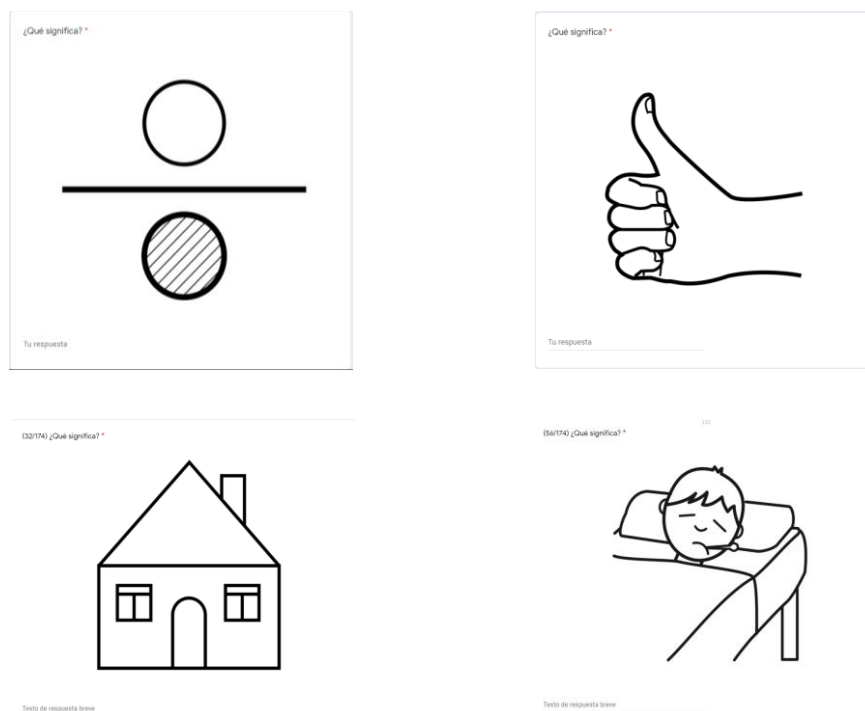
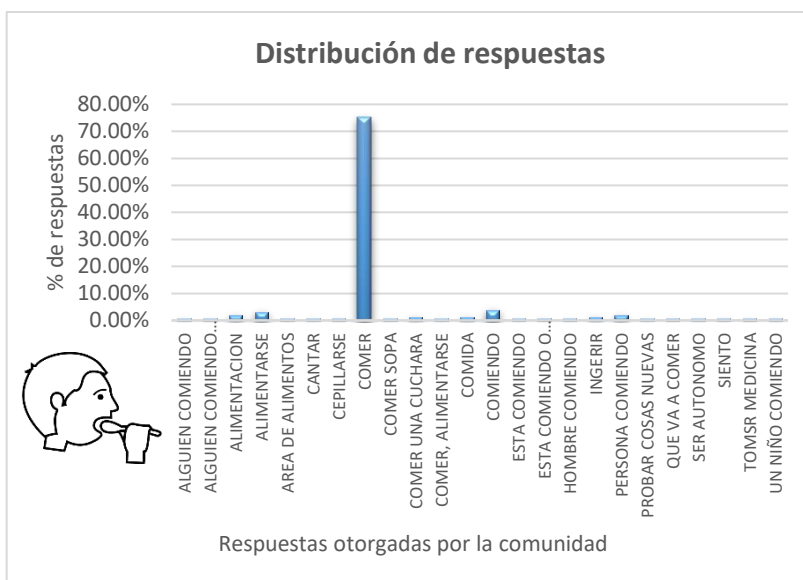


Figura 6.4. Ejemplos de pictogramas presentados a la comunidad, con el objetivo de saber que es lo primero que llega a su mente cuando observan los pictogramas.



Gráfica 6.1. Porcentaje obtenido al clasificar las respuestas de los usuarios para el pictograma de "comer".



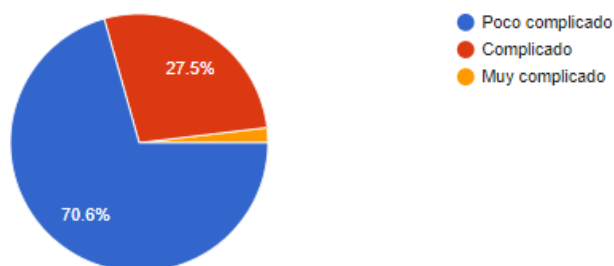
Gráfica 6.2. Porcentaje obtenido al clasificar las respuestas de los usuarios para el pictograma de "casa".



Gráfica 6.3. Porcentaje obtenido al clasificar las respuestas de los usuarios para el pictograma de "iniciar", en el que se presenta demasiada ambigüedad en el significado, las personas presentan confusión al momento de responder.

(174/174) ¿Qué tan complicado le resultó responder el cuestionario anterior basado en pictogramas?

160 respuestas



Gráfica 6.4. Nivel de dificultad que representó el cuestionario para la población.

A pesar de la reducción drástica del vocabulario humano a través del lenguaje Toki Pona y posterior a la encuesta para conocer el grado de dificultad de la comprensión de los pictogramas, se obtuvo un resultado reducido en cuanto al número de conceptos, aproximadamente 75 conceptos no ambiguos para la sociedad. Pero este número sigue siendo muy elevado para implementar un clasificador de conceptos eficiente, es necesario reducir el número de conceptos orientados al proceso de comunicación humana. Para esto se realiza la investigación de cuáles son las necesidades fundamentales de las personas, y se encontró que existe la teoría de las necesidades humanas planteada por Abraham Maslow. Maslow comprendió la obligación de conocer las necesidades del individuo, tales como; las fisiológicas, de seguridad, sociales, de estima y autorrealización; haciendo especial énfasis en percibir al hombre como un ser innovador ya que Maslow manifestaba que un “hombre autorrealizado” tiene una capacidad de aceptarse a sí mismo y a los demás, tiene una percepción superior de la realidad, es franco, tiene la capacidad para apreciar y la habilidad para tomar distintos puntos de vista frente al mundo, ser humanizado y sobre todo manifestar su independencia.

Aplicando esta teoría a un ambiente hospitalario, donde el personal de salud interactúa con personas enfermas o con discapacidad, se genera una nueva lista de necesidades observada por el área de enfermería, donde Virginia Henderson es una de las exponentes más importantes para la enfermería, ya que expone en su teoría que las personas tienen capacidades y recursos para lograr la independencia y la satisfacción de las 14 necesidades básicas, con el propósito de mantener una salud integral. En la Fig. 6.5 se presenta la relación de las necesidades básicas

Virginia Henderson desarrolló su modelo conceptual influenciada por la corriente de integración y está incluido dentro de la escuela de necesidades. Los componentes de esta escuela se caracterizan por utilizar teorías sobre las necesidades y el desarrollo humano para conceptualizar la persona, y porque aclara la ayuda a la función propia de las enfermeras. Incorpora a su concepto de enfermería, los procesos fisiológicos y psicopatológicos; además plantea la teoría de las necesidades humanas, que pueden verse afectadas por factores permanentes como el sexo, la edad, el medio socio-cultural o la capacidad física o por factores variables como las lesiones, estados de inconsciencia, enfermedades, cirugías, inmovilización, entre otras.

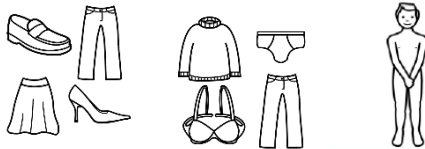
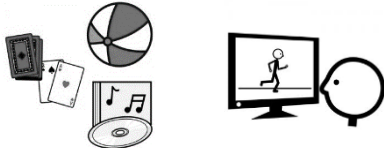
Con base en lo analizado en los artículos de investigación, donde se contrastara lo encontrado con las necesidades básicas expuestas por Virginia Henderson, se utilizarán el sistema pictográfico de comunicación. Para la elaboración de los signos, se utilizaron los propuestos en la página del Portal ARASAAC. De donde se escogieron treinta y ocho símbolos de acuerdo con las necesidades básicas del paciente expuestas por Henderson (ver Tabla 6.3) (Richard, 2012).

VIRGINIA HENDERSON	ABRAHAM MASLOW
1. Respirar con normalidad. 2. Comer y beber adecuadamente. 3. Eliminar los desechos del organismo. 4. Movimiento y mantenimiento de una postura adecuada 5. Descansar y dormir. 6. Mantener la temperatura corporal. 7. Mantener la higiene corporal.	NECESIDADES FISIOLÓGICAS
8. Evitar los peligros del entorno.	NECESIDADES DE SEGURIDAD
9. Comunicarse con otros, expresar emociones, necesidades, miedos u opiniones. 10. Ejercer culto a Dios, acorde con la religión.	NECESIDADES DE PERTENENCIA
11. Participar en todas las formas de recreación y ocio. 12. Seleccionar vestimenta adecuada.	NECESIDADES DE AUTOESTIMA
13. Trabajar de forma que permita sentirse realizado. 14. Estudiar, descubrir o satisfacer la curiosidad que conduce a un desarrollo normal de la salud.	NECESIDADES DE AUTORREALIZACIÓN

Figura 6.5. Relación de las necesidades básicas según Virginia Henderson.

Tabla 6.3. Necesidades básicas humanas.

Acción	Pictogramas
Respirar	
Comer y beber	
Eliminar los desechos del organismo.	
Moverse y mantener postura adecuada.	

Dormir y descansar	
Vestirse y desvestirse	
Mantener la temperatura corporal.	
Higiene corporal.	
Evitar los peligros del entorno.	
Comunicarse	
Creencias y religión	
Trabajar	
Recreación y ocio	
Aprender	

6.3. Diseño del experimento EEG

Una característica fundamental para el procesamiento de las señales EEG, es tratar de que al momento de ser recolectadas el usuario esté en un ambiente completamente aislado de elementos externos que puedan presentar ruidos a la señal, como los son distintas temperaturas, objetos visuales que representen distracción del usuario y este pierda de vista el objetivo, o que haga que el usuario se desconcentre. Un aspecto importante para eliminar estos datos no deseados es tener un espacio con las mismas condiciones siempre, para que el experimento se pueda replicar de manera muy similar en todas las ocasiones y con todas las personas de la población previamente seleccionada.

Un elemento más que permite la correcta recopilación de las señales EEG está representada por una “línea base” o “baseline”, mediante la cual se determina el proceso y los tiempos exactos en el que los estímulos serán presentados al usuario. En este caso, al trabajar con pictogramas el estímulo es visual, y se requiere una pantalla para presentarlos ante el usuario, un tiempo determinado de presentar cada estímulo para que el cerebro sea capaz de entender y clasificar lo que está viendo, y aunado a esto, un tiempo necesario para que la recopilación de información sea adecuada.

En la literatura existen múltiples líneas base reportadas, con las cuales se obtienen resultados satisfactorios en cuanto a la adquisición de las ondas cerebrales, y cada una de ellas está relacionada al tipo de objetivo que se requiere. En la tabla 6.4 se presentan ejemplos de la estructura del paradigma o diseño del experimento de distintos artículos, en los cuales los objetivos son la categorización de objetos. La mayoría de estos experimentos son similares, constan de una pantalla en blanco con un marcador en el centro (una “x”, un “+”, un “*” etc.) con el objetivo de hacer que el usuario no pierda la atención y se mantenga enfocado en el centro de la pantalla, es para preparar al usuario y evitar que este voltee a otros lados o pierda de vista el punto que se quiere presentar. Esta pantalla tiene una duración en milisegundos y posterior a esta, se presenta la pantalla con el estímulo objetivo, el cual se quiere presentar al usuario para su posterior categorización, de igual forma la duración de esta pantalla con el objetivo es de milisegundos. Y de esta manera se realizan las líneas base para la categorización de objetos, donde la primera pantalla se repite nuevamente antes del próximo estímulo objetivo. Hay trabajos que varían en el sentido de que posterior al objetivo, el usuario tiene que realizar una acción, como presionar un botón, para decidir si lo que vio corresponde o no a su objetivo inicial, ya que la mayoría de trabajos reportados realizan categorizaciones binarias, en los cuales comparan conceptos de campos semánticos muy alejados (animales vs casas, objetos vivos vs objetos no vivos, caras vs herramientas, etc.), y requieren de una señal que proporcione información adicional de que el usuario sabe que está viendo su objetivo. En estos casos, en la línea base se agrega una pantalla en blanco después del objetivo, para que el usuario pueda realizar la acción correspondiente.

Otros trabajos presentan una pantalla en blanco al inicio, con el fin de tratar de despejar la mente del usuario, y posterior a este la pantalla con el elemento de fijación al centro.

Tabla 6.4. Diseño de los experimentos o paradigmas más implementados en la literatura, para la presentación de estímulos visuales a una población determinada.

Pantalla en blanco inicial	Pantalla con cruz en el centro (Fijación)	Estímulo	Pantalla en blanco para pensar y realizar alguna acción	Fuente	Repeticiones
	500 ms	750 ms		(Blair, Marcos,, Hyung-Suk, Anthony, & Patrick, 2015)	
	500 ms	1000 ms		(Kyung, Hyo, & Hyun, 2005)	3 sesiones de 94 estímulos, 10 minutos de descanso entre sesiones
	1000 ms	3000 ms		(Regina & I-Ning, 2016)	6 sesiones para repetir los estímulos
	1200 ms	100 ms		(Andrei, et al., 202)	
	500 ms	2000 ms	500 ms	(Sabine, Matthias, & Ina, 2011)	
	500 ms	150 ms	1300 ms	(Asaid, Alan, Christoph, Theodor, & Jean-Marie, 2002)	
	600-900 ms	20 ms	1000 ms	(Hsin-Mei, Stephanie, Robert, & Wolfgang, 2011)	
300-700 ms		1000 ms		(Qingguo, Liangchao, & Shenyi, 2008)	
	750 ms	500 ms		(Blair, Marcos, Hyung-Suk, Anthony, & Patrick, 2015)	12 repeticiones de cada imagen, descansos después de cada 36 ensayos, 72 ensayos de 72 imágenes
	500-900 ms	80 ms	1000 ms		
		300 ms	1000-1200 ms	(Irina, Marcel, Robert, & Peter, 2010)	

En el análisis del diseño de un experimento apropiado se analiza el número de muestras que se les presenta a los usuarios, así como el número de usuarios disponibles, es decir, la población.

Se analiza el estado del arte, donde se demuestra que conforme el número de repeticiones de la presentación incrementa, se presenta un incremento proporcional (ver Tabla 6.5), así la presentación de diferentes estímulos de objetos visuales puede provocar cambios detectables en las grabaciones de EEG (ver Fig. 6.6). Las línea base estudiadas en la literatura, en la que

se presentan estímulos visuales, ya sea a través de imágenes o pictogramas, confirman después del procesamiento de datos, que estas presentaciones de estímulos visuales provocan un cambio potencial evocado visualmente, por lo que esta línea de estudio se basará precisamente en Potenciales evocados visualmente.

Tabla 6.5. Análisis de parámetros de evaluación, que son determinados por el tiempo disponible de la muestra (Zheng & Chen, 2021).

Tiempo de visualización	Exactitud
40–200 ms	87.60%
40–360 ms	90.78%
360–480 ms	95.82%

Fue necesario el estudio exhaustivo de la literatura para determinar el número adecuado de estímulos y la metodología que los autores implementan en caso multicategoricos, así como el tamaño de la población. Este análisis demuestra que conforme el número de presentaciones aumenta, es decir, que el número de pruebas que se realice con el sujeto sea mayor, la exactitud incrementa proporcionalmente.

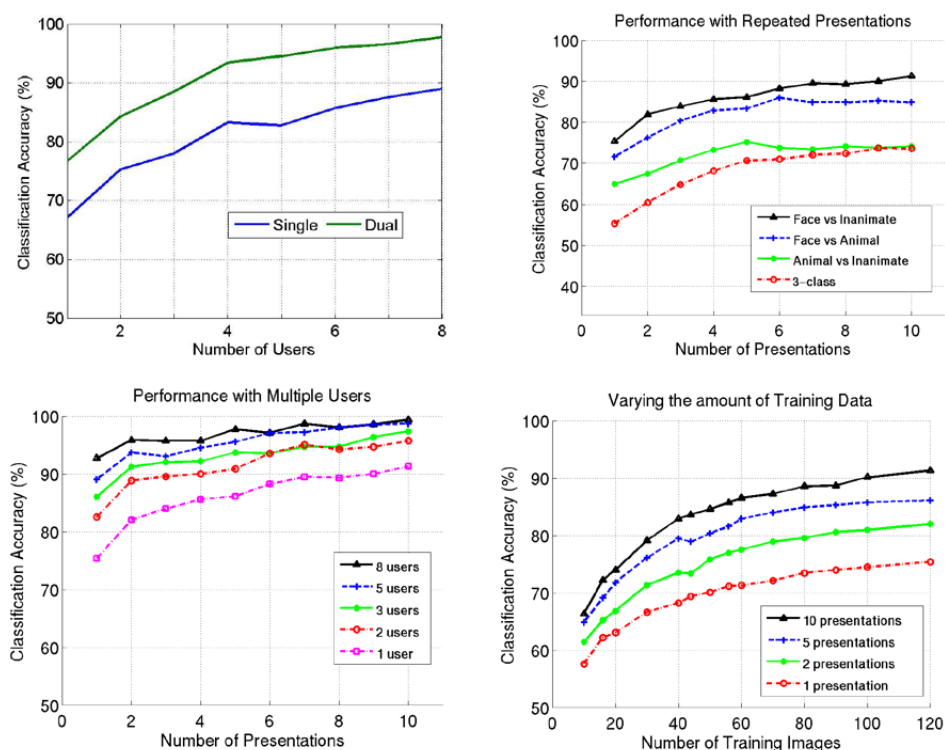


Figura 6.6. Alteración en el porcentaje de clasificación de señales EEG basado en el número de presentaciones repetidas, el número de sujetos y la cantidad de datos de entrenamiento.

La línea base propuesta como experimento inicial se basa en el trabajo (Blair, Marcos, Hyung-Suk, Anthony, & Patrick, 2015). Se realizará cada prueba con múltiples imágenes, con un fondo blanco que coincide con el fondo de la imagen. Como primer paso, se presenta

una pantalla con fondo blanco con una duración de 1 segundo para despejar la mente de la persona, posterior a esta se presenta una cruz de fijación negra en el centro de la pantalla blanca durante 500 ms, para minimizar los movimientos oculares durante y entre los ensayos, seguido del primer estímulo con duración de 750 ms, para un intervalo total entre pruebas de 1,250 ms. En la Fig. 6.7 se presenta el diseño del paradigma o de la línea base propuesta como experimento inicial.

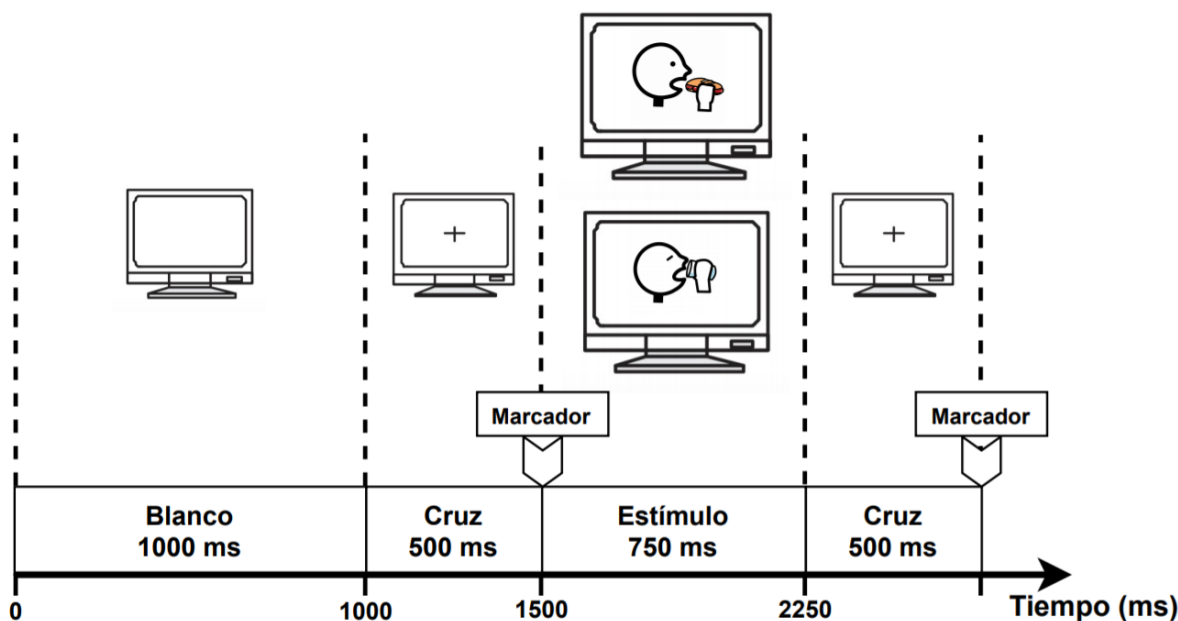


Figura 6.7. Diseño del experimento para la presentación de estímulos objetivos al usuario.

La base de datos de estímulos visuales de EEG incluye 14 categorías diferentes. Para la generación de las señales EEG se utilizarán 140 pictogramas como estímulos, donde cada categoría de imagen contiene la misma cantidad de ejemplos, específicamente, 10 imágenes de cada una de las siguientes seis categorías: Respirar, Comer y beber, Eliminar los desechos del organismo, Moverse y mantener postura adecuada, Dormir y descansar, Vestirse y desvestirse, Mantener la temperatura corporal, Higiene corporal, Evitar los peligros del entorno, Comunicarse, Creencias y religión, Trabajar, Recreación y ocio y Aprender.

En el experimento participarán 10 personas, de 20 a 40 años.

Cada prueba/ensayo consiste en una sola imagen, colocada sobre un fondo gris blanco que coincide con el fondo de la imagen, mostrada en pantalla durante 500 ms, seguida de 750 ms de una pantalla gris en blanco, para un intervalo total entre pruebas/ensayos de 1250 ms. Para minimizar los movimientos oculares durante y entre las pruebas/ensayos, se superpone una cruz de fijación blanca que abarca $0,76^\circ \times 0,76^\circ$ de ángulo visual en el centro de cada imagen y en la misma posición en la pantalla en blanco que se muestra entre las presentaciones de imágenes. Los estímulos se presentarán utilizando el software psychopy en bloques de 1,400 pruebas/ensayos, en los que cada imagen se mostrará doce veces en orden aleatorio. Se dieron breves descansos después de cada 70 pruebas/ensayo. Cada participante tendrá que completar dos sesiones experimentales, cada una con tres bloques, espaciados entre seis y ocho días. En

total, cada participante completó 60 pruebas/ensayos de cada una de las 140 imágenes, para un total de 8,400 pruebas/ensayos por participante.

6.4. Dispositivo para la adquisición de las señales EEG

La diadema Emotiv EPOC + avanzado por los sistemas Emotiv es uno de los dispositivos EEG desarrollado específicamente para aplicaciones BCI. Fue lanzado en el mercado de los EE. UU. En 2010 para abrir canales de comunicación directa entre el cerebro de un usuario y una computadora. Tiene 14 canales para cubrir la mayor parte del área de la corteza. El auricular Emotiv EPOC + es inalámbrico y el tiempo de preparación es corto. Antes de ponerse el dispositivo Emotiv EPOC +, los electrodos están equipados con una almohadilla de fieltro húmeda. Cuando las almohadillas aún no están húmedas, los electrodos se mojan con solución salina antes de insertarlos en el auricular. Con la diadema Emotiv EPOC +, no es necesario limpiar el cuero cabelludo antes y después del uso, por lo tanto, es mucho más ecológico en comparación con los de grado médico.

El panel de control de Emotiv muestra la señal inalámbrica y la recepción de energía de la batería y la calidad de contacto de cada sensor mediante un código de color. (Negro: sin señal, rojo: señal muy pobre, naranja: señal pobre, amarillo: señal regular, verde: señal buena). El auricular Emotiv EPOC + adquiere señales EEG de 14 canales de ubicaciones estándar del sistema internacional 10-20 (ver Fig. 6.8) con una frecuencia de muestreo de 128 Hz. Los canales de EEG son AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 y AF4. Referencias: P3/CMS, P4/DRL e incluye un giroscopio que registra los movimientos en el eje x e y (Emotiv, 2020).

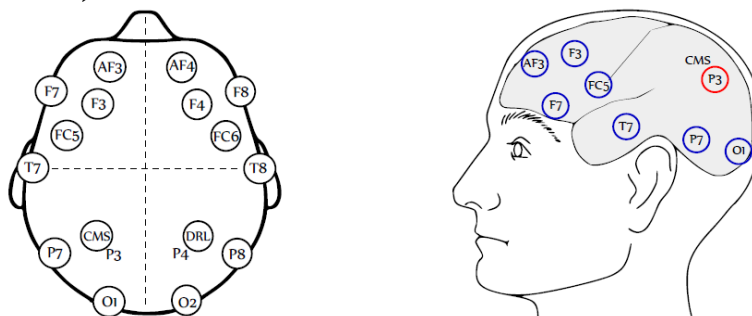


Figura 6.8. Localización de los electrodos en el kit EMOTIV (Alejandro, 2016)

El auricular Emotiv EPOC + se ha preferido en esta tesis por los motivos que se describen a continuación;

- Los sistemas BCI, basados en EEG, han sido ampliamente estudiados por investigadores de todo el mundo. Los amplificadores de EEG de grado médico son tan caros que, en particular, las personas con discapacidad no pueden pagar. El dispositivo Emotiv EPOC + es una de las alternativas de bajo costo que se pueden encontrar en el mercado.
- Otros amplificadores EEG de grado médico necesitan una conexión de cable entre cada electrodo, el amplificador y la computadora. Eso es inconveniente para el uso diario, y también presenta el problema de los artefactos creados por los movimientos de la cabeza.

- A diferencia de las aplicaciones médicas, no se requieren expertos para la recopilación de datos utilizando la diadema Emotiv EPOC +. Una persona puede recopilar sus datos de EEG fácilmente. El tiempo de preparación requerido para que el auricular Emotiv EPOC + esté listo para usar es de aproximadamente 2 minutos. En contraste, la mayoría del equipo médico requiere más de 20 minutos para la preparación.
- La calidad de los datos recopilados con el dispositivo Emotiv se ha probado en distintos estudios recientes, donde controlan sillas de ruedas, robots humanoides y diversos prototipos robóticos con resultados bastante aceptables (Apdullah, 2017).

6.5. Base de datos pública de señales EEG

Debido a la situación de la actual contingencia, que ha tardado más de dos años, resultó imposible concretar la recopilación de señales con usuarios para concretar la base de datos para los experimentos, por lo que se decidió trabajar con una base de datos de acceso libre en la literatura.

Después de una búsqueda en el estado del arte y la literatura de una base de datos de señales EEG que cuente con las características adecuadas para el trabajo aquí presente, se encuentra una base de datos de estímulos visuales de EEG que incluye 6 categorías diferentes. Para la generación de las señales EEG se utilizaron 72 fotografías de objetos reales como estímulos, donde cada categoría de imagen contiene la misma cantidad de ejemplos, específicamente, 12 imágenes de cada una de las siguientes seis categorías: cuerpo humano (CH), rostro humano (RH), cuerpo animal (CuA), rostro animal (CA), fruta vegetal (FV) y objeto inanimado (OI).

En el experimento participaron 10 personas, de 21 a 57 años (edad media 30,5 años; 3 mujeres; 1 zurdo). Todos los participantes informaron visión de color normal y visión normal o corregida a normal. Esta investigación fue aprobada por la Junta de Revisión Institucional de la Universidad de Stanford como parte del estudio titulado "Reconocimiento de ondas cerebrales del lenguaje, la música y las imágenes visuales mediante grabaciones de EEG". Se obtuvo el consentimiento informado por escrito de cada participante antes de su participación en las sesiones experimentales.

Cada prueba/ensayo consistió en una sola imagen, colocada sobre un fondo gris medio que coincidía con el fondo de la imagen, mostrada en pantalla durante 500 ms, seguida de 750 ms de una pantalla gris en blanco, para un intervalo total entre pruebas/ensayos de 1250 ms. Para minimizar los movimientos oculares durante y entre las pruebas/ensayos, se superpuso una cruz de fijación blanca que abarcaba $0,76^\circ \times 0,76^\circ$ de ángulo visual en el centro de cada imagen y en la misma posición en la pantalla en blanco que se muestra entre las presentaciones de imágenes. Los estímulos se presentaron utilizando el software Neurobehavioral Systems Presentation en bloques de 864 pruebas/ensayos, en los que cada imagen se mostró doce veces en orden aleatorio. Se dieron breves descansos después de cada 36 pruebas/ensayos; Se dieron descansos más largos entre bloques para comprobar las impedancias de los electrodos. Cada participante completó dos sesiones experimentales, cada una con tres bloques, espaciados entre seis y ocho días. En total, cada participante completó 72 pruebas/ensayos de cada una de las 72 imágenes, para un total de 5184 pruebas/ensayos

por participante. En la Fig. 6.9 se muestra la base de datos que se implementó como estímulos visuales.

El preprocesamiento se implementó con el apoyo de la caja de herramientas de código abierto EEGLAB y se realizó de la siguiente manera:

- El electroencefalograma denso se registró utilizando redes EGI HCGSN 110 de 128 canales sin blindaje.
- Los electrodos se montaron de acuerdo con el sistema estándar internacional 10-20.
- Los datos se muestrearon a la frecuencia de 1 KHz (Kaneshiro, M, H-S, AM, & P, 2015).

Images 1 – 12: Human Body



Images 13 – 24: Human Face



Images 25 – 36: Animal Body



Images 37 – 48: Animal Face



Images 49 – 60: Fruit Vegetable



Images 61 – 72: Inanimate Object

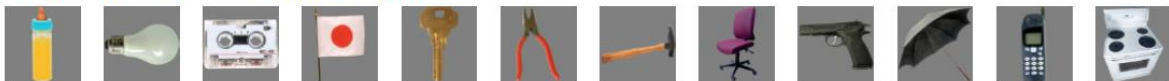


Figura 6.9. Estímulo visual presentado ante los usuarios (Kaneshiro, M, H-S, AM, & P, 2015).

6.6. Preprocesamiento de las señales EEG

La representación de las señales EEG obtenidas de la base de datos son procesadas en el software MATLAB, a través de una herramienta de código abierto denominada EEGLAB, que cuenta con múltiples paquetes que permiten examinar las señales de manera más práctica, con funciones predeterminadas que ayudan a generar procesos de filtrado, reducción de ruido y análisis de los componentes de la señal.

6.6.1. Implementación de filtros

Los filtros pueden diseñarse para tener una respuesta de impulso finita (FIR) o una respuesta de impulso infinita (IIR). Una respuesta de impulso es simplemente cómo el filtro maneja una señal de impulso unitario en el dominio del tiempo (lo que hace a la estructura de la señal) y su transformación de Fourier se conoce como la respuesta de frecuencia. Conocer la respuesta al impulso puede decirle lo que necesita saber para caracterizar el comportamiento del filtro.

La respuesta al impulso de un filtro es la forma de onda que aparece en la salida de un filtro cuando se presenta un pulso Dirac (pulso de unidad infinitamente corto) en la entrada.

Un filtro FIR tiene una respuesta de impulso por una duración finita, después de lo cual la salida pasa a cero y produce demoras iguales en todas las frecuencias (también conocida como respuesta de fase lineal). No es recursivo. La señal de salida $y[n]$ es una convolución de la señal de entrada $x[n]$ con la respuesta al impulso b_k del filtro.

$$y[n] = \sum_{k=0}^{n-1} b_k x_{n-k} \quad (6.1)$$

Ventajas del filtro FIR

- Un filtro FIR siempre dará una salida estable, ya que no es recursivo.
- Un filtro FIR tiene un retraso constante en todas las frecuencias; la forma de la señal no se verá influenciada por los cambios de fase

Desventajas del filtro FIR

- Un filtro FIR es ineficiente, utiliza muchas entradas para calcular la salida, por lo tanto:
- Un filtro FIR tiene un gran retraso

En contraste, los filtros IIR (también conocidos como filtros recursivos) tienen una respuesta de impulso infinita donde parte de la salida del filtro se usa como retroalimentación. Esto produce retrasos desiguales a diferentes frecuencias (características de fase no lineal). Esto significa que la señal de salida se desplaza en el tiempo con respecto a la entrada con algunos componentes de frecuencia desplazados más que otros (Singh, Veer, Sharma, & Kumar, 2016). La señal de salida $y[n]$ es una convolución de la señal de entrada $x[n]$ con la respuesta de impulso de longitud infinita b_k del filtro.

$$y[n] = \sum_{k=0}^M b_k x[n-k] - \sum_{k=0}^N a_k y[n-k] \quad (6.2)$$

Ventajas del filtro IIR

- Un filtro IIR es muy eficiente, utiliza solo unos pocos valores de salida anteriores para calcular la salida actual, por lo tanto:
- Un filtro IIR tiene solo un pequeño retraso

Desventajas del filtro IIR

- Un filtro IIR puede volverse inestable si los coeficientes no se eligen bien
- Debido al requisito de causalidad y una respuesta de impulso infinita que converge a 0 para $n \Rightarrow \infty$, la respuesta de impulso no puede ser simétrica. Esto implica una característica de fase que no es constante sobre las frecuencias. Por lo tanto, aparecerán diferentes frecuencias en la salida con un cambio de fase diferente. La forma de la señal de salida cambiará como consecuencia de diferentes cambios de fase.

En la Fig. 6.10 se presenta el comportamiento de 10 electrodos y su forma de onda de las señales con respecto del tiempo. Para observar el comportamiento de las señales adquiridas se grafica la densidad espectral de potencia de las señales, con el objetivo de identificar los ruidos externos generados en la señal a analizar. En la Fig. 6.11 se presenta la densidad espectral de potencia.

La elección del filtro está basada en un análisis previo de comparación de filtros y su eficiencia (ver Fig. 6.12), presentado en (Khatter, Bansal, & Mahajan, 2019), mediante la medición de distintos parámetros. Su análisis de rendimiento se ha realizado en matlab calculando el error cuadrático medio, el error absoluto medio, la relación señal / ruido, la relación señal / ruido pico y la correlación cruzada.

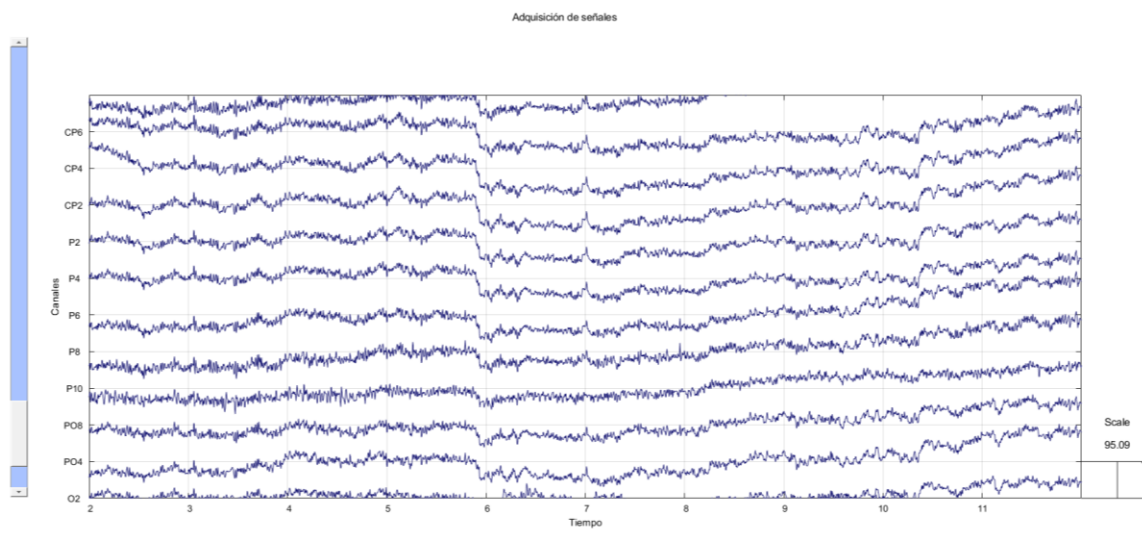


Figura 6.10. Señales EEG representadas en EEGLAB, generadas por los estímulos visuales presentados al usuario, donde las líneas verticales representan el momento en que el usuario concluyó de ver el estímulo.

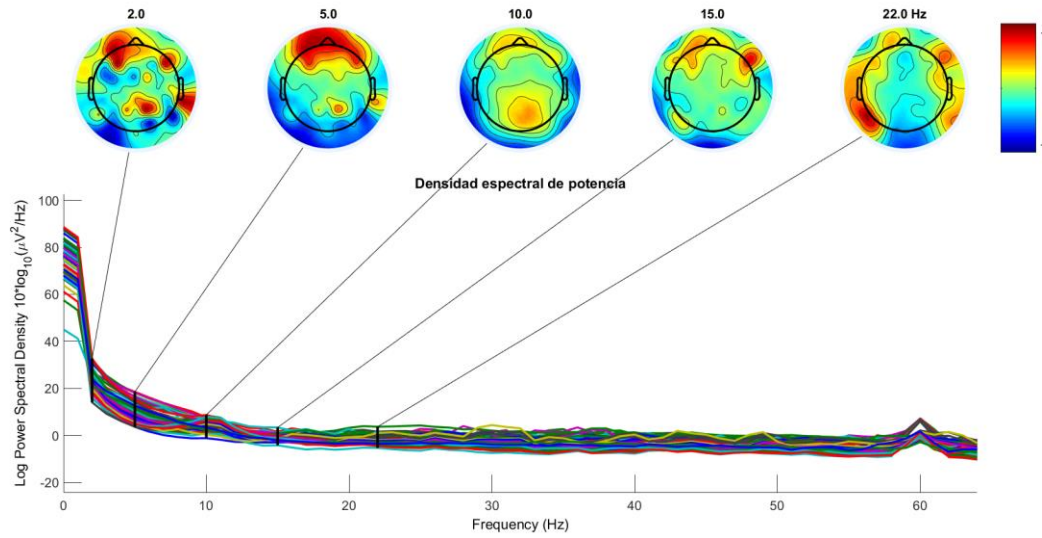


Figura 6.12. Densidad espectral de potencia de las señales EEG de todos los canales de la diadema utilizada.

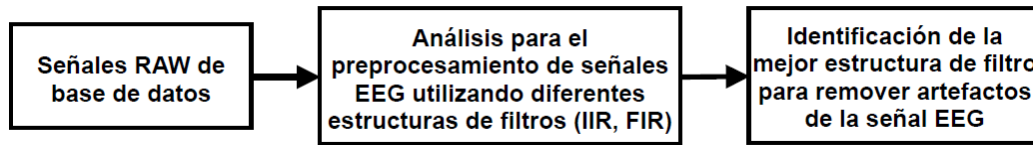


Figura 6.11. Diagrama de bloques del sistema de identificación de la estructura del filtro

Los siguientes parámetros se han analizado y propuesto para evaluar el rendimiento de las estructuras de filtro diseñadas para eliminar el ruido de las señales de EEG. Los parámetros se seleccionan para evaluar el rendimiento de los filtros IIR y FIR propuestos. El filtro que proporciona un error medio absoluto y un error cuadrático medio menores, una relación señal / ruido, una relación señal / ruido pico y una correlación cruzada más alta se considera un filtro eficaz para la eliminación de artefactos de las señales de EEG. Los siguientes parámetros de rendimiento se han calculado para evaluar el rendimiento de los filtros:

- (i) El error absoluto medio (MAE) es el error absoluto entre la señal reconstruida y la señal original y refleja el rendimiento de la estructura del filtro. MAE se utiliza como un indicador para saber cuánta divergencia hay en la señal reconstruida de la señal original.

$$MAE = \frac{\text{Señal original} - \text{Señal reconstruida}}{\text{Longitud de la señal original}} \quad (6.3)$$

- (ii) El error cuadrático medio (MSE) es el error cuadrático medio entre la señal original y la señal reconstruida. Es una medida de la fidelidad de la señal y proporciona la altitud de error / distorsión entre ellas

$$MSE = \left[\frac{(\text{Señal original} - \text{Señal reconstruida})^2}{\text{Longitud de la señal original}} \right] \quad (6.4)$$

- (iii) La relación señal / ruido se define como la relación entre la potencia de la señal y la potencia del ruido.

$$SNR(indb) = 10\log_{10}\left[\frac{Señal\ original^2}{(Señal\ original - Señal\ reconstruida)^2}\right] \quad (6.5)$$

- (iv) La relación señal pico a ruido se define como la relación entre la potencia de señal máxima posible y la potencia de ruido corruptor que afecta la fiabilidad de su representación.

$$PSNR = 10\log_{10}\left[\frac{(Valor\ absoluto\ máximo\ de\ la\ original)^2}{MSE}\right] \quad (6.6)$$

- (v) La correlación cruzada es la correlación entre la señal original y la reconstruida.

Basado en este análisis y en lo resumido en la tabla 3, los filtros más implementados y eficientes para el proceso de categorización de objetos con EEG con el fin de eliminar desviaciones lentas en la señal y ruido de alta frecuencia, son los filtros de respuesta finita (FIR) pasa bandas con frecuencias entre 1 y 100 Hz, con variaciones de uso de las ventanas de Hamming.

En la Fig. 6.9 se observa un pico pronunciado alrededor de los 60 Hz, que hace alusión al ruido de la línea. Para eliminar el ruido externo no deseado se diseña un filtro pasa bandas de respuesta al impulso finito, que permita eliminar las frecuencias bajas provocados por el desplazamiento DC, originado por la digitalización de las señales, y las frecuencias altas que se presentan por el ruido de la línea externa (60 Hz aproximadamente). En la Fig. 6.13 se presenta el filtro implementado, donde la frecuencia de corte del paso de bandas es de 0.1 Hz a 40 Hz.

Al aplicar el filtro diseñado a las señales originales, se observa que efectivamente el ruido de la red eléctrica es eliminado y únicamente permanece el espectro de frecuencia con el que deseamos trabajar (ver Fig. 6.14), de esta manera nos enfocamos en la banda de frecuencia de 0.1 a 40 Hz, donde se encuentran los datos relevantes de las señales EEG adquiridas.

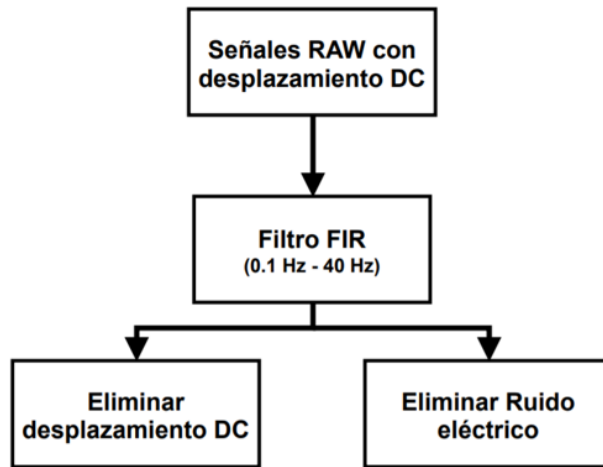


Figura 6.14. Diseño de filtro pasa bandas FIR para la eliminación de ruido indeseado.

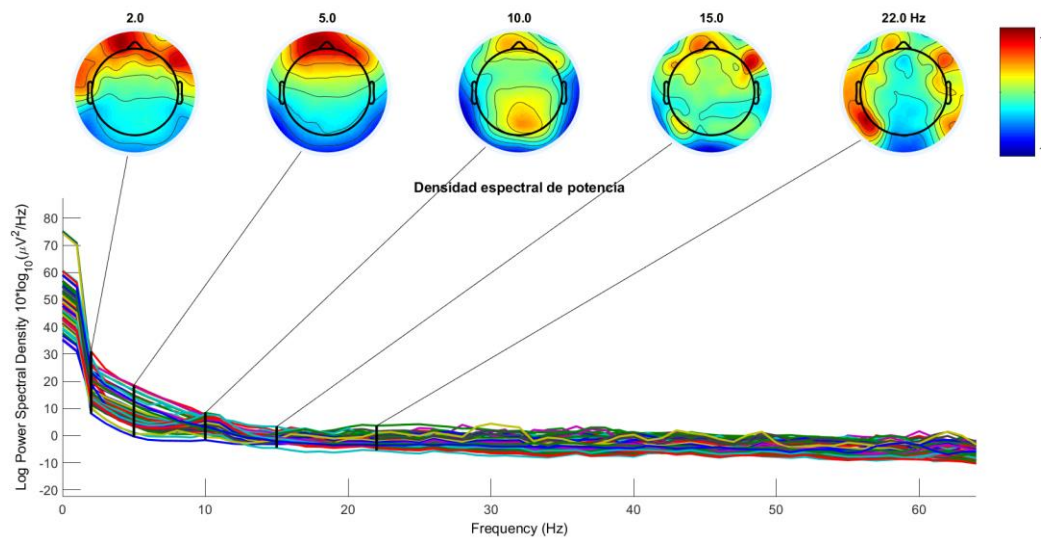


Figura 6.13. Densidad espectral de potencia. Se presentan las señales después de aplicar el FIR diseñado para eliminar el desplazamiento de DC y el ruido de la señal de la red eléctrica.

6.6.2. Eliminación de artefactos

Una dificultad principal en el análisis de señales EEG como se menciona en la sección 3.10 son los diversos artefactos que ocurren en la señal de EEG como movimientos de los ojos, parpadeos, actividad de los músculos, latidos del corazón, ruido de línea, alta impedancia de electrodos e intromisión de dispositivos eléctricos. Los artefactos se dividen en dos categorías: fisiológicos y no fisiológicos. Las principales causas de aparición de artefactos fisiológicos son que el paciente mueva la cabeza, parpadeo de ojos, sudoración, rotación del globo ocular, latido del corazón y contracción muscular, y las principales causas de aparición de artefactos no fisiológicos son fallas externas, por ejemplo, falla de electrodos, ventilación y suministro de energía. Como los artefactos oculares tienen la misma frecuencia que las

señales de EEG, por lo tanto, es problemático identificarlos, por lo que se requieren algoritmos competentes para eliminar los artefactos (Khatter, Bansal, & Mahajan, 2017).

Las grabaciones multicanal de los campos electromagnéticos que surgen de las corrientes neuronales del cerebro generan grandes cantidades de datos. Los métodos de extracción de características adecuados son, por lo tanto, útiles para facilitar la representación e interpretación de los datos. Se ha demostrado que el análisis de componentes independientes (ICA) desarrollado recientemente es una herramienta eficaz para la identificación y extracción de artefactos de registros EEG (ver Fig. 6.15) y uno de los algoritmos más utilizados en este proceso (Urigüen & Garcia-Zapirain, 2015).

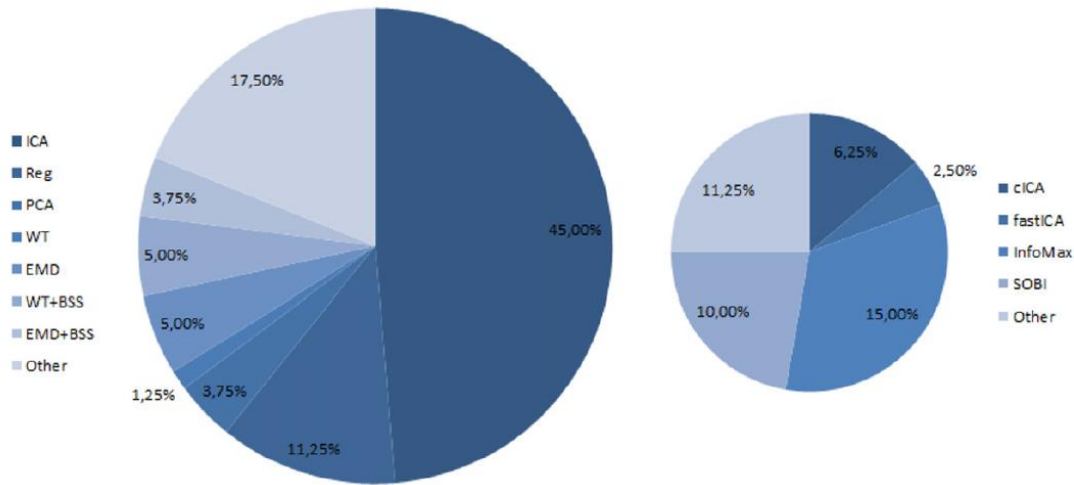


Figura 6.15. Porcentajes de tipos de algoritmos de eliminación de artefactos. El gráfico de la izquierda muestra la división de ciertas técnicas a las que se hace referencia en (Urigüen & Garcia-Zapirain, 2015). El gráfico de la derecha muestra la división de los algoritmos ICA.

ICA es un método de procesamiento de señal para separar fuentes independientes mezcladas linealmente en varios sensores. Es una técnica estadística novedosa que tiene como objetivo encontrar proyecciones lineales de los datos que maximicen su independencia mutua. Sus principales aplicaciones son la extracción de características y la separación ciega de fuentes (BSS por sus siglas en inglés). Se implementa en análisis de EEG para detectar artefactos en señales registradas originalmente. Por lo general, los artefactos oculares, musculares, electrocardiográficos y de líneas eléctricas se rastrean con el uso de este método. Es un hecho bien conocido que en el caso de EEG multicanal, ICA mejora la calidad de la señal (Delorme, EEGLAB, 2020).

Como en muchas otras transformaciones lineales, se supone que en el instante de tiempo k el vector de datos n -dimensional observado, $[x_1(k), \dots, x_n(k)]^T$ está dado por el modelo:

$$x(k) = \sum_{i=1}^m a_i s_i(k) = As(k) \quad (6.7)$$

Se supone que las señales fuente $s_i(k), \dots, s_m(k)$, son estacionarias, independientes y, junto con los coeficientes de la matriz de mezcla $A = [a_1, \dots, a_m]$, desconocidos. El objetivo

es estimar ambas incógnitas a partir de $x(k)$, con supuestos adecuados sobre las propiedades estadísticas de las distribuciones de fuentes. La solución se busca en la forma:

$$\hat{s}(k) = \mathbf{B}x(k) \quad (6.8)$$

donde $\mathbf{B}$ se llama la matriz de separación.

El problema general de BSS requiere que $\mathbf{A}$ sea una matriz $n \times m$ de rango completo, con $n \geq m$ (es decir, hay al menos tantas mezclas como fuentes independientes). En la mayoría de las derivaciones algorítmicas, se supone un número igual de fuentes y sensores. Además, solo hasta una fuente puede ser gaussiana.

Mientras que PCA simplemente correlaciona las salidas (usando una matriz ortogonal $\mathbf{B}$) ICA intenta hacer las salidas estadísticamente, sin colocar restricciones en la matriz $\mathbf{B}$ (ver Fig. 6.16 y 6.17). (Delorme, EEGLAB, 2020).

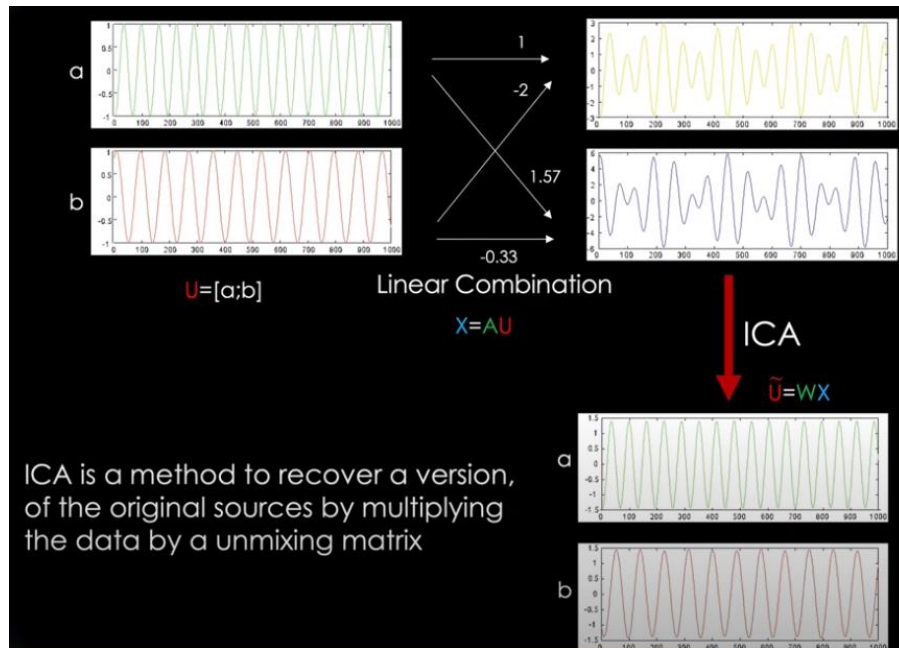


Figura 6.16. Método ICA para desacoplar señales EEG y trabajar sobre las fuentes originales. (Delorme, EEGLAB, 2020)

Cuando un artefacto tiene una distribución estable en el cuero cabelludo, generalmente es posible utilizar ICA para descomponer los datos del EEG en un conjunto de componentes subyacentes y luego reconstruir los datos sin el componente correspondiente al artefacto. Esto elimina efectivamente el artefacto del EEG. Este método de corrección de artefactos se ha vuelto muy popular para eliminar los artefactos de parpadeo, que tienen una distribución del cuero cabelludo muy estable y son tan grandes que generalmente aparecen como un solo componente ICA.

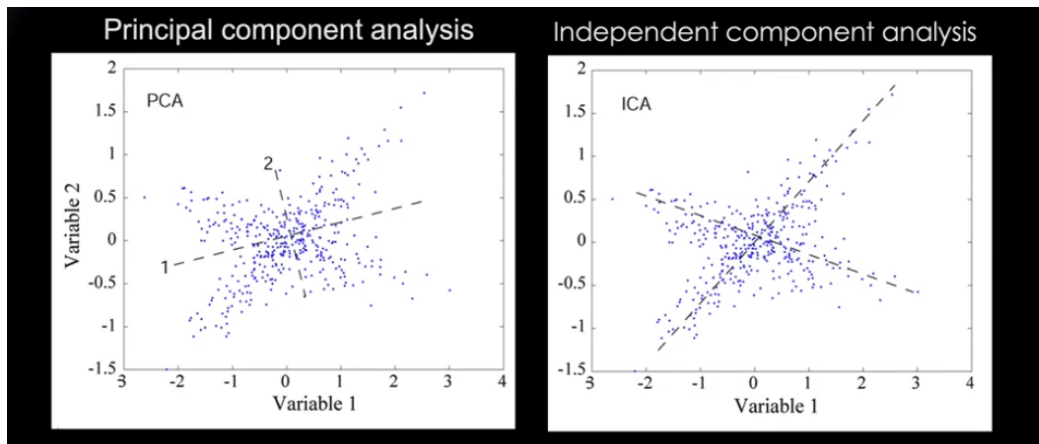


Figura 6.17. Comparación entre ICA y PCA, donde ICA intenta hacer las salidas estadísticamente, sin colocar restricciones en la matriz W (Delorme, EEGLAB, 2020).

Sin embargo, la corrección de artefactos basada en ICA tiene algunas limitaciones importantes. Primero, ICA no puede funcionar para artefactos que no tienen una distribución consistente en el cuero cabelludo para un sujeto dado (por ejemplo, potenciales cutáneos). En segundo lugar, el número de muestras de tiempo necesarias para que ICA aísle eficazmente los componentes es una función del cuadrado del número de electrodos, y los estudios con un gran número de electrodos pueden carecer del número de muestras de tiempo necesarias. En tercer lugar, el número de componentes ICA es siempre igual al número de electrodos, y el hecho de que el número de fuentes es mayor que el número de electrodos introduce algún error, aunque este error puede ser insignificante en muchas condiciones reales. En cuarto lugar, aunque varios estudios han validado la eficacia de ICA para la corrección del parpadeo, hay menos estudios de validación para otros tipos de artefactos. Finalmente, los parpadeos y los movimientos oculares no solo crean voltajes artificiales; también cambian la entrada sensorial. Por ejemplo, la entrada sensorial cambia enormemente si los ojos se cierran cuando aparece un estímulo. Los cambios en la posición de los ojos también tienen un impacto sustancial en la visibilidad de un estímulo visual (Lopez & J.Luck, 2014).

EEGLAB contiene sus propios algoritmos de detección de artefactos, incluido un método para inspeccionar visualmente los datos y marcar segmentos que contienen artefactos (o desmarcar segmentos que fueron marcados por los algoritmos de detección automática de artefactos). Para el caso particular de esta tesis, se implementan las herramientas que ofrece EEGLAB para la identificación automática de artefactos. Los artefactos oculares se eliminaron del registro EEG de cada participante utilizando el algoritmo extendido Infomax ICA. Se analizaron los datos de cada electrodo para identificar los canales dañados. Estos se identificaron calculando el cuarto momento estandarizado (curtosis) a lo largo de la señal de cada electrodo. En la Fig. 6.18 y 6.19 se presentan artefactos detectados en las señales EEG adquiridas. Posteriormente se eliminan las secciones que presentan ruido introducido por los artefactos, y se reconstruye la señal.

La extracción de características y la clasificación de las señales de EEG son cuestiones fundamentales en las aplicaciones de interfaz cerebro-computadora (BCI). Un problema desafiante cuando se trata de la extracción de características de EEG es la complejidad de la señal, ya que es de naturaleza no lineal, no estacionaria y aleatoria. Las señales se consideran estacionarias solo en intervalos cortos, por eso la mejor práctica es aplicar una técnica de ventana de tiempo corto para cumplir con este requisito. Sin embargo, todavía se considera una suposición que se mantiene durante una condición cerebral normal. Se pueden observar señales no estacionarias durante el cambio en el estado de alerta y vigilia, durante el parpadeo de los ojos y también durante las transiciones de los estados mentales (Bird, Manso, Ribeiro, Ekárt, & Faria, 2018). La señal se dividió en épocas de un segundo y se calculó el nivel de curtosis intracanal de cada época para rechazar las épocas muy dañadas por el ruido. Además, se aplicó la detección automática de artefactos, que rechaza la época cuando más de 2 canales tienen muestras que superan un umbral absoluto de $> 100 \mu V$ y un gradiente de $70 \mu V$ entre muestras. En la Fig. 6.20 se presentan las señales reconstruidas después de eliminar los artefactos detectados.

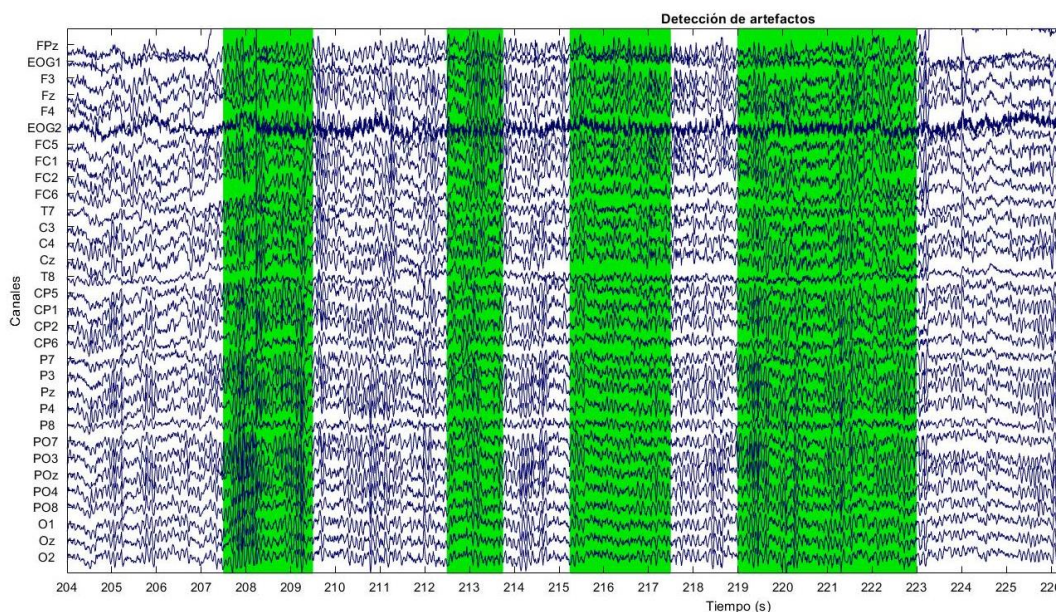


Figura 6.18. Detección automática de artefactos implementando toolbox EEGLAB en MATLAB.

Es posible observar el comportamiento de las señales de cada electrodo y la actividad correspondiente de estos, para detectar las zonas de mayor actividad cuando las personas observan un determinado concepto, y de esta manera conocer qué electrodos y en qué zonas cerebrales se presenta mayor información requerida para los objetivos de la tesis. Verificando los datos en las Figs. 6.21-6.23, se observa mayor actividad en los lóbulos occipitales, que es el centro de procesamiento visual de nuestro cerebro, incluido el procesamiento visuoespacial de bajo nivel (orientación, frecuencia espacial), la diferenciación de color y la percepción del movimiento y en el lóbulo parietal, que hace alusión a la integración de información proveniente de fuentes externas.

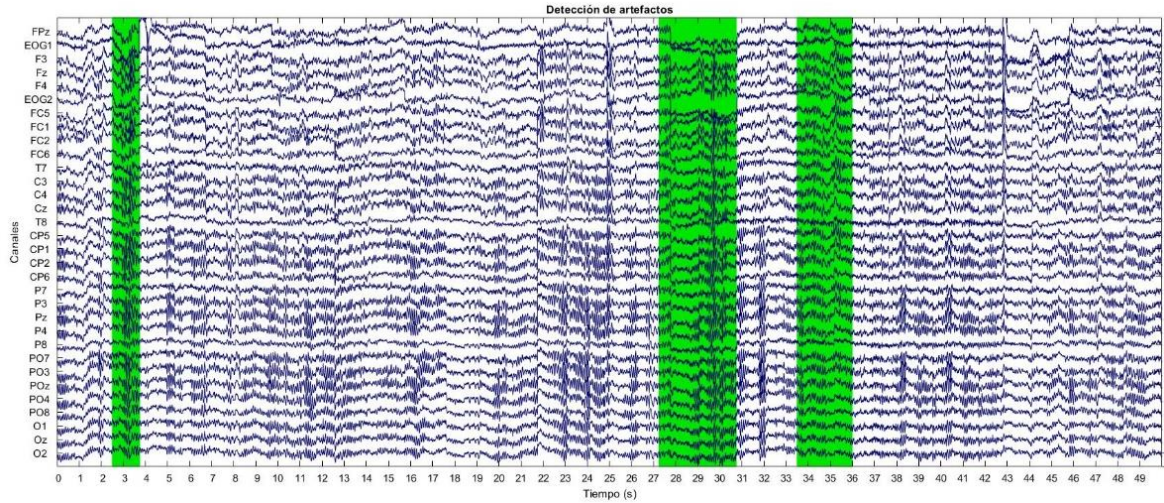


Figura 6.19. Detección automática de artefactos implementando toolbox EEGLAB en MATLAB.

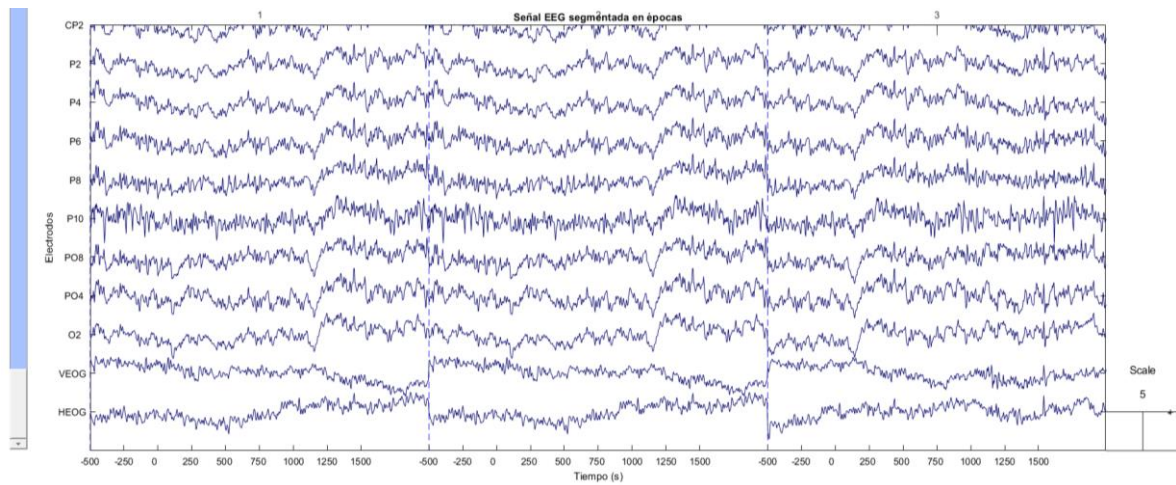


Figura 6.20. Señales EEG reconstruidas después de aplicar el método ICA para detección de artefactos y la eliminación de estos y segmentación de la señal en épocas.

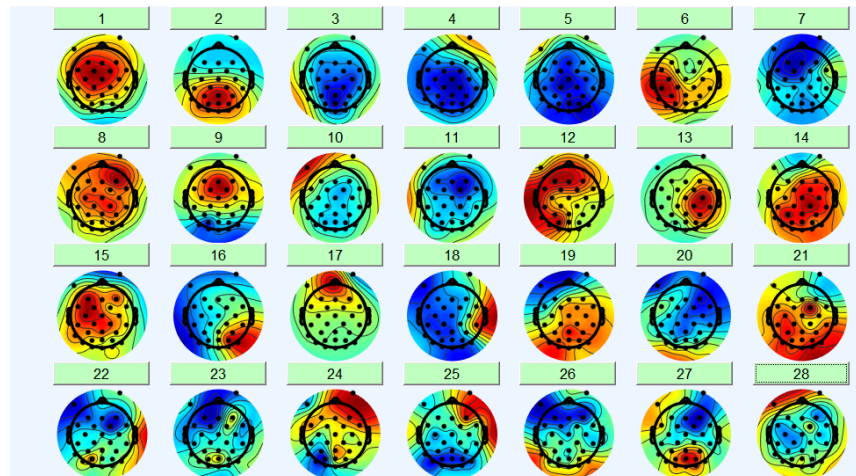


Figura 6.21. Análisis de la actividad cerebral en cada uno de los electrodos de la diadema con la que fueron capturados los datos, para observar a través del procesamiento ICA, el comportamiento de dicha actividad.

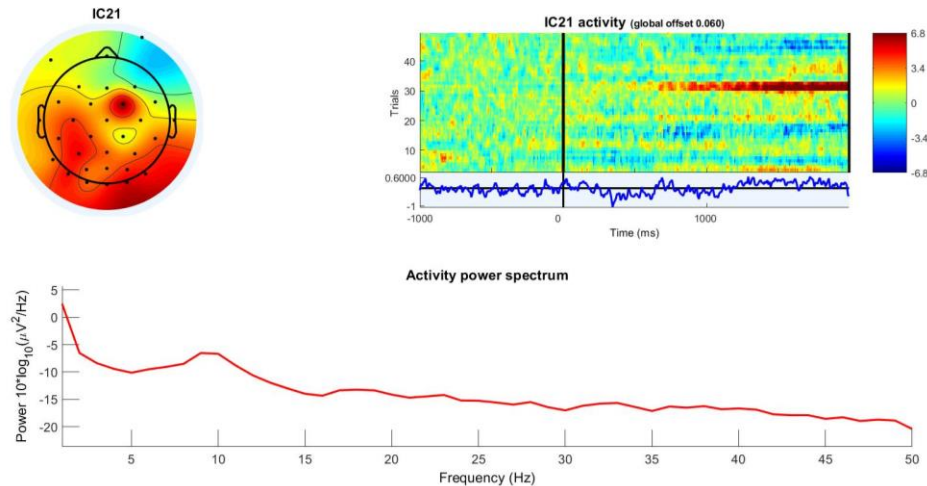


Figura 6.22. Espectro de potencia de la actividad cerebral en el electrodo 21.

Al graficar la actividad de las señales en el espectro del tiempo, se observan cambios de amplitud aproximadamente a los 250 ms, lo que implica generación de potenciales relacionados con eventos (ver Fig. 6.23). Los ERP proporcionan un medio no invasivo de medir la actividad cerebral en humanos, y su resolución temporal de milisegundos y su resolución espacial aproximada complementan la resolución temporal aproximada y la resolución espacial fina de fMRI, lo que representa una ventaja en este tipo de estudios.

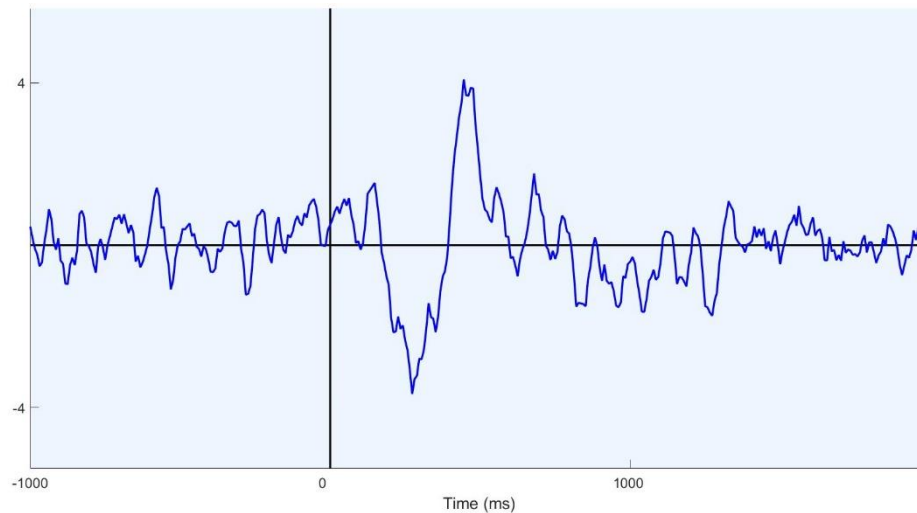


Figura 6.23. Análisis del espectro de potencia de la actividad cerebral del electrodo 21, graficado en el tiempo.

El objetivo de los paradigmas de EEG relacionados con eventos es recopilar los procesos cerebrales que se desencadenan por estímulos externos, en este caso, es precisamente lo que se tiene, la respuesta ante el estímulo externo.

La forma de onda EEG media restante es el potencial relacionado con el evento, que refleja la actividad EEG media relacionada con el estímulo desencadenado por un estímulo

específico. Los ERP pueden describirse por varias características: apariencia y forma, número, latencia, amplitudes de los "meneos", componentes del ERP (picos positivos y negativos) y topografía (que es la distribución de voltaje en los momentos pico en todos los electrodos). Los componentes ERP como el N400, P300 o N170 representan algunos de los componentes ERP más ampliamente analizados y mejor entendidos en la investigación académica (iMotions, 2019).

En el análisis de las ondas adquiridas, los ERP presentan variaciones en sus componentes P300. Se calcula la potencia de las señales para observar su comportamiento en el tiempo y mediante EEGLAB se obtiene la intensidad de actividad cerebral en cada momento del tiempo. En la Fig. 6.24 se observa esta intensidad en las regiones cerebrales a lo largo del tiempo, donde la mayor actividad cerebral se presenta aproximadamente en los 400 ms después de observar el estímulo, lo que da lugar al componente P300.

En la Fig. 6.25 se grafica la potencia de las señales y en efecto se observa que la mayor amplitud de potencia se encuentra alrededor de los 400 ms después de presentar el estímulo, lo que implica enfocarse en trabajar con los potenciales P300.

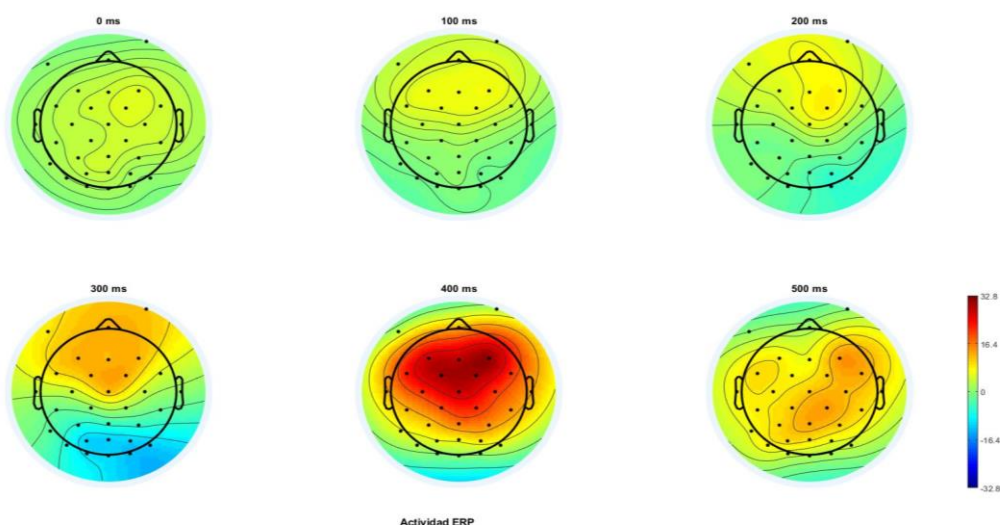


Figura 6.24. Intensidad de actividad en las distintas regiones cerebrales a lo largo del tiempo.

Después del preprocesamiento, se realizaron análisis de conectividad espectral y funcional. Se realizó un análisis espectral de EEG, utilizando el método de Welch, para estimar los espectros de potencia en cada época, dentro del ancho de banda de frecuencia clásico θ (4–8 Hz), α (8–12 Hz), β (13–25 Hz), γ (25–40 Hz). La banda de frecuencia δ (menos de 4 Hz) no se tuvo en cuenta en este estudio porque se relaciona con etapas más profundas del sueño. En la Fig. 6.26 se observa el espectro de la señal en el tiempo, dividido en cada banda de frecuencia, mientras que en la Fig. 6.27 se grafica cada banda en el dominio de la frecuencia mediante el método de Welch.

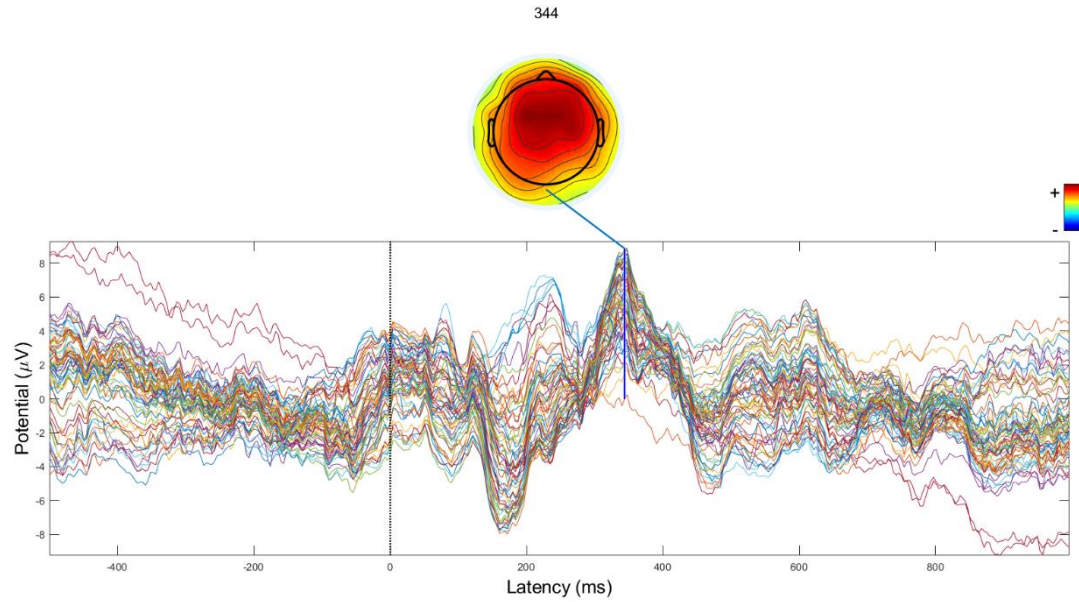


Figura 6.26. Potencia de las señales EEG, donde se observa la mayor amplitud de esta, alrededor de los 400 ms, que representa la mayor actividad cerebral.

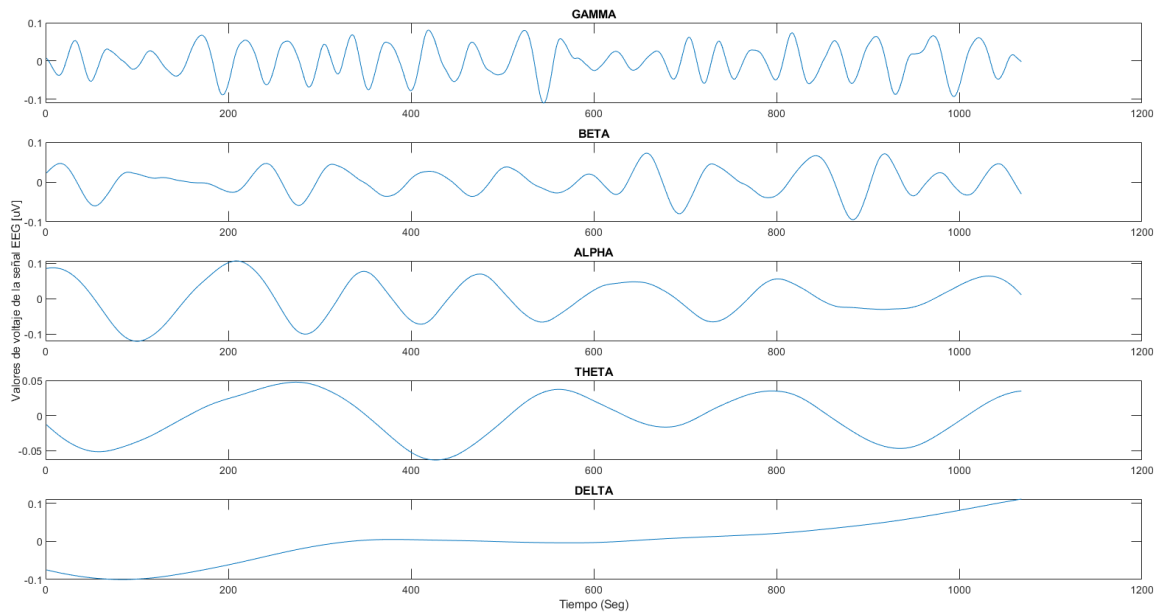


Figura 6.25. Señal graficada de un solo canal en el dominio del tiempo de la base de datos, en las diferentes bandas clásicas de comportamiento de la señal EEG.

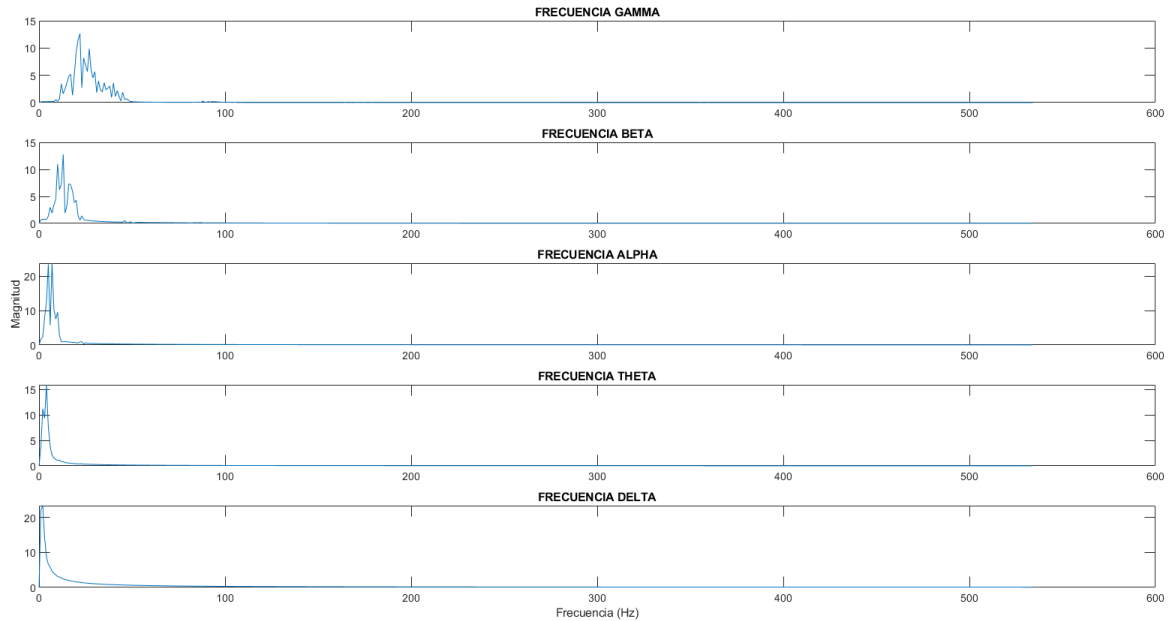


Figura 6.27. Análisis espectral mediante el método de Welch de un solo canal de la base de datos, en las diferentes bandas clásicas de comportamiento de la señal EEG.

6.7. Extracción de características

En este apartado se describe el conjunto de características consideradas en este trabajo para discriminar adecuadamente diferentes señales obtenidas por la presentación de los estímulos visuales. Estas características se basan en técnicas estadísticas, frecuencia de tiempo basada en la transformada rápida de Fourier (FFT), entropía de Shannon, características máximo-mínimo en secuencias temporales, covarianza logarítmica y otras. Todas las características propuestas para clasificar los estados mentales se calculan en términos de la distribución temporal de la señal en una ventana de tiempo determinada. Esta ventana de diapositiva se define como un período de 1 segundo, es decir, todas las características se calculan dentro de este instante de tiempo.

La extracción de características y la clasificación de las señales de EEG son cuestiones fundamentales en las aplicaciones de interfaz cerebro-computadora (BCI). Un problema desafiante cuando se trata de la extracción de características de EEG es la complejidad de la señal, ya que es de naturaleza no lineal, no estacionaria y aleatoria. Las señales se consideran estacionarias solo en intervalos cortos, por eso la mejor práctica es aplicar una técnica de ventana de tiempo corto para cumplir con este requisito. Sin embargo, todavía se considera una suposición que se mantiene durante una condición cerebral normal.

Las señales no estacionarias se pueden observar durante el cambio en el estado de alerta y la vigilia, durante el parpadeo de los ojos y también durante las transiciones de los estados mentales.

Los registros multicanal de las señales EEG continuamente generan una gran cantidad de datos. Los métodos de extracción de características pueden facilitar la representación e interpretación de estos datos para la posterior clasificación de los patrones de interés. Un gran

número de técnicas se han empleado para la extracción de características EEG en tareas orientadas a la categorización de objetos que, aunque no han evidenciado estadísticamente resultados diferenciables, se pueden esperar en efecto diferentes grados de éxito por la elección de los algoritmos adecuados. Por ello, la utilización en conjunto de un buen método de extracción y clasificación puede influir considerablemente en los resultados de clasificación de los patrones de interés (Medina, Sierra, & Barrios, 2018). En la tabla 6.5 se presenta la clasificación de las características extraídas de cada una de las señales.

Tabla 6.6. Tipo de características comúnmente extraídas en señales EEG.

Tipo de característica	Características extraídas
Características del dominio tiempo-frecuencia	1. Media de pico a pico.
	2. Valor cuadrado medio.
	3. La varianza.
	4. Desviación estándar
	5. Parámetro de Hjorth: actividad
	6. Parámetro Hjorth: movilidad.
	7. Parámetro Hjorth: Complejidad.
	8. Frecuencia espectral de potencia máxima.
	9. Densidad espectral de potencia máxima.
	10. Suma de potencia.
Características del sistema dinámico no lineal.	11. Entropía aproximada.
	12. Complejidad de C0.
	13. Dimensión de correlación.
	14. Entropía de Kolmogorov.
	15. Exponente de Lyapunov.
	16. Entropía de permutación.
	17. Entropía singular.
	18. Entropía de Shannon.
	19. Entropía espectral.

Características del dominio tiempo-frecuencia

Se consideraron diez tipos de características en los dominios de tiempo y frecuencia de cada señal. La media de pico a pico es la media aritmética de la longitud vertical desde la parte superior hasta la parte inferior de la serie temporal. El valor medio al cuadrado es la media aritmética de los cuadrados de las series de tiempo. La varianza mide el grado de dispersión de las series de tiempo.

La desviación estándar:

$$\sigma = \sqrt{\frac{1}{N} \sum_i^N (x_i - \mu)^2} \quad (6.6)$$

Después de transformar la serie temporal en el dominio de frecuencia a través de la transformada de Fourier, se calcula la suma de la potencia espectral y se extrae la densidad espectral de potencia máxima junto con su correspondiente valor de frecuencia. Según el trabajo de Hjorth (1970), también se extrajeron tres parámetros de Hjorth que pueden reflejar características de actividad, movilidad y complejidad: el parámetro de actividad refleja la información de la potencia de la señal, el parámetro de movilidad es una estimación de la frecuencia media y la complejidad refleja el ancho de banda y el cambio de frecuencia. Los parámetros de Hjorth se consideran adecuados para analizar señales de EEG no estacionarias.

Los Parámetros de Hjorth(): desarrolla las características mostradas en (6.7-6.9) para una serie de tiempo dada por $x(t)$ y un tamaño de muestra T .

Actividad:

$$A_x = \frac{\sum_{t=1}^T (x(t) - \mu)^2}{T} \quad (6.7)$$

Movilidad

$$M_x = \sqrt{\frac{\text{var}(\dot{x}(t))}{\text{var}(x(t))}} \quad (6.8)$$

Complejidad

$$C_x = \frac{\dot{M}(x(t))}{M(x(t))} \quad (6.9)$$

$\dot{x}(t)$ Representa la derivada de la serie de tiempo $x(t)$.

Características del sistema dinámico no lineal.

También se extrajeron nueve tipos de características que pueden reflejar las características de los sistemas dinámicos no lineales. Los investigadores han descubierto que el cerebro humano manifiesta muchas características que pertenecen específicamente a sistemas dinámicos no lineales y caóticos; por lo tanto, la señal EEG es intrínsecamente compleja, no lineal, no estacionaria y de naturaleza aleatoria. La entropía aproximada (ApEn) es una medida no lineal de la regularidad de una señal; cuanto más regular es una señal, más pequeña será la ApEn. La complejidad de C0 se adopta para medir la cantidad de componentes estocásticos, lo que supone que una señal consta de una parte regular y una parte estocástica.

La dimensión de correlación determina el número de dimensiones (variables independientes) que pueden describir la dinámica del sistema y refleja la complejidad del proceso y la distribución de los estados del sistema en el espacio de fase. El exponente de Lyapunov se usa para medir la dinámica aperiódica de un sistema caótico. Esta característica puede capturar la separación y la evolución de los estados iniciales del sistema en el espacio de fases. El exponente positivo de Lyapunov indica el caos en el sistema. La entropía de

Kolmogorov también es una medida del grado de caos y mide la velocidad a la que el sistema produce información, así como la velocidad a la que el sistema pierde información. El cálculo de la entropía de Komolgorov está muy influenciado por el ruido y la dimensionalidad de los datos. La complejidad de la actividad neuronal también se puede medir utilizando la teoría simbólica dinámica, en la que se puede mapear una serie temporal a una secuencia simbólica, a partir de la cual se puede derivar la entropía de permutación (PE). El valor más grande de PE es 1, lo que indica que la serie temporal es completamente aleatoria, mientras que el valor más pequeño de PE es 0, lo que indica que la serie temporal es completamente regular. La entropía del espectro singular se calcula mediante una descomposición de valor singular (SVD) de la trayectoria, que se obtiene al reconstruir la serie de tiempo unidimensional en un espacio de fase multidimensional. Esta característica refleja la incertidumbre y la complejidad de la distribución de energía y es un indicador de desincronización relacionada con eventos (ERD) y sincronización relacionada con eventos (ERS). La entropía de Shannon es una cuantificación clásica de la incertidumbre y se usa con frecuencia para medir el grado de caos en la señal EEG. La entropía espectral de potencia se basa en la entropía de Shannon y mide la complejidad espectral del sistema. Después de que se realiza la transformación de Fourier, la señal se transforma en un espectro de potencia, y la entropía de información del espectro de potencia se denomina entropía espectral de potencia (Xiang, et al., 2018). El análisis no lineal como la entropía de Shannon ha demostrado su eficacia en el procesamiento de señales y series de tiempo, ya que la aleatoriedad de los datos no lineales está bien incorporada mediante el cálculo de entropías sobre la serie de tiempo. La entropía es una medida de incertidumbre y en aplicaciones de interfaz cerebro-máquina, se utiliza para medir el nivel de caos del sistema, ya que es una medida no lineal que cuantifica el grado de complejidad de los datos. En teoría de la información, la entropía de Shannon viene dada por:

$$h = \sum_j S_j \times \log(S_j) \quad (6.10)$$

donde h es una característica calculada en cada ventana de tiempo de 1 seg. y S_j es cada elemento (normalizado) de esta ventana temporal. Luego, dada la misma ventana de tiempo, nos dividimos en dos para calcular la entropía de energía logarítmica de la siguiente manera:

$$\log e = \sum_i \log(S_i^2) + \sum_j \log(S_j^2) \quad (6.11)$$

donde i representa un índice para los elementos de la primera subventana (0 - 0,5 seg.) y j representa un índice para la segunda subventana (0,5 - 1 seg.).

Características estadísticas: para tener una representación compacta de los datos brutos del sensor en un rango de tiempo dado, estamos utilizando un conjunto de características estadísticas clásicas, ya que son útiles con eficacia probada para complementar un conjunto de características múltiples con el fin de reconocer patrones una serie del tiempo. Las características estadísticas son: (i) dado un conjunto de valores de datos $\{x_1, x_2, \dots, x_N\}$ adquiridos en cada ventana temporal, el valor medio de cada secuencia es calculado:

$$\mu = \frac{1}{N} \sum_I^N x_i \quad (6.12)$$

Momentos estadísticos de 3º y 4º orden, que nos da el sesgo para medir la asimetría de los datos, y también la curtosis para medir el pico de la distribución de probabilidad de los datos, respectivamente. Los momentos estadísticos empleados se calculan de la siguiente manera:

$$y = \frac{\mu^k}{\sigma^k} \quad (6.13)$$

$$\mu^k = \frac{1}{N} \sum_I^N (x_i - \mu)^k \quad (6.14)$$

Donde μ^k es el momento $k = \{3^\circ, 4^\circ\}$ sobre la media y $y = \{\text{asimetría}, k = 3; \text{curtosis}, k = 4\}$.

Máximos, mínimos y derivadas: Dada una ventana de tiempo de 1 segundo, los valores máximo y mínimo se calculan para aumentar la diversidad de los tipos de características. Las derivadas también se calculan como características temporales. Para cada ventana de tiempo, dividimos la ventana de tiempo por 2, por ejemplo, $w / 2 = 0.5$ seg. y $w = 1$ seg., lo que da como resultado dos secuencias de datos a ~ 125 Hz, luego calculamos:

$$uv = \frac{\mu^w - \mu^{w/2}}{2} \quad (6.15)$$

donde w y $w/2$ indican la primera y segunda mitad de la secuencia de datos en una ventana de tiempo de 1 seg. Se emplea la misma estrategia para obtener las derivadas, dada las características máximas y mínimas en las subventanas de tiempo (Bird, Manso, Ribeiro, Ekárt, & Faria, 2018):

$$\max_t = \frac{\max^w - \max^{w/2}}{2} \quad (6.16)$$

$$\min_t = \frac{\min^w - \min^{\frac{w}{2}}}{2} \quad (6.17)$$

6.8. Selección de clasificador de señales EEG

Para la etapa de clasificación se propone implementar un clasificador SVM, un método eficiente utilizado ampliamente en la comunidad científica. SVM encuentra hiperplanos discriminantes y separa los datos que pertenecen a diferentes clases con el máximo margen posible. Maximizar los márgenes aumenta las capacidades de generalización del clasificador SVM. En su formulación general, SVM requiere resolver el siguiente problema de optimización:

$$\text{Min} \frac{1}{2} \|a\|^2 + c \sum_{j=1}^N \xi_i \quad (6.18)$$

$$\text{sujeto a } w_i y(x_i) \geq 1 - \xi_i \text{ donde } i = 1 \dots N \text{ y } \xi_i \geq 0 \quad (6.19)$$

donde y es el modelo aprendido, c es el factor de penalización, a es el vector que representa los parámetros del modelo adaptativo, w_i es la etiqueta asociada con un punto de datos, i es el índice asociado con un punto de datos, ξ_i es la variable de holgura, x_i es el vector que representa un punto de datos, y N es el número de puntos de datos.

SVM lineal puede hacer límites de decisión no lineales usando el "truco del núcleo". Generalmente se hace mapeando los datos a un espacio de mayor dimensionalidad, con la ayuda de una función del núcleo. La función de base radial (RBF) tiene el menor número de hiperparámetros en comparación con otras funciones del núcleo (como el núcleo polinomial), por lo que RBF es menos costoso computacionalmente. La representación matemática del kernel RBF es:

$$k(y_i, y_j) = \exp(-\gamma \|y_i - y_j\|^2) \quad (6.20)$$

donde γ es el parámetro del kernel y y_i, y_j son vectores de entrenamiento.

El rendimiento de clasificación de la SVM depende de dos parámetros: 1) el parámetro del núcleo gamma γ , que determina la tolerancia de convergencia, y 2) el peso de penalización c , que es un parámetro de regularización que controla la tasa de clasificación errónea de los datos de entrenamiento. La influencia de estos parámetros se probó mediante una búsqueda en cuadrícula, junto con una validación cruzada, en el intervalo (0, 3) para γ y (0, 100) para c . Es necesario seleccionar los valores de los parámetros que den como resultado la mayor precisión de validación cruzada para construir el clasificador de SVM óptimo.

Para implementar la clasificación multiclase usando SVM, se puede ocupar la estrategia "uno contra uno". Más específicamente, la clasificación se realiza utilizando una estrategia de votación de victorias máximas, en la que cada clasificador asigna los datos a una de dos clases y la clase asignada recibe un voto. La clase que recibe el mayor número de votos se selecciona como clase final. Por lo tanto, considerando n como el número de clases, este enfoque implica la construcción de $n(n-1) / 2$ clasificadores. Por ejemplo, en un problema de clasificación de tres clases, la salida del clasificador tendrá uno de los tres estados discretos "0", "1" o "2" y no será una función continua.

La relación entre los datos de entrenamiento y los datos de las pruebas debe ser consistente con la literatura que muestra una relación es apropiada y generalmente utilizada para BCI basados en EEG en tiempo real. El uso de este procedimiento de validación cruzada podría ayudar a evitar un ajuste excesivo (Tavakolan, Frehlick, Yong, & Menon, 2017).

Todas las clasificaciones se realizaron dentro de los participantes mediante el análisis discriminante lineal (LDA). Las clasificaciones de clases múltiples se realizaron utilizando

una estrategia de uno contra todos. Para tener una estimación no sesgada de las tasas de clasificación, utilizamos un procedimiento de validación cruzada.

Cada clasificación utilizó una validación cruzada de diez veces, mediante la cual los datos se aleatorizaron y dividieron en diez subconjuntos de aproximadamente el mismo tamaño, sobre los cuales ocurrieron diez iteraciones de prueba de entrenamiento. Cada partición se utilizó como conjunto de prueba exactamente una vez, y las nueve particiones restantes se utilizaron para el entrenamiento en ese pliegue.

Al implementar un clasificador SVM con la base de datos previamente procesada como se explicó con anterioridad, se obtiene una exactitud de clasificación aproximadamente del 39%. La matriz de confusión generada se puede apreciar en la Fig. 6.28.

En la Fig. 6.29 se presenta la ubicación de los electrodos, así como la exactitud de cada uno de ellos, donde a partir del electrodo número 60, al 100 generan una mayor exactitud, los cuales hacen referencia a los lóbulos temporales y occipitales. La región subyacente a los electrodos occipitales mediales se clasifica por encima del azar, aunque no de manera significativa. En este caso se permite comprobar que la categorización de objetos visuales se presenta en la zona del lóbulo occipital, que se encuentra debajo del centro posterior del cerebro y es el centro principal de estimulación visual, es decir, en la corteza visual.

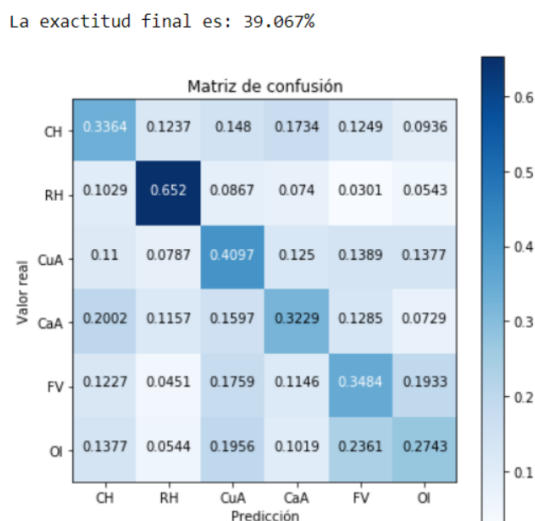


Figura 6.28 Métrica de exactitud y matriz de confusión de la clasificación utilizando el método de SVM.

Para comparar con LDA, realizamos un subconjunto de clasificaciones usando Support Vector Machine (SVM) con un kernel radial Gaussiano, y usando los mismos procedimientos de optimización de componentes que las clasificaciones LDA. LDA produjo resultados superiores y también fue aproximadamente diez veces más rápido de ejecutar. En la Fig. 6.30 se presenta la matriz de confusión generada por el método LDA y su exactitud, la cual es aproximadamente de 40.4%.

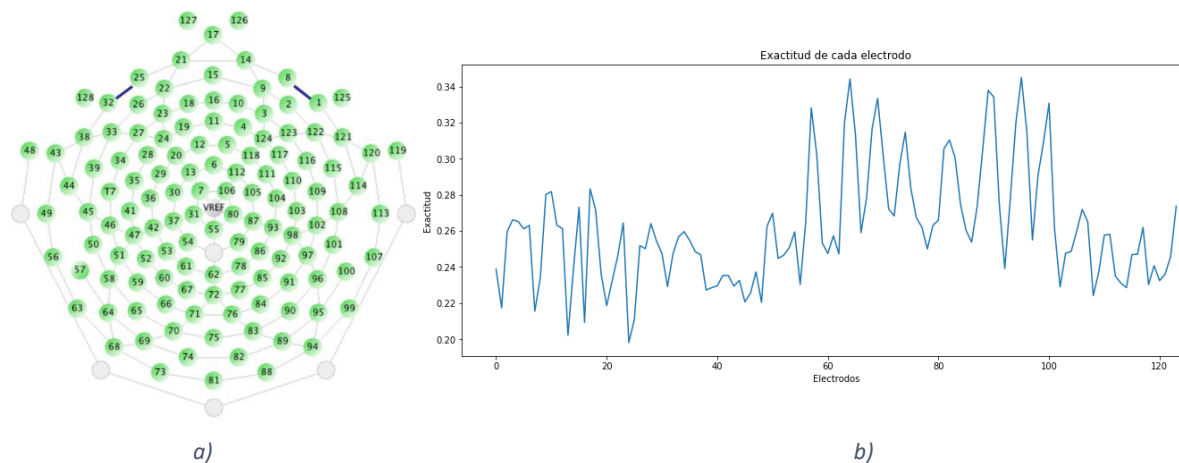


Figura 6.29. a) Localización espacial de los electrodos implementados en la adquisición de las señales EEG. b) Exactitud individual de los electrodos.

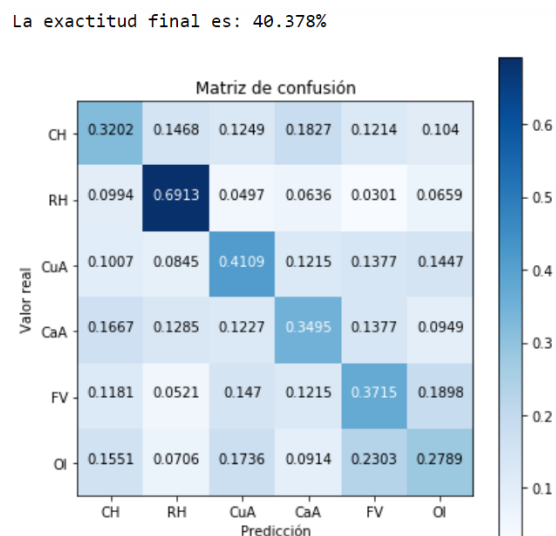


Figura 6.30. Métrica de exactitud y matriz de confusión de la clasificación utilizando el método de LDA.

Se realizaron clasificaciones a nivel de categoría para conocer entre qué clases el cerebro realiza una discriminación con mayor facilidad. La categoría de Rostro Humano se clasificó con mayor éxito, con el 63.14% de los ensayos etiquetados correctamente (ver Fig. 6.31). Como puede verse en la parte inferior derecha de la matriz de confusión, las dos categorías inanimadas muestran una confusión notable.

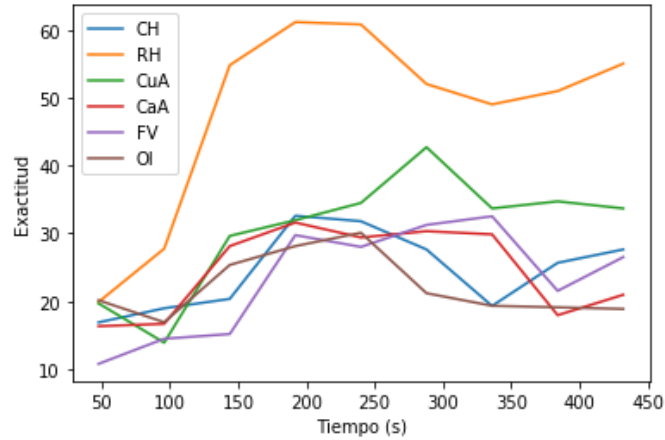


Figura 6.31. Gráfica de exactitud en la clasificación a nivel categoría.

El porcentaje de exactitud que entrega el clasificador es relativamente bajo, esto se puede deber a múltiples factores, como lo son la elección de las características extraídas, el proceso de filtrado y eliminación de artefactos, es posible que el clasificador esté detectando ruido en las señales y esto impida un desempeño óptimo, pero es una primera aproximación del sistema, en donde podemos decir que efectivamente se comienza a clasificar estos estímulos visuales percibidos por el usuario a través de las ondas cerebrales que estos generan.

6.9. Redes neuronales convolucionales (CNN)

En la actualidad, existen dos tipos de métodos de análisis para señales de EEG. Uno es el método tradicional de aprendizaje automático, que extrae características de EEG en el dominio del tiempo, el dominio de la frecuencia y el dominio del espacio, y luego se utiliza el clasificador para clasificarlos. El otro es el método de aprendizaje profundo, que consiste en diseñar una red neuronal artificial de extremo a extremo para permitir que la máquina capture y filtre automáticamente las características de la señal. Cada método tiene sus propias ventajas (Zheng & Chen, 2021).

El aprendizaje profundo con redes neuronales convolucionales (deep ConvNets) ha revolucionado la visión por computadora a través del aprendizaje de un extremo a otro, es decir, el aprendizaje de los datos sin procesar. Existe un interés creciente en el uso de ConvNets profundos para el análisis de EEG de extremo a extremo, pero aún se necesita una mejor comprensión de cómo diseñar y entrenar ConvNets para la decodificación de EEG de extremo a extremo y cómo visualizar las características informativas de EEG que aprenden los ConvNets. . Aquí, estudiamos ConvNets profundos con una variedad de arquitecturas diferentes, diseñadas para decodificar señales EEG que son generadas por distintas categorías de imágenes a través de estímulos visuales.

En general, las redes neuronales convolucionales van a estar construidas con una estructura que contendrá 3 tipos distintos de capas (ver Fig. 6.32):

- Una capa convolucional, que es la que le da el nombre a la red. Estas consisten en tomar “grupos de datos cercanos” de la matriz de entrada e ir operando matemáticamente (producto escalar) contra una pequeña matriz que se llama kernel.
- Una capa de reducción o de pooling, la cual va a reducir la cantidad de parámetros al quedarse con las características más comunes.
- Una capa clasificadora totalmente conectada, la cual nos va dar el resultado final de la red.

Originalmente las CNN fueron diseñadas para el procesamiento de imágenes. Actualmente son requeridas en múltiples tareas, dentro de ellas el análisis de señales EEG. Para su implementación es necesario adecuar los datos de entrada, de manera que la red pueda aplicar las diferentes convoluciones y operaciones sin perder información, o pasar por alto características importantes, y aunado a esto, la correcta interpretación de dichos datos.

Se realizó la investigación en el estado del arte para identificar las principales arquitecturas para que la CNN sea capaz de interpretar adecuadamente los datos, se requiere realizar una convolución temporal a los datos de entrada, que permita realizar la convolución a lo largo del eje del tiempo. Esto corresponde a un prefiltrado lineal de la señal de EEG de cada canal. Y una segunda capa que realice la convolución a lo largo del eje del canal EEG, es decir, solo se consideran los diferentes valores del canal al mismo tiempo. En la fig. 6.33 se presenta una adecuación de los datos de entrada y la forma en la que se distribuyen las convoluciones mencionadas.

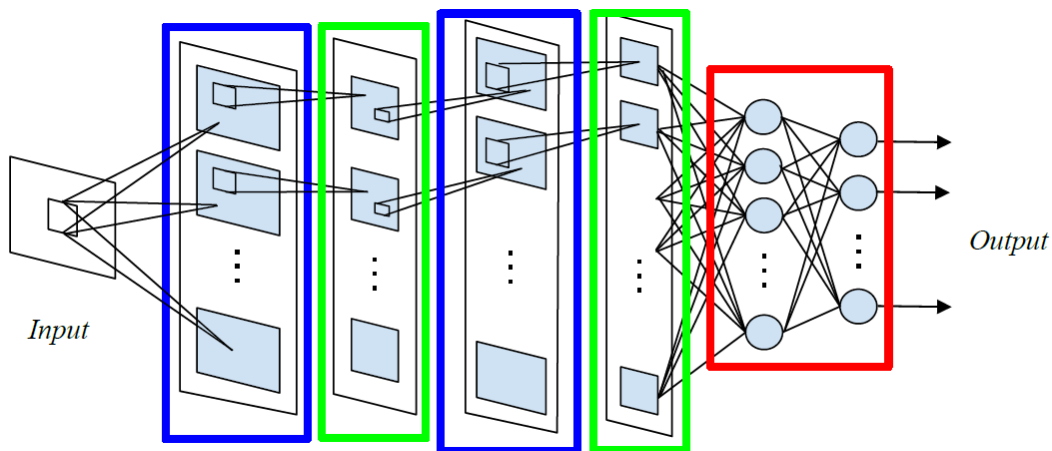


Figura 6.32. Red convolucional: Capas convolucionales (azul), Pooling (verde), y Fully-connected (rojo). Fuente: (Dagallier & Campillos, 2018).

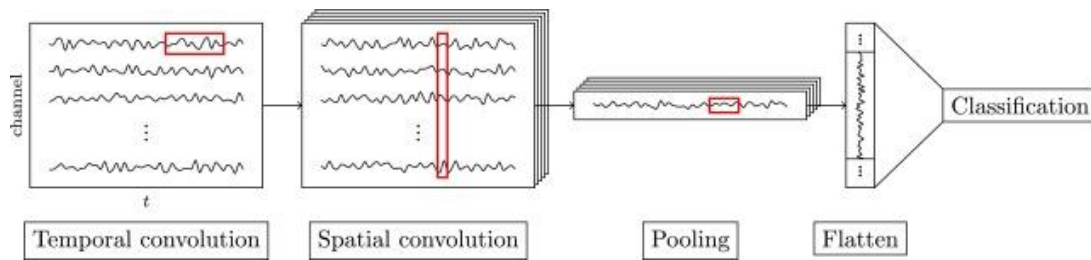


Figura 6.33. Arquitectura general: Las capas de la CNN se pueden describir como diferentes tipos de pasos de filtrado y procesamiento antes de la clasificación real. Los rectángulos rojos indican el núcleo del filtro y las dimensiones de la agrupación. Fuente: (Dosea, Møllera, Iversenb, & Puthusserypady, 2018)

6.9.1. Preparación de los datos para la entrada a la CNN

Las señales de la base de datos implementada, han sido previamente procesados y recopilados en un fichero .mat (Matlab) bajo diversas variables:

- N: Número de muestras de tiempo por prueba. Este valor siempre es igual a 64.
- Fs: Frecuencia de muestreo de los datos. Siempre es igual a 128 Hz.
- T: Número de pruebas realizadas. Tiene un valor de aproximadamente 5184 pruebas por sujeto.
- X_2D: Matriz de dos dimensiones representando las muestras en el número de pruebas y las muestras en el tiempo de todos los electrodos concatenados. Sus dimensiones para un sujeto son: (T, 124*N)
- X_3D: Matriz tridimensional anteriormente mencionada con dimensiones: (124, N, T)
- categoryLabels: Vector que incluye la categoría a la que se corresponde cada prueba. Cada número se corresponde con una categoría siguiendo la tabla III. Sus dimensiones son: (T,1).

Para el desarrollo de esta tesis se trabaja con la matriz de datos X_3D y la variable 'categoryLabels'. En el caso de analizar la precisión de la clasificación de imágenes de los modelos empleando todos los sujetos simultáneamente, se concatenan estas matrices de cada uno de los sujetos en la dimensión del número de pruebas T. Posteriormente, esta nueva matriz concatenada se mezcla de forma aleatoria en la dimensión anteriormente mencionada. Así, los datos de todos los sujetos se encuentran entremezclados, buscando un enfoque más general. Respecto a las etiquetas que acompañan a cada prueba, también son mezcladas en el mismo orden que la matriz de datos y procesadas con la función one-hot antes mencionada.

La capa de entrada a la CNN toma una entrada bidimensional con dimensiones espaciales (número de canales) y temporales (puntos de datos por muestra) para cada muestra.

A continuación se presenta de manera gráfica el proceso de convoluciones de las entradas, y el por qué de la estructura de estas. En la parte superior de la Fig. 6.34 se define la entrada como la representación de las señales en todos los canales, en el eje "y" se representan los canales, y en el eje "x" se representan las muestras de tiempo (denominadas ventanas o épocas). A esta estructura de entrada se le aplica una convolución temporal de "F" filtros (parte inferior de la figura), que dará una profundidad al arreglo de salida, lo que permitirá

posteriormente aplicar una convolución espacial, que implique toda la dimensión de los canales, en este caso 124:

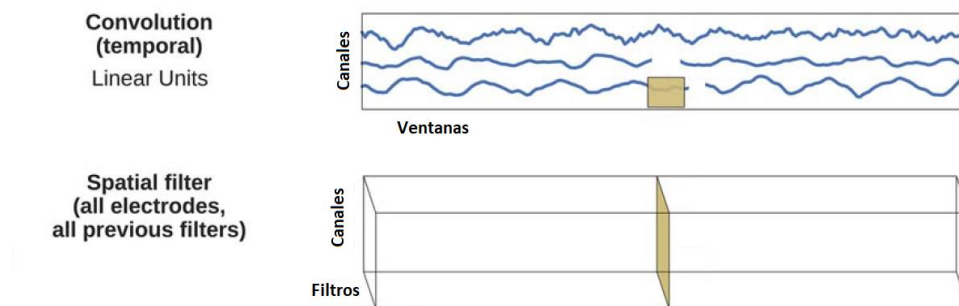


Figura 6.34. Representación de las convoluciones aplicadas a la matriz de entrada a la CNN, donde el eje “y” representa el número de canales, y el eje “x” representa los puntos de tiempo de la ventana. Fuente: (Schirrmester R. T., et al., 2017).

Codificación One-hot

En tareas de clasificación de datos en categorías, cada muestra del conjunto de datos va acompañada por una etiqueta que determina la categoría a la que pertenece. En numerosas ocasiones, esta etiqueta es una palabra (‘triste’, ‘feliz’, ‘aburrido’). Sin embargo, muchos algoritmos de aprendizaje automático solo son capaces de trabajar con datos numéricos.

Una solución es asignar a cada categoría un número entero diferente, empezando por ‘0’ hasta N-1 clases, técnica conocida como “integer encoding”. La desventaja de este método es que los algoritmos entienden que un número mayor implica que esa categoría es más importante. Esto puede conllevar predicciones muy poco certeras y problemas con el aprendizaje. Para solucionar esto existe la codificación one-hot. En lugar de asignar un número entero a cada clase, le atribuye un valor binario en función de su pertenencia a la clase. De esta forma, todas las categorías tienen la misma importancia para el modelo. A continuación se presenta un ejemplo de esta codificación.

Tabla 6.7. Ejemplo de codificación one-hot.

Triste	Feliz	Aburrido
1	0	0
0	1	0
0	0	1

6.9.2. Arquitectura del modelo

La arquitectura del modelo propuesto se basa en el modelo CNN de paradigma cruzado basado en EEG, EEGNet (Lawhern, et al., 2018). Según los autores, la arquitectura original de EEGNet se adapta bien a cualquier tipo de problema de clasificación basado en EEG (ver Fig. 6.35). La estructura básica de las ramas se puede describir de la siguiente manera: las líneas denotan la conectividad del núcleo convolucional entre entradas y salidas (llamadas

mapas de características). La red comienza con una convolución temporal (segunda columna) para aprender los filtros de frecuencia, luego usa una convolución en profundidad (columna central), conectada a cada mapa de características individualmente, para aprender los filtros espaciales específicos de frecuencia. La convolución separable (cuarta columna) es una combinación de una convolución en profundidad, que aprende un resumen temporal para cada mapa de características individualmente, seguido de una convolución puntual, que aprende cómo mezclar de manera óptima los mapas de características.

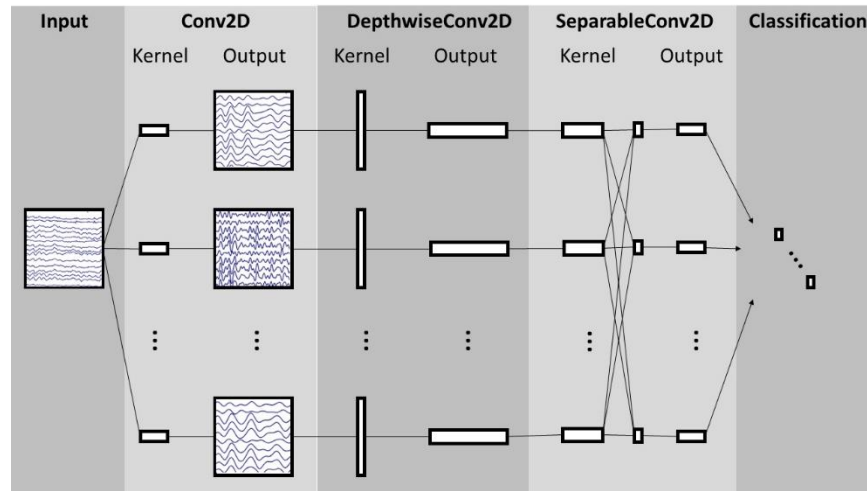


Figura 6.35. Arquitectura del modelo original de EEGNet.

Es necesario ajustar la CNN con los parámetros de entrada de nuestra base de datos. En este caso particular las dimensiones de los datos de entrada son de 64 muestras de tiempo y 124 canales. A continuación se presenta un diagrama de la arquitectura EEGNet con los valores correspondientes a la base de datos aquí implementada.

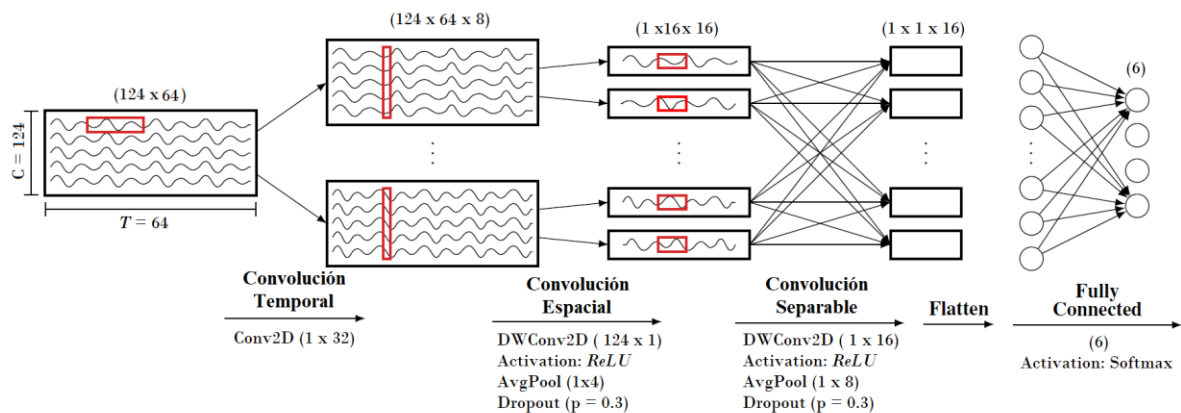


Figura 6.36. Arquitectura EEGNet adaptada a los parámetros previamente definidos de la base de datos.

Como primer paso, el modelo aprende filtros de frecuencia con convoluciones temporales utilizando capas convolucionales bidimensionales con tamaño de filtro con tamaño (1, 32),

con la longitud de filtro que corresponde a una cuarta parte de la frecuencia de muestreo (128 Hz aproximadamente), generando mapas de características F1 (número de filtros temporales =8), que contienen la señal EEG en diferentes frecuencias de paso de banda. La capa convolucional temporal precede a una capa convolucional en profundidad con filtros (124, 1), cuya longitud hace referencia al número de canales (electrodos). El principal beneficio de las convoluciones en profundidad es reducir el número de parámetros entrenables, al tiempo que proporciona una forma de extraer filtros espaciales específicos de frecuencia para cada filtro temporal. Se define una convolución espacial como una convolución a lo largo del eje del tiempo que extrae filtros espaciales, y la salida representa el número de filtros espaciales para aprender dentro de cada convolución temporal.

Finalmente, se utilizó una capa convolucional separable con tamaño de filtros (1, 16). La convolución separable combinó convoluciones en profundidad con convoluciones puntuales para aprender resúmenes temporales de cada mapa de características individualmente y mezclar de manera óptima los mapas de características juntos.

Las convoluciones separables consisten en realizar primero una convolución espacial en profundidad (que actúa en cada canal de entrada por separado) seguida de una convolución puntual que mezcla los canales de salida resultantes. A cada capa convolucional también le siguió la normalización por lotes, que es un método eficaz para reducir el sobreajuste y mejorar la velocidad de entrenamiento de la red.

Además, cada capa convolucional separable y en profundidad fue seguida por una función de activación de unidad lineal rectificadora (ReLU) y es más eficiente computacionalmente en la clasificación basada en CNN (Roots, Muhammad, & Muhammad, 2020).

Además, para reducir el sobreajuste, a cada capa de agrupación le siguió una función dropout con el valor 0,3 y su salida se proporcionó como entrada a la capa de clasificación softmax, produciendo las probabilidades para cada una de las 6 clases.

6.9.3. Evaluación del modelo

Se realiza la concatenación de los datos de los 10 sujetos para generar una base de datos de mayores dimensiones, donde el modelo se entrena con los datos de todos los participantes.

El 70% de las muestras totales se seleccionaron aleatoriamente para entrenar el modelo, el 20% para validación y el 10% para pruebas. Se usó la misma división de datos para cada experimento, ya que la selección aleatoria se implementó usando un valor de división fijo.

La división del conjunto de datos en dos conjuntos es bastante práctica, pero no constituye en gran medida a una buena clasificación. Para reducir en gran medida las posibilidades de sobreajuste es recomendable particionar el conjunto de datos en los tres subconjuntos mencionados anteriormente ya que al enfrentarse a nuevos datos es necesario asegurarse que el desempeño sea lo más óptimo. Al realizar el entrenamiento es necesario intentar reproducir las condiciones reales tanto como sea posible. Por esta razón, es una buena práctica dividir nuestro conjunto de datos en tres partes como se menciona. Se usa el conjunto de validación para evaluar los resultados del conjunto de entrenamiento. A continuación, se usa el conjunto de prueba para verificar la evaluación de la red. El uso de datos completamente nuevos

permite obtener una opinión imparcial sobre qué tan bien funciona nuestro algoritmo. En la Fig. 6.37 se muestra el nuevo flujo de trabajo.

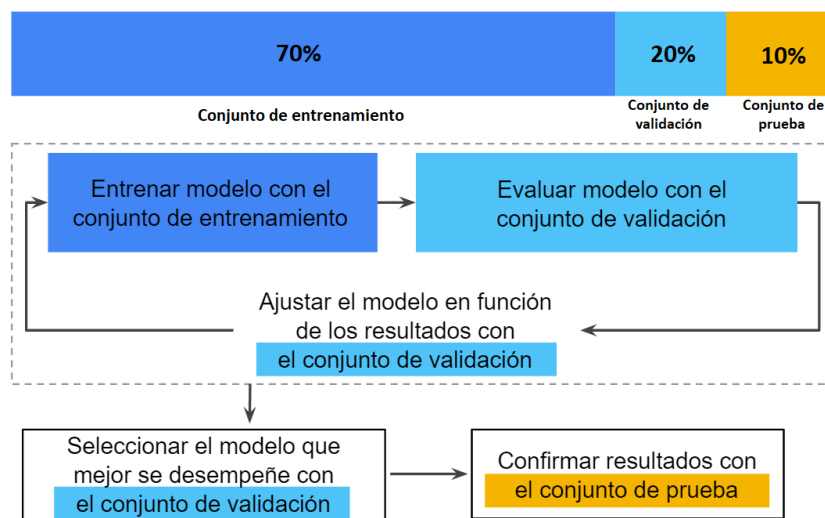


Figura 6.37. Separación del conjunto de datos para un flujo de trabajo eficaz. Fuente: (Developers, 2021)

Durante la fase de entrenamiento, se evaluó la pérdida de validación y se guardaron los pesos del modelo con las mejores precisiones de validación. Antes de la fase de prueba, se cargaron los pesos del modelo con las mejores precisiones de validación y el modelo se evaluó en el conjunto de datos de prueba mediante la predicción de las etiquetas de los objetivos y el cálculo de la precisión como la proporción de predicciones correctas del número total de objetivos, así como la precisión, recuperación y tiempo computacional por muestra para el modelo.

Con este flujo de trabajo mejorado, se realiza lo siguiente:

- Selecciona el modelo que mejor se desempeñe con el conjunto de validación.
- Verifica el modelo con respecto al conjunto de prueba.
- Este flujo de trabajo es más eficaz porque crea menos exposiciones al conjunto de prueba.

6.9.4. Overfitting o sobreajuste

De los problemas más frecuentes que se generan al entrenar una red neuronal, es el sobreajuste u “overfitting”. Consiste en ajustar excesivamente bien los hiperparámetros, parámetros que se fijan antes de realizar el entrenamiento al set de datos de entrenamiento, es decir, se presenta una situación en la que la red neuronal está tan ajustada al conjunto de entrenamiento que es difícil generalizar y hacer predicciones para nuevos datos. Este sobreajuste conlleva una mala adquisición de características generales a todos los datos, y, por tanto, reduce la calidad de respuesta ante el conjunto de datos de prueba y ante datos futuros (Dagallier & Campillos, 2018).

A continuación se ejemplifica un caso de overfitting de una red, para visualizar la importancia de la corrección de este error:

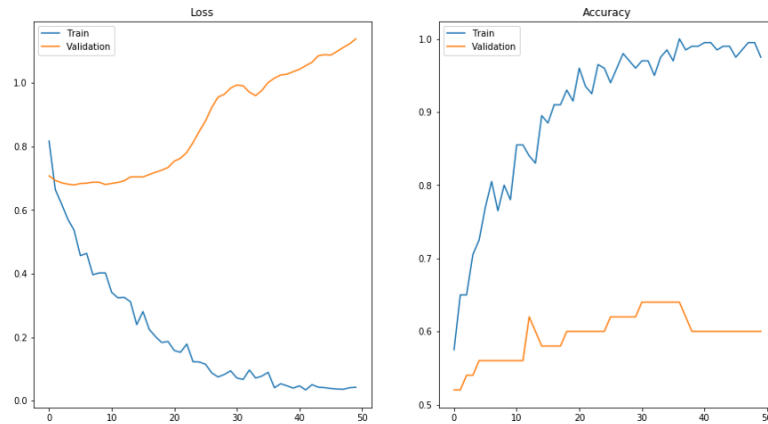


Figura 6.38. Gráficas con señal de sobreajuste: la pérdida de entrenamiento está disminuyendo, pero la pérdida de validación está aumentando. Fuente: (Schluter, 2019).

Si se observa un comportamiento similar en el desempeño de la red, es una señal clara de que el modelo está sobreajustado: está aprendiendo los datos de entrenamiento realmente bien, pero no generaliza el conocimiento a los datos de prueba.

Al realizar las primeras pruebas con los modelos de redes neuronales convolucionales se presentó este error. En la Fig. 6.39 se muestra la gráfica de exactitud y pérdida del caso particular de una red con dos capas convolucionales.

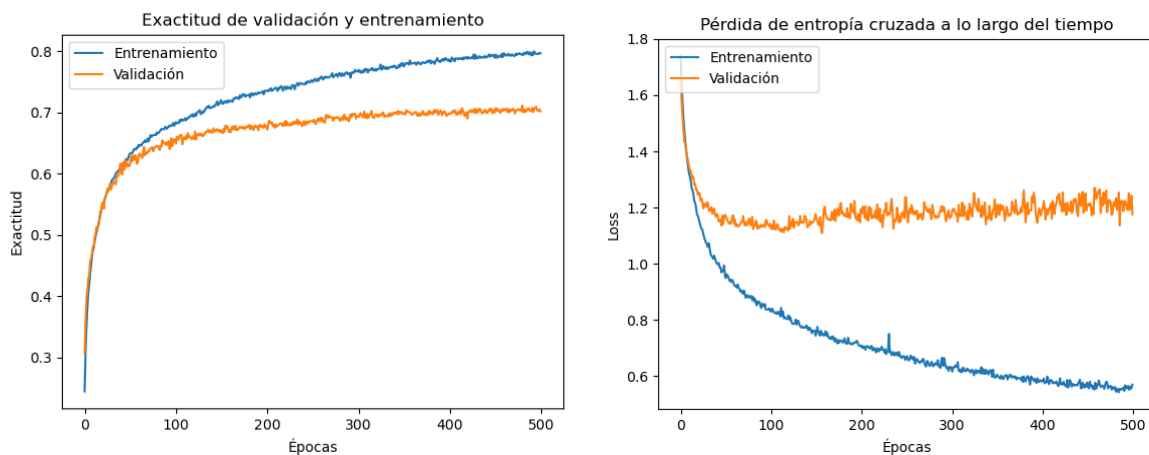


Figura 6.39. Gráfica de exactitud y pérdida implementando una red EEGNet.

Los resultados de la gráfica de pérdida presentan un claro problema de sobreajuste en la red. Para mejorar la puntuación, básicamente se pueden hacer dos cosas:

- Mejorar el modelo
- Mejorar nuestros datos

1. Simplificando el modelo

El primer paso cuando se trata de sobreajuste es disminuir la complejidad del modelo. En el modelo base dado, hay 2 capas ocultas, una con 128 y otra con 64 neuronas. Además, la capa de entrada tiene 300 neuronas. Esta es una gran cantidad de neuronas. Para disminuir la complejidad, se pueden eliminar capas o reducir la cantidad de neuronas para hacer la red más pequeña. No existe una regla general sobre cuánto eliminar o qué tan grande debe ser su red.

2. Adición de capas de abandono (Dropout)

Dropout Layers puede ser una forma fácil y eficaz de evitar el sobreajuste en los modelos. Una capa de abandono deja caer aleatoriamente algunas de las conexiones entre capas. Esto ayuda a evitar el sobreajuste, porque si se cae una conexión, la red se ve obligada a hacerlo. Afortunadamente, con keras es muy fácil agregar una capa de abandono (Schluter, 2019).

La regularización de abandonos es una forma computacionalmente económica de regularizar una red neuronal profunda. El abandono funciona mediante la eliminación probabilística o el “abandono” de las entradas de una capa, que pueden ser variables de entrada en la muestra de datos o activaciones de una capa anterior. El funcionamiento de esta herramienta consiste en obviar una fracción aleatoria “p” de nodos pertenecientes a una capa oculta, en cada muestra de entrenamiento y en cada iteración. De esta forma, se eliminan tanto la neurona elegida aleatoriamente, como sus conexiones con las capas anterior y posterior, y sus activaciones (ver Fig. 6.40). Durante la fase de validación, se usan todas las activaciones de la red neuronal, pero reduciéndose por un factor p, para tener en cuenta el descarte hecho durante el entrenamiento.

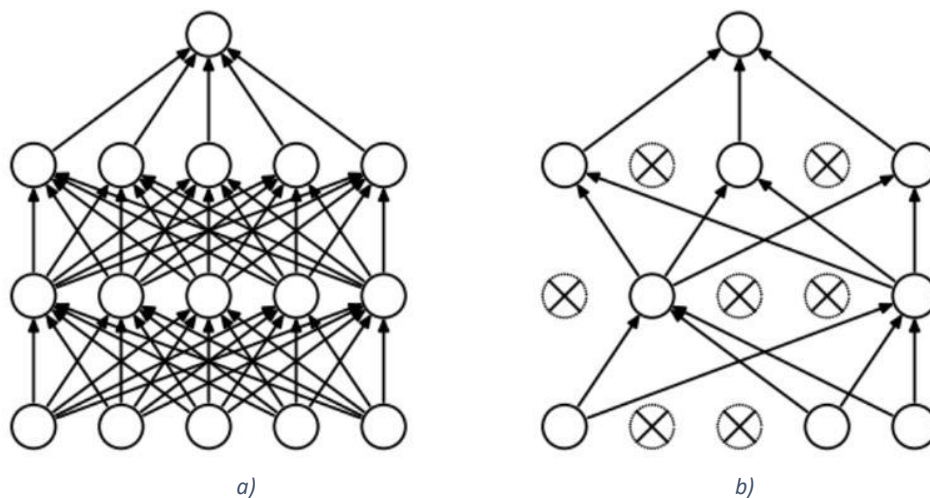


Figura 6.40. a) Red neuronal estándar. b) Red neuronal después de aplicar “Dropout”. Fuente: (Pelayo & Gordillo, 2018).

Una de los puntos negativos de incluir esta función, es que el modelo necesita más iteraciones para alcanzar una convergencia. Sin embargo, cada iteración se ejecuta en un

tiempo más reducido, y las características extraídas son más robustas (Pelayo & Gordillo, 2018).

Conociendo esta información, para prevenir el sobre entrenamiento de la red, se incluye un descarte aleatorio de un porcentaje de los datos entre cada capa. Se proponen 3 porcentajes diferentes de descarte (15%, 30% y 50%), obteniéndose los resultados de la figuras 6.38-6.40. La tasa de acierto del conjunto de validación, muy similar en los tres casos, presenta una menor variabilidad con un dropout de 0.3 para el modelo. Cabe mencionar que este valor es el que comúnmente se implementa en aprendizaje profundo.

Se observa una clara mejora en las tasas de acierto para el caso de la implementación de las redes neuronales convolucionales. Esta tasa experimenta un incremento de aproximadamente un 30% respecto a la obtenida en el supuesto anterior. En la tabla 6.7 se presenta la métrica de exactitud obtenida de la evaluación de los modelos previamente entrenados, donde la arquitectura DeepCovNet presenta mejores resultados.

Tabla 6.8. Parámetros evaluados en las CNN y exactitud del entrenamiento.

	<i>EEGNet</i>	<i>ShallowConvNet</i>	<i>DeepConvNet</i>
<i>Parámetros</i>	2,934	199,766	507,631
<i>Optimizador</i>	Adam	Adam	Adam
<i>Exactitud</i>	66.06	72.73%	73.19%

A continuación se presentan las gráficas que representan las métricas de exactitud y pérdida del entrenamiento de cada una de las arquitecturas utilizadas en este trabajo, donde se aprecia que la línea de validación se aproxima con buena optimización a la línea de entrenamiento, esto quiere decir que no existe un sobreajuste en el entrenamiento, o si existe este es mínimo, el entrenamiento de cierta manera fue exitoso, y se realiza una correcta validación de las redes con el conjunto de prueba, previamente mencionado.

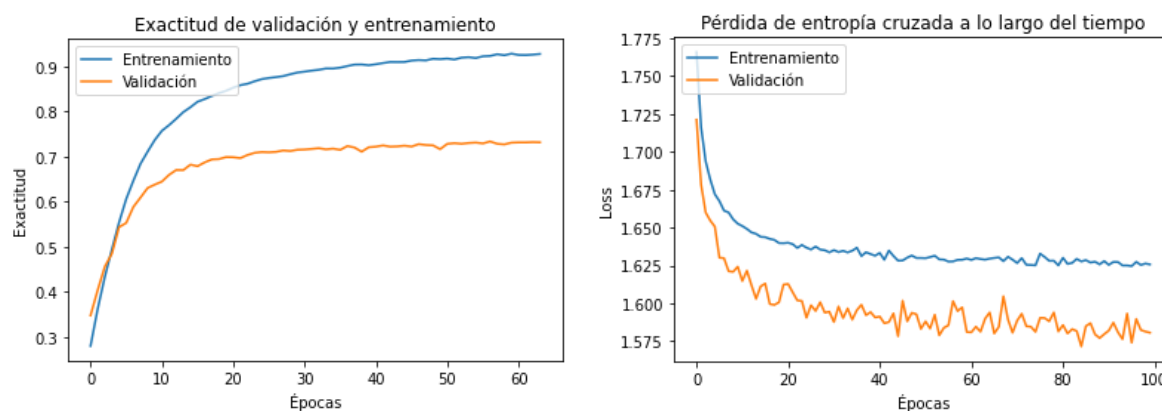


Figura 6.41. Gráficas de exactitud y pérdida de la arquitectura EEGNet.

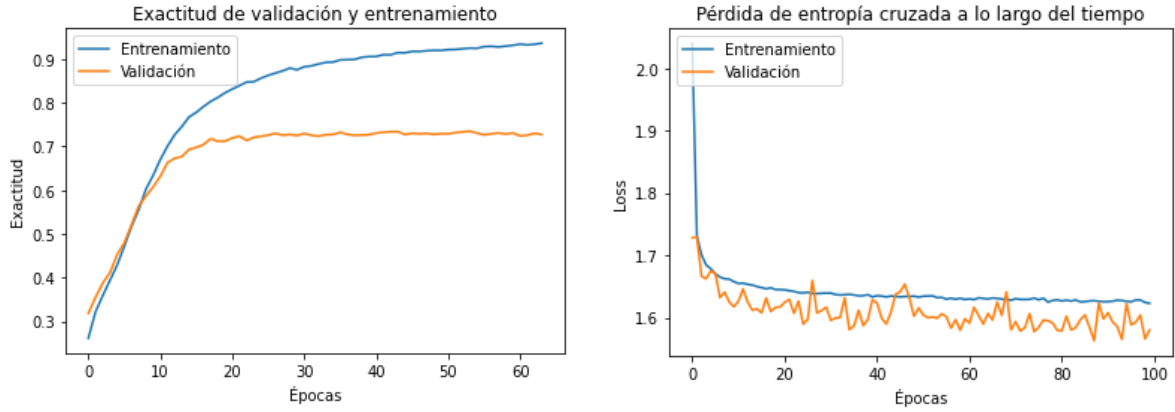


Figura 6.42. Gráficas de exactitud y pérdida de la arquitectura DeepConvNet.

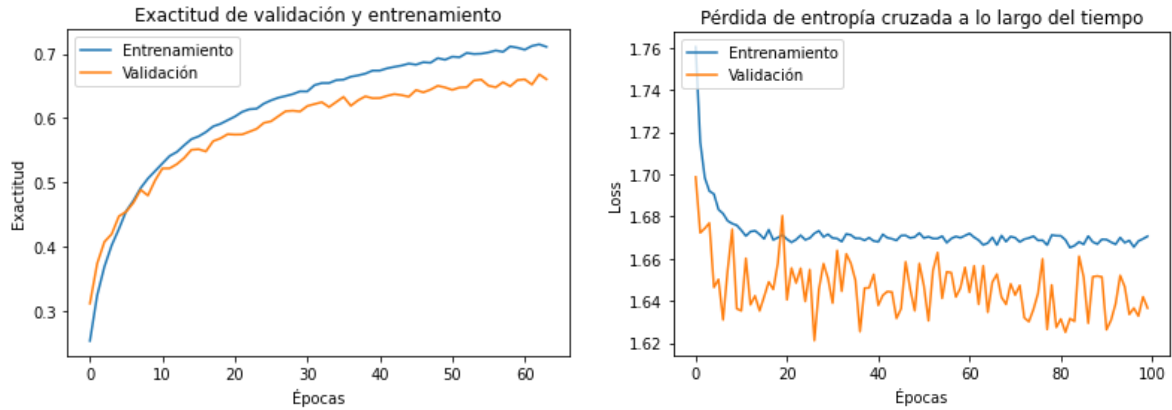


Figura 6.43. Gráficas de exactitud y pérdida de la arquitectura ShallowConvNet.

Se presentan las matrices de confusión de cada modelo implementado, la diagonal que representa la correcta clasificación de las imágenes en las distintas categorías se muestra más marcada, indicando un mayor grado de precisión (Ver Fig. 6.44 y 6.45.)

Posteriormente se extraen las métricas de precisión, recall y f1-score para observar el comportamiento del entrenamiento y su eficiencia.

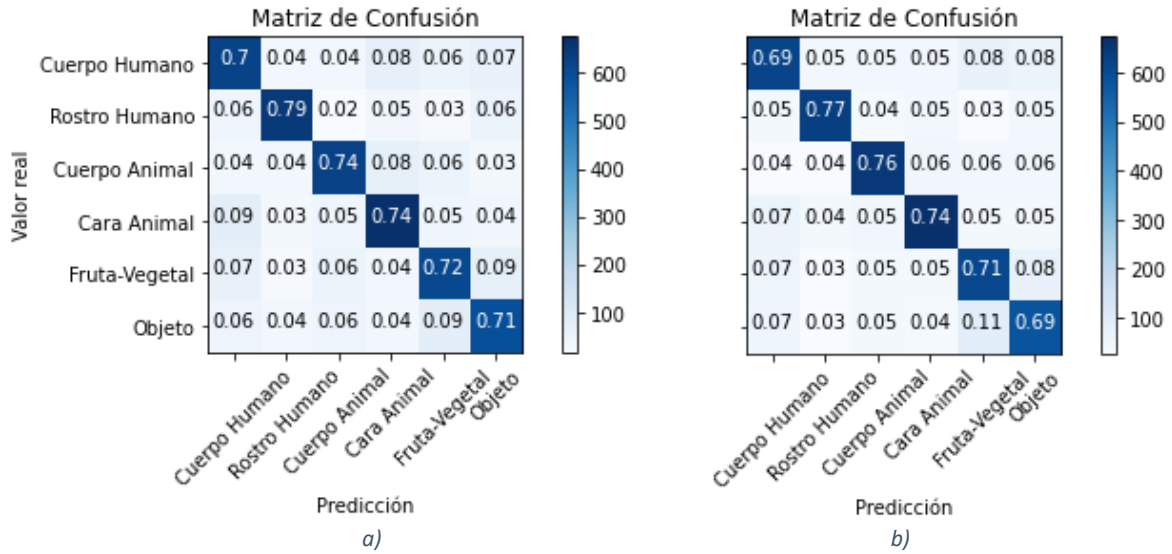


Figura 6.44. a) Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: EEGNet. b) Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: DeepConvNet.

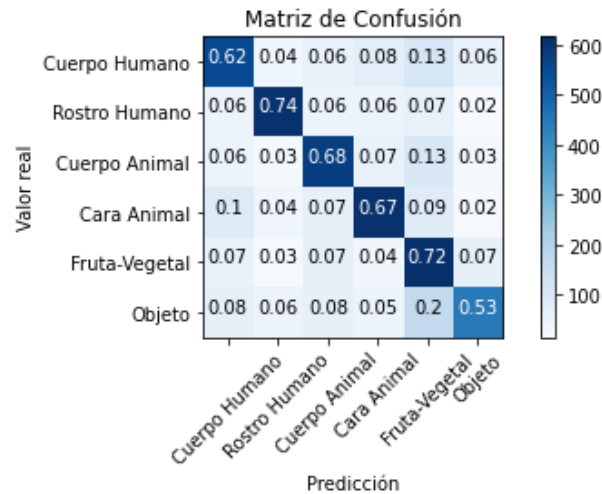


Figura 6.45. Matriz de confusión normalizada aplicando una arquitectura previamente diseñada de CNN: ShallowConvNet.

Se analizan otras métricas de validación de la matriz de confusión. A continuación se presenta un resumen de las métricas obtenidas a partir de la matriz de confusión.

Reporte de métricas de clasificación para la arquitectura EEGNet:

	precision	recall	f1-score	support
0	0.69	0.70	0.70	893
1	0.81	0.79	0.80	827
2	0.76	0.74	0.75	853
3	0.73	0.74	0.74	911
4	0.70	0.72	0.71	858
5	0.70	0.71	0.70	844

accuracy			0.73	5186
macro avg	0.73	0.73	0.73	5186
weighted avg	0.73	0.73	0.73	5186

Reporte de métricas de clasificación para la arquitectura DeepConvNet:

precision	recall	f1-score	support		
	0	0.71	0.69	0.70	893
	1	0.79	0.77	0.78	827
	2	0.75	0.76	0.76	853
	3	0.76	0.74	0.75	911
	4	0.68	0.71	0.70	858
	5	0.68	0.69	0.69	844
accuracy				0.73	5186
macro avg	0.73	0.73	0.73	0.73	5186
weighted avg	0.73	0.73	0.73	0.73	5186

Reporte de métricas de clasificación para la arquitectura ShallowConvNet:

Reporte de métricas de clasificación:

	precision	recall	f1-score	support
0	0.64	0.62	0.63	893
1	0.78	0.74	0.76	827
2	0.66	0.68	0.67	853
3	0.70	0.67	0.69	911
4	0.53	0.72	0.61	858
5	0.71	0.53	0.61	844
accuracy			0.66	5186
macro avg	0.67	0.66	0.66	5186
weighted avg	0.67	0.66	0.66	5186

En los últimos años, las redes convolucionales han demostrado ser el modelo de red neuronal artificial más apto para resolver tareas de clasificación de imágenes. Como se ha mencionado con anterioridad, este tipo de redes fueron diseñadas para tareas específicas de clasificación de imágenes, debido a esto se ha decidido implementar un modelo sencillo de este tipo de red con la base de datos descrita, con la diferencia de que la entrada o tipo de imagen serán mapas de calor correspondiente a cada una de las señales EEG, en lugar de una fotografía o dibujo.

Para comparar la eficacia de este modelo con la red EEGNet, se han desarrollado dos redes convolucionales, como se muestran en las figuras 6.46 y 6.47, construidas sobre el modelo denso con capas profundas dense o fully-connected. La última capa cuenta con solo seis neuronas, debido a que se desea clasificar seis categorías mencionadas anteriormente. Las otras capas densas funcionan como extractores de parámetros. El número de neuronas se reduce de una capa a la siguiente con el objetivo de reducir las posibles salidas predichas progresivamente. El primer diseño cuenta con una única capa convolucional seguida de una reducción de dimensiones (Max Pooling) y un descarte aleatorio de datos. El segundo diseño y ligeramente más profunda, incluye una segunda capa convolucional con mayor número de filtros entre la primera reducción de dimensiones y la capa de dropout.

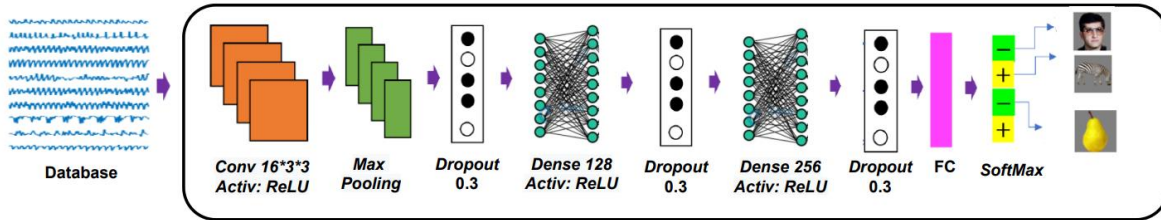


Figura 6.47. Esquema del modelo convolucional simple.

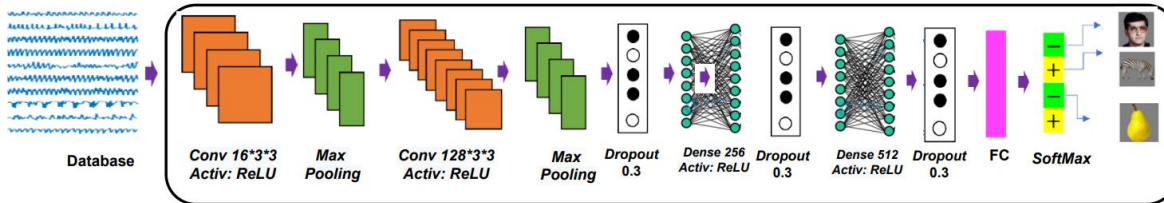


Figura 6.46. Esquema del modelo convolucional doble.

Para realizar múltiples pruebas al variar los parámetros de la red, así como los porcentajes de abandono se utilizaron recursos de hardware y software proporcionados de manera gratuita por la empresa Google, así como por la Benemérita Universidad Autónoma de Puebla (BUAP).

Google Colaboratory es una plataforma online de la empresa Google, que pone a disposición de los usuarios una GPU gratuita Tesla K80. Se trata de un programa de investigación que busca facilitar el acceso a un entorno de aceleración para los desarrolladores de inteligencia artificial. Cuenta con una memoria RAM limitada a 11.17 GB, pero permite una ejecución de la librería Tensorflow hasta 10 veces más rápida que la realizada mediante una CPU (Pelayo & Gordillo, 2018).

Asimismo, la BUAP cuenta con recursos a disposición de los alumnos, como lo es la Supercomputadora Cuetlaxcoapan del Laboratorio Nacional de Supercómputo (LNS) del Sureste de México. Está compuesta por 136 nodos estándares de cálculo, 6 nodos especiales y un sistema de almacenamiento de 520 TB, interconectados en una red Infiniband con topología de estrella mediante un switch central. La unidad básica de ejecución es un core de procesamiento. Cada nodo estándar de cálculo posee 24 cores (2 sockets Intel Xeon E5-2680 v3 a 2.5 Ghz con 12 cores por socket) y 128 GB de memoria RAM compartida. Por lo tanto, el tiempo de cálculo se mide en horas-core. Esta supercomputadora permite realizar una conexión inalámbrica por protocolo de comunicación SSH (Secure SHell), desde nuestra computadora en casa hasta el servidor, lo que facilita enormemente el tiempo de procesamiento, ya que nos presenta una velocidad de hasta 27 veces más rápida que la realizada con un CPU, y hasta 3 veces más rápida que la GPU proporcionada por Google.

Gracias a la velocidad de procesamiento de la supercomputadora de la BUAP, se realizaron entrenamientos de las redes de hasta 500 épocas, para observar el comportamiento con un alto número de épocas y conocer cuando estos convergen y se estabilizan, con el objetivo de saber cuántas épocas son necesarias para su entrenamiento, garantizando la continuidad de este.

Se comparó la exactitud de cada uno de los modelos al clasificar las imágenes durante el entrenamiento al variar los hiper parámetros de las capas dense (el número de neuronas) y el porcentaje de las capas dropout, y con base a estos resultados se ha determinado que la mejor combinación es la que se muestra en la figura 6.48.

Si se ajusta un número alto de neuronas en la primera capa densa, se consigue una mayor exactitud del modelo, pero la contraparte es que si este número es demasiado alto, se produce sobreaprendizaje del modelo. Por otro lado, si la cantidad de neuronas es baja, el modelo requerirá un tiempo mayor para alcanzar la convergencia, siendo esta inferior a la óptima debido a que el modelo puede aprender un número menor de parámetros. En las figuras 6.48 a) y 6.48 b) se presentan las gráficas de exactitud de las redes propuestas, y se observa que se alcanza una exactitud del 79.02% para el modelo convolucional simple (a)), y una exactitud del 82.15% para el modelo convolucional doble (b)).

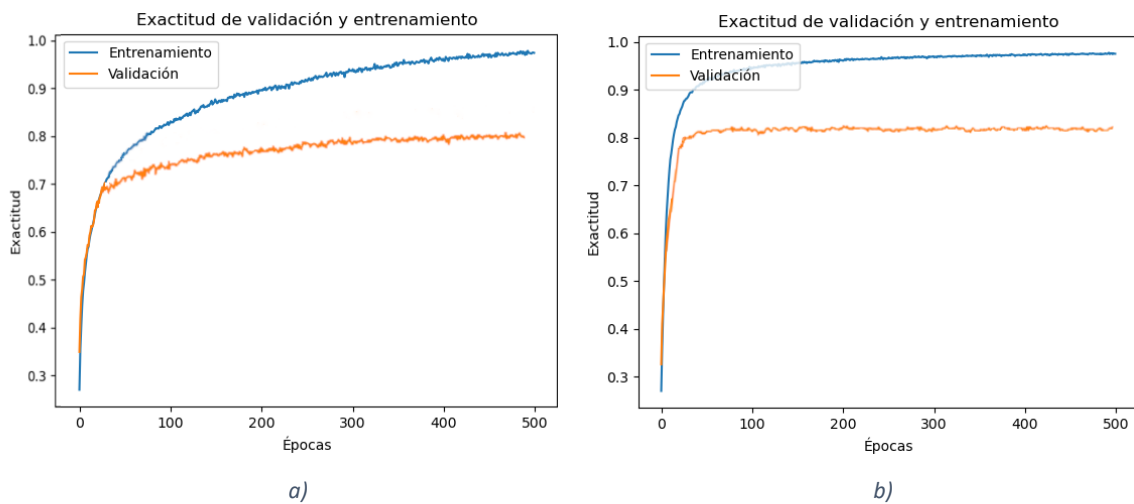


Figura 6.48. a) Gráfica de exactitud de la arquitectura de una red convolucional simple para datos EEG. b) Gráfica de exactitud para la arquitectura de una red convolucional doble.

A continuación se presenta una tabla comparativa del tiempo de ejecución de las épocas para el modelo de red neuronal presentado en la Fig. 6.48 con los diferentes hardwares implementados para estos experimentos, y observar la diferencia en tiempo de ejecución de cada una.

Finalmente se presenta nuevamente una tabla comparativa entre el número de parámetros de los diseños de redes neuronales convolucionales implementadas en esta tesis, para analizar la y comparar la exactitud que presenta cada una de estas, así como observar las dimensiones que tienen, y cómo varia la exactitud al modificar la estructura de cada una, el tipo de capas que incluyen y la cantidad de estas. En la tabla 6.9 se presentan los detalles.

Tabla 6.9. Comparación entre los tiempos de ejecución de una CPU y diferentes GPU's.

Tiempos de ejecución	
Acelerador por hardware	Tiempo
Ninguno (CPU)	2.99 ms/step o 136 s/epoch
GPU Tesla K80	308 us/step o 14 s/epoch
Supercomputadora del LNS	110 us/step o 5 s/epoch

Tabla 6.10. Parámetros evaluados en las CNN y exactitud del entrenamiento.

	<i>EEGNet</i>	<i>ShallowConvNet</i>	<i>DeepConvNet</i>	<i>ConvSimple</i>	<i>ConvDouble</i>
<i>Parámetros</i>	2,934	199,766	507,631	1,882,854	5,737,446
<i>Optimizador</i>	Adam	Adam	Adam	Adam	Adam
<i>Exactitud</i>	66.06	72.73%	73.19%	79.02%	82.15%

6.10. Pruebas de laboratorio

Como se mencionó anteriormente en este capítulo, debido a la contingencia actual fue imposible generar una base de datos propia, ya que esto implica inversión de tiempo por cada una de las personas que participarían en el experimento previamente diseñado en la sección 6.3. Y alineándonos a los protocolos de sanidad para prevenir contagios, resultó prácticamente imposible trabajar con las personas. Debido a la reducción de cifras en dichos contagios últimamente, nos dimos a la tarea de realizar el experimento con una sola persona, por cuestiones de tiempo y para analizar los datos generados por el paradigma propuesto.

Como se mencionó en la sección 6.3 y 6.4, el experimento se programa en el software psychopy3, es un paquete de software de código abierto escrito en el lenguaje de programación Python principalmente para su uso en investigación en neurociencia y psicología, que resulta bastante versátil debido a que cuenta con un compilador de programación con bloques para diseñar los experimentos.

Otra de las ventajas de este software es que es completamente compatible con las diademas Emotiv, ya que cuenta con componentes de registro de señales de estas diademas, los cuales hacen que Psychopy se conecte a estas para poder enviar los marcadores al flujo de datos (Peirce & Ltd, 2022). Este componente debe añadirse al inicio del paradigma, antes de que se presente cualquier estímulo en la parte superior del primer ensayo del experimento para comenzar la grabación. Además es posible insertar marcadores en el flujo de datos del EEG (ver Fig. 6.45) en el momento en que se presentan los estímulos. Estos marcadores se pueden añadir con etiquetas y valores, así como especificar un intervalo de tiempo enviando un marcador de parada.

De esta manera, al grabar los datos, las señales estarán etiquetadas automáticamente dependiendo de los valores configurados, así como los marcadores ajustados al inicio y final de cada estímulo, lo que permite que en el procesamiento de las señales sea más fácil segmentar las regiones de interés de las señales, teniendo sincronizado el inicio y fin del estímulo con estos marcadores. De esta forma tenemos diseñado el experimento de una forma rápida y con pocos elementos. En la Fig. 6.49 se presenta una imagen del diseño del experimento en la plataforma psychopy con los bloques mencionados.



Figura 6.49. Diseño del experimento en la interfaz psychopy a través de bloques de programación y componentes de Emotiv.

En la siguiente figura se muestra una secuencia del paradigma propuesto desplegado en el software psychopy:

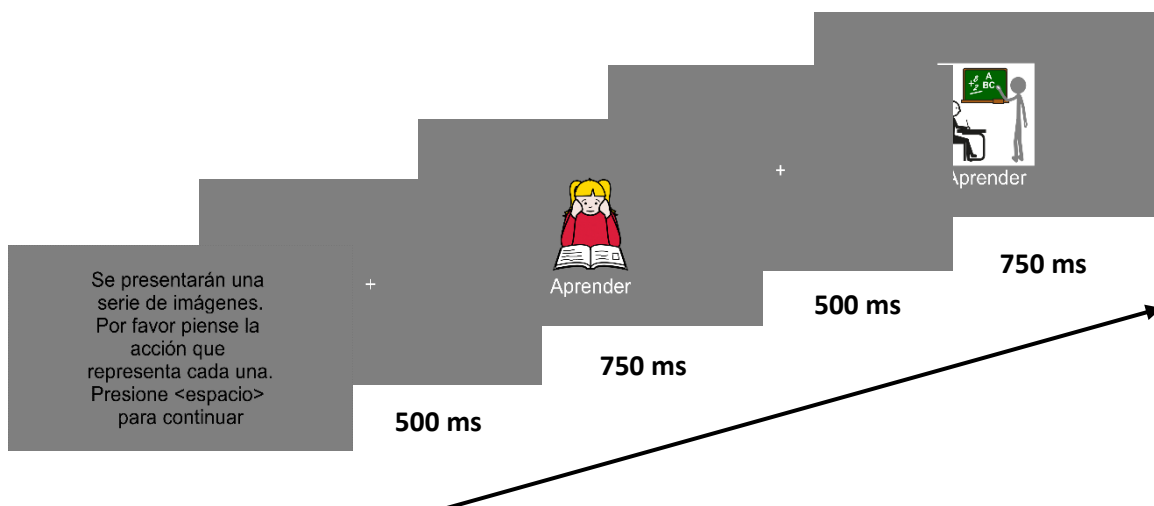


Figura 6.50. Visualización del paradigma diseñado para la presentación de estímulos a los usuarios.

Para este experimento se implementó una diadema de contacto en seco, Emotiv Insight de 5 electrodos (ver características en la tabla 5.1) y a través del software EmotivPro se grabaron las señales generadas al observar los estímulos. En la Fig.6.51 se presenta la recopilación de señales a través de este software, el cual nos permite exportar los datos en formatos “.csv” o “.edf”, así como un archivo adicional que contiene los marcadores del inicio y fin de los estímulos sincronizados con las marcas de tiempo.

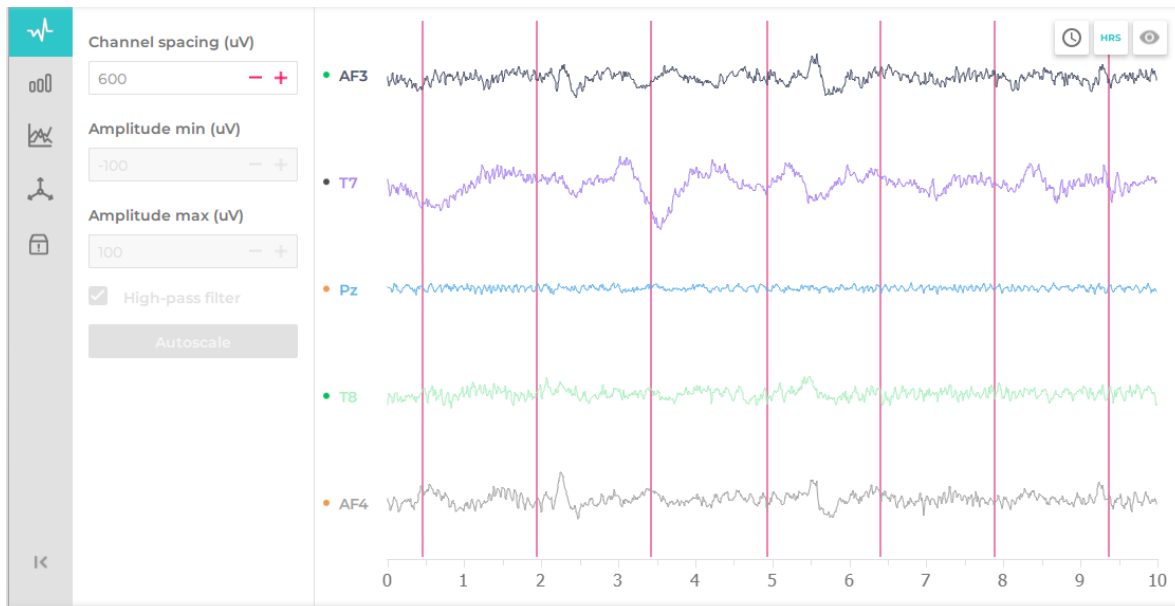


Figura 6.51. Adquisición de las señales EEG de los 5 canales de la diadema Insight, a través del software EmotivPro.

Como se mencionó, psychopy tiene la versatilidad de adjuntar los marcadores temporales para el inicio y fin de los estímulos en un archivo externo, esto permite visualizar y segmentar los datos de una forma más sencilla, donde la trama de datos se divide en cada intervalo de tiempo definido, y se adjunta la etiqueta correspondiente al estímulo presentado. En el ejemplo de la Fig. 6.52 se presenta la grabación de algunos datos donde el usuario observa el concepto “Aprender”, y en cada época segmentada se anexa la leyenda “AP” correspondiente a este estímulo. El preprocesamiento de los datos se realizó en la caja de herramientas de Matlab, EEGLAB.

Para seguir analizando los datos, se convierten del dominio del tiempo al dominio de la frecuencia mediante una FFT, pero antes de realizar una Transformada Rápida de Fourier (FFT) es necesario eliminar el desplazamiento de CC de los datos.

Uno de los métodos de eliminación comúnmente utilizado es restar el valor medio de todo el canal de datos. Este método puede dar lugar a un aumento del ruido de baja frecuencia. Lo ideal es aplicar también un filtro de paso alto que se ajuste a las características de la electrónica, es decir, utilizar un filtro de paso alto de primer orden de 0,16 Hz para eliminar la señal de fondo (esto también elimina cualquier deriva a largo plazo, lo que no se consigue con el método de sustracción del promedio).

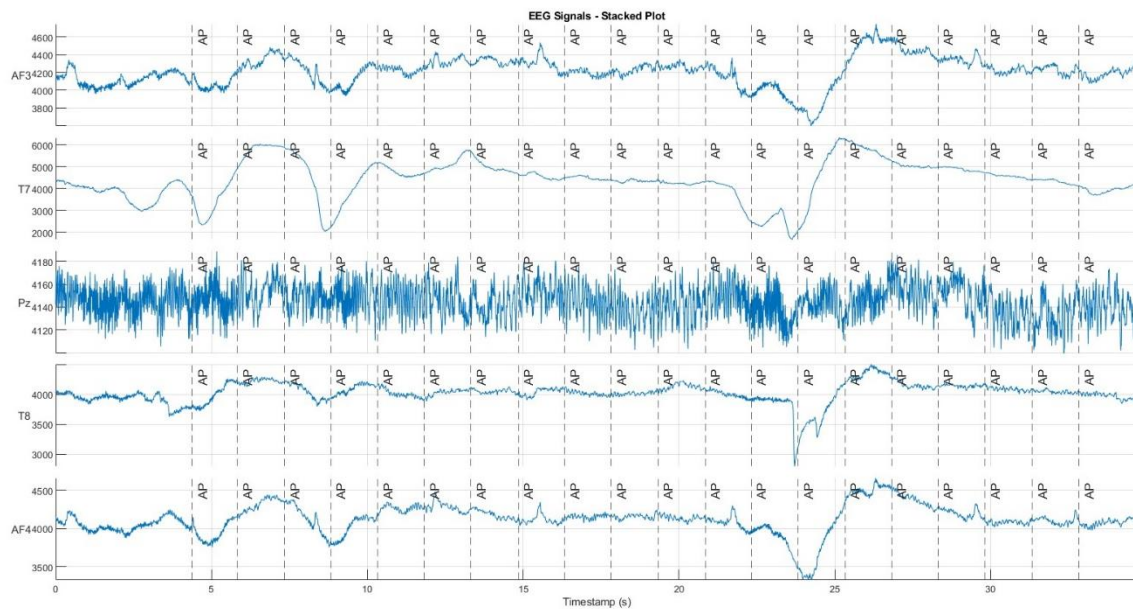


Figura 6.52. Representación de los canales EEG segmentados en inicio y fin de cada estímulo, para un procesamiento más eficiente.

Para el resto del preprocesamiento se implementaron los pasos descritos a partir de la sección 6.6, con la ventaja de tener el protocolo previamente diseñado para el tratamiento de los datos.

Debido al tiempo necesario para la recopilación de datos de múltiples usuarios y a la contingencia actual, no fue posible realizar la base de datos propuesta en la metodología, pero se realizó todo el procedimiento con la base de datos pública que se menciona en la sección 6.5, lo que genera una estructura diseñada para el procesamiento y análisis de futuras grabaciones de EEG, como resultó en el caso de la recopilación de señales de un usuario aquí mencionada. Se pretende posteriormente completar una base de datos propia orientada específicamente al proceso de comunicación humana, e implementar la estructura de los algoritmos aquí desarrollados para su análisis y clasificación.

CAPÍTULO 7

7.Conclusiones

Los datos actuales muestran que la categoría de objeto se puede decodificar con éxito a partir de los primeros componentes visuales del EEG del cuero cabelludo. Esta contribución es relevante para la investigación de la interfaz cerebro-computadora. Para los análisis de datos que involucran fMRI, el alto costo de los estudios puede hacer que tales exploraciones sistemáticas no sean prácticas. Afortunadamente, estudios como este que demuestran que el conocimiento conceptual y el análisis de categorías semántico pueden analizarse mediante EEG, representan un área de investigación potencial, en la que se puede profundizar en métodos de filtrado de las señales adquiridas, en la combinación de metodologías existentes con el fin de crear procedimientos de análisis híbridos y evaluar su precisión. Dado que los estudios de EEG se pueden realizar a un costo mucho menor que la resonancia magnética funcional, pueden ser una metodología más factible para la investigación léxica y la categorización a gran escala.

Por el contrario, algunos autores destacan que las técnicas de EEG implican numerosos esfuerzos para mejorar la precisión de la ubicación de la fuente neuronal, y que la información obtenida en algunos de los experimentos concretados no es suficiente para proporcionar una estimación completa y robusta de la distribución espacial de las respuestas neuronales que subyacen a la percepción de diferentes tipos de objetos; pero que existen métodos y aportes específicos para este tipo de señales, en los que se involucran técnicas de clasificación de patrones a través del aprendizaje automático.

El protocolo de procesamiento diseñado en esta tesis presenta soluciones aceptables para la categorización de las señales generadas por la visualización de conceptos, en el cuál se argumentan las siguientes observaciones:

- En la etapa de filtrado se ha presentado un algoritmo de implementación de filtro digital preciso y eficiente para eliminar el ruido / artefactos de las señales de EEG. Los parámetros de rendimiento (relación señal / ruido y correlación cruzada) se han calculado para evaluar el mejor desempeño de las estructuras de filtro diseñadas. Los resultados muestran que un filtro FIR de orden 4 a 5 posee una mejor capacidad de filtrado que el filtro IIR. Un filtro FIR siempre dará una salida estable, ya que no es recursivo, además

de tener un retraso constante en todas las frecuencias; la forma de la señal no se verá influenciada por los cambios de fase.

- Cuando un artefacto tiene una distribución estable en el cuero cabelludo, generalmente es posible utilizar ICA para descomponer los datos del EEG en un conjunto de componentes subyacentes y luego reconstruir los datos sin el componente correspondiente al artefacto. Este método de corrección de artefactos se ha vuelto muy popular para eliminar los artefactos de parpadeo, que tienen una distribución del cuero cabelludo muy estable y son tan grandes que generalmente aparecen como un solo componente ICA.

Sin embargo, la corrección de artefactos basada en ICA tiene algunas limitaciones importantes. Primero, ICA no puede funcionar para artefactos que no tienen una distribución consistente en el cuero cabelludo para un sujeto dado (por ejemplo, potenciales cutáneos). En segundo lugar, el número de muestras de tiempo necesarias para que ICA aísle de manera eficaz los componentes es una función del cuadrado del número de electrodos, y los estudios donde se utiliza hardware con un alto número de electrodos pueden carecer del número de muestras de tiempo necesarias. Como tercer punto, el número de componentes ICA es siempre igual al número de electrodos, y el hecho de que el número de fuentes es mayor que el número de electrodos introduce algún error, aunque este error puede ser insignificante en muchas condiciones reales. Por último, aunque varios estudios han validado la eficacia de ICA para la corrección del parpadeo, para otros tipos de artefactos hay menos estudios de validación.

- En este proyecto se ha empleado una base de datos reducida que se obtuvo para otra línea de investigación, con la que se consiguió en primera instancia un porcentaje de acierto en la clasificación de un 40 % aproximadamente, empleando el método de Análisis de Componentes Principales (PCA). Se ha conseguido mejorar la tasa de acierto en aproximadamente un 30% al emplear redes neuronales profundas con capas convolucionales.
- Se ha detectado en las matrices de confusión, una tendencia a clasificar con un mayor acierto las imágenes pertenecientes a la categoría de ‘Rostro humano’. Esto se debe a que los seres humanos estamos acostumbrados a reconocernos entre nosotros, lo que genera una respuesta más marcada y distintiva en nuestro cerebro al visualizar un rostro humano.
- Se evalúan los resultados obtenidos a través de todo el procedimiento, y al final se implementan modelos de redes neuronales convolucionales, que presentan un incremento en la exactitud considerable, en comparación con los métodos clásicos de machine learning, alcanzando exactitudes cercanas al 80%. El diseño de este tipo de redes es completamente basado en el tipo de datos de entrada, ya sea que las señales EEG se tomen como series temporales directamente, o se conviertan a imágenes basadas en mapas de calor y utilizar las CNN exactamente como si se trabajara con fotos o imágenes reales. La gran ventaja que se tuvo en el análisis de estas arquitecturas fue la implementación de

GPU's que permitieron un entrenamiento en tiempos cortos, y de manera simultánea permitió entrenar las redes con hasta 500 épocas.

Además, se ha podido observar que un mismo modelo puede generar respuestas muy variadas en función del sujeto. Esto se debe a la diferencia existente entre cerebros de personas distintas. Como consecuencia, para aplicar el aprendizaje profundo a un problema de este tipo, se requiere un alto número de datos para su análisis, así como con un modelo adecuado de red y sistemas de validación robustos, como el de validación cruzada, capaz de extraer características generales a varios sujetos, pudiendo generalizar los resultados obtenidos de esta forma.

Sin embargo, aunque se ha logrado mucho progreso para los BCI no invasivos, las señales cerebrales disponibles que pueden usarse como señales de control a la vez son aún limitadas. Creemos que son necesarias exploraciones sistemáticas a gran escala del léxico mental y la categorización con datos neuronales, que implican tanto un análisis más cuidadoso de las separaciones conceptuales como un número mayor de categorías, porque la mayoría de los artículos presentan análisis comparativos entre un número reducido de categorías, además de que ciertas categorías están completamente aisladas unas de otras. En este sentido, el aporte principal de esta tesis es fue realizar estudios y experimentos que permitan conocer si es posible identificar conceptos a través de señales neuronales, asociados al proceso de comunicación humana, es decir, ver la posibilidad de detectar una acción completa relacionada al lenguaje humano.

El sistema visual humano es capaz de reconocer los objetos rápidamente y la actividad neuronal del cerebro genera señales que proporcionan información sobre las categorías de objetos vistos por los sujetos, los resultados proporcionan información útil sobre los canales y la parte de las señales que son afectados por diferentes categorías de objetos en términos de señales cerebrales.

En términos de procesos neuronales de semántica, se han identificado ubicaciones del cuero cabelludo, intervalos de tiempo y bandas de frecuencia que entregan información valiosa sobre las diferencias de categorías. El objetivo final del análisis de las señales EEG es lograr una respuesta razonable directamente a la intención del usuario. La tecnología actual y los procesos involucrados en la interfaz cerebro-computadora se basan en el acceso directo a la información de actividad física en los procesos de pensamiento del cerebro humano, proporcionando un camino neurológico efectivo capaz de interpretar las señales cerebrales. Esto se ha convertido en una importante oportunidad de desarrollo en el campo del procesamiento del lenguaje natural.

Referencias

- AAMODH, S. (2016). *Body swarm interface (BOSI): controlling robotic swarms using human bio-signals*. BOSTON UNIVERSITY.
- Abdelhakim, R., & Daoud, B. (2017). EEG Signals Classification Based on Time Frequency Analysis. *Journal of Circuits, Systems, and Computers*.
- Aguilar, E. (2013). Plataforma para la interacción con dispositivos de electroencefalografía.
- Ahmed, H., Wilbur, B., R, B., M, H., & Siskind, J. (2020). Object classification from randomized EEG trials.
- Alejandro, T. (2016). Análisis y clasificación de electroencefalogramas (EEG) registrados durante el habla imaginada.
- Andrei, S., Barry, H., Vladimir, N., Rihana, W., Stephen, F., & Braun, R. (202). fMRI Study Comparing Names Versus Pictures. *Human Brain Mapping* 16, 168-175.
- Angana, S., & Sudip, P. (2019). EEG Signal Processing and Its Classification for Rehabilitation Device Control. En *Application of Biomedical Engineering in Neuroscience*. Springer Nature Singapore.
- Apdullah, Y. (2017). Enhancing Extreme Learning Machine: Novel Extensions and Applications to Optimizing Visual Stimuli for Brain Computer Interface.
- Aramayo, S. (25 de Abril de 2016). *Diagnóstico de trastornos orgánicos a nivel neurológico en niñas y adolescentes víctimas de violencia sexual como método optimizador en intervención psicológica*. Obtenido de ENDAELALTO:
http://www.endaelalto.org/archivos/Suplemento%20Boletin9_Trastornos%20organicos%20a%20nivel%20neurol%C3%B3gico.pdf
- Arango, C. P. (2004). EL CEREBRO: DE LA ESTRUCTURA YLA FUNCIÓN A LA PSICOPATOLOGÍA. *Revista Colombiana de Psiquiatría, Suplemento No. 1, Vol. 33*, 102-125.
- ARASAAC. (2014). ¿QUÉ SON LOS SISTEMAS AUMENTATIVOS Y ALTERNATIVOS DE COMUNICACIÓN (SAAC)? Obtenido de Portal Aragonés de la Comunicación Aumentativa y Alternativa:
<http://www.arasaac.org/aac.php#1>
- ARASAAC. (2016). *Portal Aragonés de la Comunicación Aumentativa y Alternativa*. Obtenido de PICTOGRAMAS, IMÁGENES, LOCUCIONES, VÍDEOS Y FOTOGRAFÍAS EN LSE:
<http://www.arasaac.org/aac.php#1>
- Argopsique. (31 de Enero de 2014). Obtenido de ARGOPSI, psicología y algo más:
<https://psicoteraphy.wordpress.com/2014/01/31/hemisferios-cerebrales-caracteristicas-y-funciones/>
- Asaid, K., Alan, J., Christoph, M., Theodor, L., & Jean-Marie, A. (2002). Dynamics of Brain Activation During an ExplicitWord and Image Recognition Task: An Electrophysiological Study. *Brain Topography*, 197-212.

- Asly, S., Gilde, M., Moctezuma, L., & Molinas, M. (2019). TOWARDS EEG-BASED SIGNALS CLASSIFICATION OF RGB. *Brain-Computer Interface Conference*.
- Bacary, G., Alice, O., & Amine, N.-a. (2020). EEG Biometrics for Person Verification. 45-69.
- Bashivan, P., Rish, I., Yeasin, M., & Codella, N. (2016). Learning representations from EEG with deep recurrent-convolutional neural networks. *International Conference on Learning Representations (ICLR)*.
- Bautista, S., Hervás, R., Hernández, A., Martínez, C., Pascua, S., & Gervás, P. (2017). AraTraductor: Text to Pictogram Translation using Natural Language Processing Techniques. *Interacción 2017. Association for Computing Machinery*.
- Bedoya, A., Giraldo, J., Torres, I., & Becerra, M. (2013). Interfaz Cerebro Computador, basado en señales EEG para el control de movimiento de una prótesis de mano usando ANFIS. 59-64.
- Behroozi, Mehdi, Daliri, mohammad, R., Shekarchi, & Babak. (2015). EEG phase patterns reflect the representation of semantic categories of objects. *Medical & biological engineering & computing*. 54.
- Binder, J., Desai, R., Graves, W., & Conant, L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19, 2767–2796.
- Bird, J. J., Manso, L. J., Ribeiro, E. P., Ekárt, A., & Faria, D. R. (2018). A Study on Mental State Classification using EEG-based Brain-Machine Interface . *International Conference on Intelligent Systems (IS)*.
- Blair, K., Marcos, P., Hyung-Suk, K., Anthony, M., & Patrick, S. (2015). A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. *PLOS ONE*.
- Calvo, Hiram, Paredes, José, Figueroa, & Nazuno, J. (2018). Measuring Concept Semantic Relatedness through Common Spatial Pattern Feature Extraction on EEG Signals. . *Cognitive Systems Research*.
- Chan, M. (2011). *Informe mundial sobre la discapacidad*. Suiza.
- Chand, B., Lamichhane, B., & Dhamala, M. (2016). Face or House Image Perception: Beta and Gamma Bands of Oscillations in Brain Networks Carry Out Decision-Making. . *Brain Connectivity*, 621–631.
- Chao, L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category related cortical activity. . *Cerebral Cortex*.
- Cichy, Radoslaw, Pantazis, & Dimitrios. (2017). Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. . *NeuroImage*. 158.
- Dagallier, A., & Campillos, E. (2018). Análisis de registros de EEG mediante redes neuronales.
- Damasio. (2010). *Psicología de la percepción visual*. Obtenido de Ondas cerebrales.

- Delorme, A. (2018, Mayo 23). *EEGLAB*. Retrieved from EEGLAB preprocessing part 5: Visualizing data : <https://www.youtube.com/watch?v=SgQxdVgryVY&t=5s>
- Delorme, A. (23 de Marzo de 2020). *EEGLAB*. Obtenido de EEGLAB. Obtenido de ICA applied to EEG part 1: What is ICA?: <https://www.youtube.com/watch?v=kWAjhXr7pT4>.
- Dugarte, N., Cristobal, S., Álvarez, A., González, A., Perez, S., Sánchez, J., . . . Álvarez, G. (2018). Avances en el Proyecto para detección precoz y neurofeedback en personas con síndrome del espectro autista TEA.
- Emotiv. (2013). *Emotiv*. Obtenido de <https://www.emotiv.com/>
- F, L., L, B., A, C., M, C., & M, C. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *Journal of Neural Engineering*.
- Fabien, L. (2014). A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces. In R. M. Eduardo, & C. Julien, *Guide to Brain-Computer Music Interfacing*.
- Fabien, L. (2015). Signal processing approaches to minimize or suppress calibration time in oscillatory activity based Brain Computer Interfaces. *Proceedings of the IEEE*, 871-890.
- Fabien, L., Marco, C., Anatole, L., Fabrice, L., & Bruno, A. (2007). A Review of Classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 24.
- Fares, A., Zhong, S., & Jiang, J. (2018). Region level Bi-directional Deep Learning Framework for EEG-based Image Classification. *International Conference on Bioinformatics and Biomedicine (BIBM)*, 368-373.
- Fernández, J., Pons, A., Bonilla, C., Calero, M., Estévez, A., Fernández, E., . . . González, V. (2010). *Trastornos del desarrollo con discapacidad intelectual*. Andalucía, España.
- Fernandino, L., Humphries, C., Seidenberg, M., Gross, W., Conant, L., & Binder, J. (2015). Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes. *Neuropsychologia*, 76, 17-26.
- Folstein, J., & Monfaredm, S. (2019). Extended categorization of conjunction object stimuli decreases the latency of attentional feature selection and recruits orthography-linked ERPs. *CORTEX*.
- Gao, S., Wang, Y., Gao, X., & Hong, B. (2014). Visual and Auditory Brain-Computer Interfaces. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 61, NO. 5*, 1436-1447.
- García, E., Cruz, N., & Coria, G. (2015). Análisis de ondas cerebrales para determinar emociones a partir de estímulos visuales.
- Garzón, J., Rojas, S., Sanabria, C., & Montenegro, D. (2013). Uso de una BCI (Brain Computer Interface) como enlace interactivo, terapéutico y de aprendizaje dirigido a personas con discapacidad cognitiva y motora. *Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI'2013)*, 1-10.

- Gerlach, C. (2007). A review of functional imaging studies on category specificity. *Journal of Cognitive Neuroscience*, 296–314.
- Germán, A., Pressel, C., Iván, G., & Leonardo, R. (2017). Open Access database of EEG signals recorded during imagined speech. *12th International Symposium on Medical Information Processing and Analysis*.
- Gifford, A., Dwivedi, K., Roig, G., & Cichy, R. (2022). A large and rich EEG dataset for modeling human visual object recognition. *bioRxiv* 2022.
- Gómez, L., García, C., & Roque, S. (2016). Análisis de señales EEG para detección de eventos oculares, musculares y cognitivos.
- Greene, M., & Hansen, B. (2018). Shared spatiotemporal category representations in biological and artificial deep neural networks. *PLoS Comput Biol* 14(7): e1006327.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annu. Rev. Neuroscience*. 49–77.
- Grootswagers, T., Zhou, I., Robinson, A. K., Hebart, M., & Carlson, T. (2022). Human EEG recordings for 1,854 concepts presented in rapid serial visual presentation streams. *Scientific Data*.
- Grootswagers, Tijn, Robinson, Amanda, Carlson, & Thomas. (2018). The representational dynamics of visual objects in rapid serial visual processing streams.
- Hanson, S., Matsuka, T., & Haxby, J. (2004). Combinatorial codes in ventral temporal lobe for object recognition. *NeuroImage* 23, 156–166.
- Hauke, D., Jakob S., M., Helle K., I., & Sadasivan, P. (2018). An End-to-end Deep Learning Approach to MI-EEG Signal Classification for BCIs. *Expert Systems With Applications*.
- Helbig, H., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. . *Experimental brain research*.
- Hillyard, S. A. (2017). Event-Related Potentials (ERPs) and Cognitive Processing☆. . *Reference Module in Neuroscience and Biobehavioral Psychology*.
- Hinojosa, José, A., Moreno, Eva, Ferré, & Pilar. (2019). Affective neurolinguistics: towards a framework for reconciling language and emotion. *Language. Cognition and Neuroscience*, 1-27.
- Hornero, R., Corralejo, R., & Álvarez, D. (2012). Brain-Computer Interface (bci) aplicado al entrenamiento cognitivo y control domótico para prevenir los efectos del envejecimiento. *Cuadernos de la Fundación General CSIC | Nº 8 | LYCHNOS*, 29-34.
- Hsin-Mei, S., Stephanie, L., Robert, D., & Wolfgang, A. (2011). Contextual influences on rapid object categorization in natural scenes. *BRAIN RESEARCH*, 40-54.
- iMotions. (2019). *Electroencephalography*. iMotions.
- INEGI. (2020). *INEGI*. Obtenido de <https://cuentame.inegi.org.mx/poblacion/discapacidad.aspx>

- Irina, S., Marcel, G., Robert, O., & Peter, H. (2010). Identifying Object Categories from Event-Related EEG: Toward Decoding of Conceptual Representations. *PLOS ONE*.
- Isik, L., Meyers, Ethan, Leibo, Joel, Poggio, & Tomaso. (2013). The dynamics of invariant object recognition in the human visual system. . *Journal of neurophysiology*.
- James, D., Davide, Z., & Nicole, R. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron perspective*, 415-434.
- Jia, W., Huang, D., Luo, X., Pu, H., Chen, & X. Bai, O. (2012). Electroencephalography(EEG)-Based Instinctive Brain-Control of a Quadruped Locomotion Robot. *34th Annual International Conference of the IEEE EMBS*, 1777-1781.
- Kaneshiro, B., M, P. G., H-S, K., AM, N., & P, S. (2015). A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. *PLoS ONE 10(8): e0135697*.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. . *Philos Trans R Soc Lond B Biol Sci* 361, 2109–2128.
- Khatteer, A., Bansal, D., & Mahajan, R. (2017). Study of Various Automatic EEG Artifact Removal Techniques. *International Journal for Research in Applied Science and Engineering Technology*, 1027-1037.
- Khatteer, A., Bansal, D., & Mahajan, R. (2019). Performance Analysis of IIR & FIR Windowing Techniques in Electroencephalography Signal Processing.
- Kiefer, M., Sim, E., Herrnberger, B., J, G., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J Neuroscience* .
- Kozunov, Vladimir, Nikolaeva, Anastasia, Stroganova, & Tatiana. (2018). Categorization for Faces and Tools—Two Classes of Objects Shaped by Different Experience—Differs in Processing Timing, Brain Areas Involved, and Repetition Effects. *Frontiers in Human Neuroscience*.
- Kumar, S., Sharma, A., Mamun, K., & Tsunoda, T. (2016). A deep learning approach for motor imagery EEG signal classification. *Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, 34–39.
- Kyung, H., Hyo, W., & Hyun, W. (2005). Event-related Potential Study of Brain Activation during Word/Pictogram. *IEEE*.
- Lang, S. (2014). *Toki Pona: The language of good*. Obtenido de Wikipedia: https://en.wikipedia.org/wiki/Toki_Pona
- Lawhern, V., Solon, A., Waytowich, N., Gordon, S., Hung, C., & Lance, B. (2018). EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer-Interface. *Journal of Neural Engineering*.
- Leonardo, G. (2016). Análisis de señales EEG para detección de eventos oculares, musculares y cognitivos.

- Lingoda. (2019). Obtenido de Lingoda: <https://www.lingoda.com/es/espanol/vocabulario>
- Lopez, C. J., & J.Luck, C. a. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 213.
- Maguire, M., Brier, M., & Ferree, T. (2010). EEG theta and alpha responses reveal qualitative differences in processing taxonomic versus thematic semantic relationships. . *Brain and Language*, 16–25.
- Mahon, B., & Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annu Rev Psychol* 60, 27–51.
- Manzano, S., Placencia, F., Ayala, P., Chicaiza, D., & Carrera, A. (2017). Controlador Basado en FPGA para Autómatas Utilizando Impulsos Neuronales Alfa y Beta.
- Marini, F., Breeding, K., & Snow, J. (2019). Distinct visuo-motor brain dynamics for real-world objects versus planar images. *NeuroImage*, 232–242.
- Martin, A. (2007). The Representation of Object Concepts in the Brain. *Annual Review of Psychology* 58, 25–45.
- Martín, J., & Martínez, I. (2015). Sistema Brain Computer Interface.
- Medina, B., Sierra, J., & Barrios, A. (2018). Técnicas de extracción de características de señales EEG en la imaginación de movimiento para sistemas BCI.
- Minguez, J. (2010). Tecnología de Interfaz Cerebro - Computador. *Jornadas II Internacionales de Mayores y Nuevas Tecnologías*. Obtenido de <https://webdiis.unizar.es>
- Mitchell, J., Heatherton, T., & Macrae, C. (2002). Distinct neural systems subserve person and object knowledge. 38–43.
- MN, H., Dickter AH, K. A., WY, K., A, C., C, V. W., & CI, B. (2019). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *Plos One*.
- Mohammad, R., Mitra, T., & Kavous, S. (2012). EEG Signature of Object Categorization from Event-related Potentials. . *Journal of Medical Signals & Sensors*. , 37-44.
- Mohedano, E., Healy, G., McGuinness, K., Giró-i-Nieto, X., O'Connor, N., & Smeaton, A. (2014). Object Segmentation in Images using EEG Signals. 417-426.
- Morán, A. (2015). *Diseño de interfaces cerebro-máquina controlados mediante registros de EEG*.
- Morillo, E. (2014). *ANÁLISIS VISUAL DEL ELECTROENCEFALOGRAMA*. Obtenido de Asociación Colombiana de Neurología: <https://www.acnweb.org>
- Moseley, Rachel, Pulvermüller, & Friedemann. (2014). Nouns, verbs, objects, actions, and abstractions: Local fMRI activity indexes semantics, not lexical categories. . *Brain and language*. 132C., 28-42.

- Murphy, B., Poesio, M., Bovolo, F., Bruzzone, L., Dalponte, M., & Lakany, H. (2011). EEG decoding of semantic category reveals distributed representations for single concepts. *ELSEVIER*, 12-22.
- Nicolas, L., & Gómez, J. (2012). Brain Computer Interfaces, a Review . *Sensors*, 1211-1279.
- Noppeney, U., Price, C., Penny, W., & Friston, K. (2006). Two distinct neural mechanisms for category-selective responses. *Cereb Cortex* , 16: 437–445.
- Nunez, P. L. (2017). Reference Module in Neuroscience and Biobehavioral Psychology. *Electroencephalography (EEG)*.
- Olague, G., Clemente, E., Hernández, D., Aaron, B., Chan-ley, M., & Bakshi, S. (2019). Artificial Visual Cortex and Random Search for Object Categorization. *Digital Object Identifier*, 54054-54072.
- Olias, J., M. Clemente, R., Sarmiento V, M. A., & Cruces, S. (2019). EEG Signal Processing in MI-BCI Applications With Improved Covariance Matrix Estimators. *NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 27*, 895-905.
- Olivares, S. (2019). *TRADUCCIÓN AUTOMÁTICA DE TEXTO A PICTOGRAMAS*. Puebla: Facultad de Ciencias de la Computacion, BUAP.
- OMS. (s.f.). Obtenido de Organización mundial de la salud:
<https://www.who.int/topics/disabilities/es/>
- OMS-. (2013, Septiembre 13). Retrieved from Organización mundial de la salud:
<https://www.who.int/features/qa/es/>
- OMS. (16 de Enero de 2018). Obtenido de Organización mundial de la salud:
<https://www.who.int/es/news-room/fact-sheets/detail/disability-and-health>
- Oñate. (2018). APLICACIÓN DEL DISPOSITIVO MINDWARE PARA LA ASISTENCIA REMOTA A PERSONAS CON DISCAPACIDAD FÍSICAS DEL INSTITUTO TECNOLÓGICO SUPERIOR STANFORD.
- Oñate, R. (2018). APLICACIÓN DEL DISPOSITIVO MINDWARE PARA LA ASISTENCIA REMOTA A PERSONAS CON DISCAPACIDAD FÍSICAS DEL INSTITUTO TECNOLÓGICO SUPERIOR STANFORD.
- Papadimitriou, A., Passalis, N., & Tefas, A. (2019). Visual representation decoding from human brain activity using machine learning: A baseline study. *Pattern Recognition Letters*, 38-44.
- Parekh, V., Subramanian, R., Roy, D., & C, V. (2018). An EEG-based Image Annotation System. *ResearchGate*.
- Patterson, K., Nestor, P., & Rogers, T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. . *Nat Rev Neurosci* 8, 976–987.
- Peirce, J., & Ltd, O. S. (2022). *Psychopy*. Retrieved from
https://www.psychopy.org/builder/components/emotiv_record.html

- Pelayo, C. E., & Gordillo, D. A. (2018). Análisis de registros de EEG mediante redes neuronales.
- Pérez, E., Minguillón, J., & López, M. (2018). Neuroingeniería lúdica e inclusión. *Researchgate*.
- Prieto, A., Pelayo, F., López, M., & Romero, S. (2013). Brain Computer Interfaces (BCI).
- Qingguo, M. X., Liangchao, S., & Shenyi, D. (2008). P300 and categorization in brand extension. *Neuroscience Letters*, 57–61.
- Regina, W., & I-Ning, L. (2016). Pictogram and ERP study - A point of view from friendly design. *IEEE*.
- Richard, G. (2012). Procedimiento operacional estandarizado (POE) de comunicación aumentativa-alternativa para pacientes intubados en UCI.
- Roldan, S. (2017). Object Recognition in Mental Representations: Directions for Exploring Diagnostic Features through Visual Mental Imagery. *Perception Science*.
- Roots, K., Muhammad, Y., & Muhammad, N. (2020). Fusion Convolutional Neural Network for Cross-Subject EEG Motor Imagery Classification. *Computers 2020*.
- Sabine, F., Matthias, S., & Ina, B. (2011). Conflicts in language processing: A new perspective on the N400–P600 distinction. *Neuropsychologia*, 574–579.
- Schalk, G., & Leuthardt, E. C. (2011). Brain-Computer Interfaces Using Electrocorticographic Signals. *IEEE Reviews in Biomedical Engineering*, 140–154.
- Schirrmeister, R. T., Springenberg, J. T., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., . . . Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 5391–5420.
- Schluter, N. (5 de Junio de 2019). *Towards datascience*. Obtenido de Don't Overfit! — How to prevent Overfitting in your Deep Learning Models: <https://towardsdatascience.com/dont-overfit-how-to-prevent-overfitting-in-your-deep-learning-models-63274e552323>
- Sciotto, A., & Niripil, B. (2018). ONDAS CEREBRALES, CONCIENCIA Y COGNICIÓN. *ResearchGate*.
- Shen, Y., Lu, H., & Jia, J. (2017). Classification of motor imagery EEG signals with deep learning models. *Intelligence Science and Big Data Engineering*, 181-190.
- Shinkareva, S., Mason, R., Malave, V., Wang, W., Mitchell, T., & Just, M. (2008). Using fMRI Brain Activation to Identify Cognitive States Associated with Perception of Tools and Dwellings. 1-9.
- Simanova, I., Gerven, M., Oostenveld, R., & Hagoort, P. (2010). Identifying Object Categories from Event-Related EEG: Toward Decoding of Conceptual Representations. *PLOS one*, 1-12.
- Simon, O., Stephen, F., & William, M. (2011). Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals. *Medical Engineering and Physics*.

- Singh, V., Veer, K., Sharma, R., & Kumar, S. (2016). Comparative study of FIR and IIR filters for the removal of 50 Hz noise from EEG signal. . *International Journal of Biomedical Engineering and Technology*, 22(3), 250.
- Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. . *Neuron*.
- Sylvain, B., Karl, F., & Oostenveld, R. (2011). Academic Software Applications for Electromagnetic Brain Mapping Using MEG and EEG. *Computational Intelligence and Neuroscience*.
- Tabar, R., & Halici, U. (2017). A novel deep learning approach for classification of EEG motor imagery signals. *Journal of Neural Engineering*.
- Tang, Z., Li, C., & Sun, S. (2017). Single-trial eeg classification of motor imagery using deep convolutional neural networks. 11–18.
- Tavakolan, M., Frehlick, Z., Yong, X., & Menon, C. (2017). Classifying three imaginary states of the same upper extremity using time-domain features. *PLoS ONE* 12(3): e0174161.
- Urigüen, J., & Garcia-Zapirain, B. (2015). EEG artifact removal—state-of-the-art and guidelines. *Journal of Neural Engineering*.
- Vargas, R. (30 de Diciembre de 2012). *Cerebros modificados*. Obtenido de Matemáticas y cerebro: <https://rv2010.wordpress.com/de-la-neurona-a-la-educacin/matematicas-y-cerebro/>
- Wu, Y., & Coulson, S. (2007). How iconic gestures enhance communication: An ERP study. *Brain and language*, 34-45.
- Xiang, L., Dawei, S., Peng, Z., Y. Z., Yuxian, H., & Bin, H. (2018). Exploring EEG Features in Cross-Subject Emotion Recognition. *Frontiers in Neuroscience*.
- Yang, Y., Chevallier, S., Wiart, J., & Bloch, I. (2017). Subject-specific time-frequency selection for multi-class motor imagery-based BCIs using few laplacian EEG channels. *Biomedical Signal Processing and Control*, 302-311.
- Zheng, X., & Chen, W. (2021). An Attention-based Bi-LSTM Method for Visual Object Classification via EEG. *Biomedical Signal Processing and Control*.