

INSTITUTO TECNOLÓGICO DE CHIHUAHUA
DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

***“DISEÑO DE ALGORITMOS PARA EXTRACCIÓN
DE CARACTERÍSTICAS DE SEÑAL P300 EN
INTERFACES CEREBRO COMPUTADORA”***

TESIS

QUE PARA OBTENER EL GRADO DE

***MAESTRO EN CIENCIAS EN INGENIERÍA
ELECTRÓNICA***

PRESENTA:

ING. LUIS MADRID HERRERA

DIRECTOR DE LA TESIS:

DR. JUAN ALBERTO RAMÍREZ QUINTANA



**TECNOLÓGICO NACIONAL DE
MÉXICO**



CHIHUAHUA, CHIH., ENERO 2018



Chihuahua, Chih., 10 de enero de 2018

MTR. LUIS CARDONA CHACÓN
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
PRESENTE

Por medio de la presente notificamos a usted que en cumplimiento de los requerimientos para la obtención del grado de **Maestro en Ciencias en Ingeniería Electrónica**, el documento de tesis del **C. LUIS MADRID HERRERA**, ha sido aprobado y aceptado para su impresión. El título de la tesis es:

"Diseño de algoritmos para extracción de características de señal P300 en interfaces cerebro computadora"

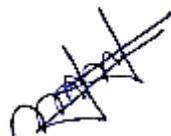
Por lo que proponemos, le sea concedida la autorización de impresión correspondiente.

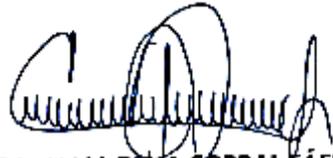
Agradeciendo la atención a la presente, quedamos de usted:

Atentamente,

"LA TÉCNICA POR EL ENGRANDECIMIENTO DE MÉXICO"

Juan Ramírez
DR. JUAN ALBERTO RAMÍREZ QUINTANA
DIRECTOR DE TESIS


DR. MARIO IGNACIO CHACÓN MURGUÍA
MIEMBRO DEL JURADO DE EXAMEN


MTRA. ALMA DELIA CORRAL SÁENZ
MIEMBRO DEL JURADO DE EXAMEN


DRA. MARÍA DEL PILAR GÓMEZ GIL
MIEMBRO DEL JURADO DE EXAMEN


**SECRETARÍA DE
EDUCACIÓN PÚBLICA
INSTITUTO TECNOLÓGICO
DE CHIHUAHUA**





Chihuahua, Chih., 10 de enero de 2018

**C. LUIS MADRID HERRERA
PRESENTE**

Por este conducto le comunico que a propuesta del Jurado de Examen, la División de Estudios de Posgrado e Investigación le ha concedido la autorización para la impresión de su tesis para obtener el grado de Maestro en Ciencias en Ingeniería Electrónica, cuyo título es:

"Diseño de algoritmos para extracción de características de señal P300 en interfaces cerebro computadora"

Con el siguiente contenido de capítulos:

- I. Introducción
- II. Metodología de obtención y generación de señales EEG con P300
- III. Extracción de características basada en la amplitud de la señal P300 y clasificación mediante una red feedforward
- IV. Extracción de características y clasificación de la señal P300 mediante una red neuronal convolucional
- V. Detección de la señal P300 para la selección de objetos
- VI. Comparación de los resultados
- VII. Resultados, conclusiones y trabajo futuro

Atentamente,

"LA TÉCNICA POR EL ENGRANDECIMIENTO DE MÉXICO"

**MTRO. LUIS CARDONA CHACÓN
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN**



SECRETARÍA DE
EDUCACIÓN PÚBLICA
INSTITUTO TECNOLÓGICO
DE CHIHUAHUA



Av. Tecnológico N° 2909, Col. 10 de Mayo C.P. 31310, Chihuahua, Chih., México
Tel. (614) 201-2000 Fax (614) 413-5187 Apartado Postal 2-1549
Correo Electrónico: director@itchihuahua.edu.mx www.itchihuahua.mx



Chihuahua, Chihuahua a 19 de enero de 2018

Dr. Enrique Cabrero Mendoza

Director de CONACYT.

Presente.

Por este conducto aprovecho la ocasión para saludarlo e informarle que a la fecha he obtenido el grado de **Maestría en Ciencias en Ingeniería Electrónica** en la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de Chihuahua. Motivo por el cual agradezco todo el apoyo brindado por esta Institución que usted representa, el otorgamiento de esta beca permitió dedicarme de tiempo completo a la realización de mis estudios de Posgrado y de esta manera lograr el cumplimiento del objetivo principal del convenio establecido.

Sin otro particular por el momento, me es grato quedar de usted como su seguro servidor, no sin antes reiterar mi agradecimiento. Muchas Gracias.

A t e n t a m e n t e

Luis Madrid Herrera
Exbecario CONACYT

Agradecimientos

Agradezco al Tecnológico Nacional de México por el apoyo a este trabajo bajo el proyecto “Adquisición, procesamiento y análisis de señales EEG dirigido a generar nuevo conocimiento y tecnología para interfaz cerebro-computadora orientada a rehabilitación virtual”, Clave 5684.16-P. Agradezco al CONACYT por el apoyo económico brindado durante mis estudios. Agradezco a la Secretaria de Educación Cultura y Deporte por la beca de créditos otorgada. Agradezco al personal de la DEPI por su apoyo. Agradezco a mi director de tesis el Dr. Juan A. Ramírez por su asesoramiento y dedicación. Agradezco a mi Comité Tutorial, Dr. Mario I. Chacón, M.C. Alma Corral y Dra. Pilar Gómez por su apoyo en este trabajo de tesis. Agradezco a mis compañeros y amigos de generación Abimael y Miguel por compartir sus conocimientos. Agradezco a mis compañeros de laboratorio, Chava, Carlos, Jorge, Brenda, Meny, Oscar, Macías, Raúl, Daniel, Andrea y Lalo, por el ambiente y convivencia. Agradezco a mis papás y hermanos por su apoyo incondicional. Agradezco a Vanessa por su apoyo brindado. Agradezco a Manuel por animarme a estudiar una maestría. También agradezco a todas aquellas personas que indirectamente ayudaron a realizar este trabajo de tesis.

RESUMEN

DISEÑO DE ALGORITMOS PARA EXTRACCIÓN DE CARACTERÍSTICAS DE SEÑAL P300 EN INTERFACES CEREBRO COMPUTADORA

Ing. Luis Madrid Herrera
Maestría en Ciencias en Ingeniería Electrónica
División de Estudios de Posgrado e Investigación del
Instituto Tecnológico de Chihuahua
Chihuahua, Chih. 2018
Director de Tesis: Dr. Juan Alberto Ramírez Quintana

La medición de la actividad cerebral ha cobrado interés en los últimos años debido a que ha permitido el desarrollo de novedosos esquemas de comunicación para personas con capacidades diferentes que no pueden hablar o moverse. Una forma exitosa de utilizar la medición de la actividad cerebral para comunicación, es el análisis de la señal P300 contenida en señales electroencefalográficas (EEG). La señal P300 es un potencial evocado relacionado a eventos (ERP por sus siglas en inglés) que se genera debido a un cambio en la actividad cerebral cuando las personas son expuestas a estímulos externos. Esta señal se manifiesta en las señales EEG como una deflexión positiva alrededor de los 300 ms después de que el sujeto se expone a un estímulo inesperado. La detección de señales P300 permite el diseño de sistemas basados en interfaces cerebro computadora (BCI por sus siglas en inglés), que pueden ser utilizados en aplicaciones relacionadas a comunicación tales como deletreo, selección de objetos, entre otras.

Existen diversos métodos en la literatura para detectar las señales P300 que obtienen altos desempeños basándose en metodologías como redes neuronales artificiales, reconocimiento de patrones o aprendizaje profundo. Sin embargo, estos métodos son difíciles de implementar en entornos reales ya que tienen una complejidad computacional alta por la gran cantidad de electrodos que utilizan. Por lo tanto, aún se requiere realizar trabajo para el desarrollo de nuevos métodos de detección de la señal P300, que puedan analizar las características de este tipo de señales obtenidas con pocos electrodos.

En consecuencia, para contribuir con el desarrollo de nuevos métodos de análisis de la señal P300, en este trabajo se proponen dos métodos basados en diferentes esquemas de extracción

de características. Un método se denomina FFp300 y su extracción de características está basada en la amplitud de la señal P300, que utiliza una red neuronal artificial *feedforward* para la clasificación. El otro es un modelo de aprendizaje profundo que se llama CNN300, se trata de red neuronal convolucional que define automáticamente sus características y realiza la clasificación. Ambos métodos se diseñaron utilizando los electrodos O1, Oz y O2 del sistema internacional 10-20 y para evocar la señal P300 se utilizó el delectreador de Donchin.

Los resultados de ambos métodos mostraron que CNN300 tiene mejor desempeño que FFp300, ya que este último puede llegar a fallar en de detección de la señal cuando la señal P300 se presenta desfasada. Debido a que CNN300 obtuvo mejor desempeño, este método se implementó con éxito en una terapia virtual para selección de objetos.

Para conocer el estado de funcionamiento de FFp300 y CNN300 respecto a los demás métodos de la literatura, se realizó un estudio comparativo, el cual consistió en medir el desempeño y las características de FFp300 y CNN300, contra los métodos de detección de P300 más populares. Los resultados arrojaron que FFp300 tiene un desempeño aceptable respecto a los demás métodos, ya que, aunque tiene una tasa de reconocimiento baja, este utiliza poca cantidad de electrodos. Por otro lado, CNN300 tiene un desempeño competitivo contra los mejores métodos debido a que su tasa de reconocimiento es alta, y utiliza una cantidad de electrodos significativamente baja respecto a los que utilizan los mejores métodos.

CONTENIDO

LISTA DE FIGURAS	<i>xi</i>
LISTA DE TABLAS	<i>xiv</i>
CAPÍTULO 1. INTRODUCCIÓN	1
1.1 Métodos de adquisición de señales cerebrales.....	1
1.2 Electroencefalografía	3
1.2.1 Clasificación de señales EEG con base en frecuencia	3
1.3 Interfaces cerebro computadora.....	5
1.3.1 Adquisición	6
1.3.2 Procesamiento de las señales EEG.....	8
1.3.3 Aplicaciones de sistemas BCI.....	9
1.4 Potenciales relacionados a eventos (ERP).....	10
1.5 Trabajos realizados relacionados a la detección de ERP	13
1.6 Justificación y objetivos de la tesis.....	15
CAPÍTULO 2. METODOLOGÍA DE OBTENCIÓN Y GENERACIÓN DE SEÑALES EEG CON P300	19
2.1 Base de datos de señales EEG	20
2.2 Dispositivo OpenBCI®	22
2.2.1 Ultracortex Mark IV.....	24
2.2.2 Electrodo de copa y gel conductor	25
2.2.3 Experimentos para la selección de electrodos.....	26
2.2.4 Conexión de OpenBCI® con Matlab®	29
2.3 Metodología para la adquisición de señales	30
2.4 Obtener señal P300 con delectreador de Donchin y el OpenBCI®	31
2.4.1 Diseño de la interfaz en Matlab® para el delectreador de Donchin	32
2.4.2 Experimentación para la asegurar la generación de la señal P300.....	35
2.4.3 Análisis de P300 en el delectreador de Donchin original	39
2.4.4 Paradigma de Donchin modificado.....	42
2.4.5 Análisis de P300 en el delectreador de Donchin modificado	42

2.4.6 Conclusiones de los experimentos de generación de P300.....	43
CAPÍTULO 3. EXTRACCIÓN DE CARACTERÍSTICAS BASADA EN LA AMPLITUD DE LA SEÑAL P300 Y CLASIFICACIÓN MEDIANTE UNA RED <i>FEEDFORWARD</i>	45
3.1 Esquema general del método propuesto	45
3.2 Preprocesamiento y extracción de características.....	46
3.2.1 Agrupamiento de señales por intensificaciones	46
3.2.2 Características con base en la amplitud	49
3.3 Clasificación	51
3.3.1 Clasificación para la detección de la señal P300	52
3.3.2 Predicción del carácter deseado	56
3.4 Resultados y discusión.....	57
3.4.1 Análisis de resultados en detección por filas ^T y columnas	58
3.4.2 Análisis de resultados en la predicción del carácter	66
CAPÍTULO 4. EXTRACCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE LA SEÑAL P300 MEDIANTE UNA RED NEURONAL CONVOLUCIONAL.....	69
4.1 Esquema general del método propuesto	69
4.2 Preprocesamiento de las señales EEG	70
4.3 Agrupamiento de las señales por intensificación.....	70
4.4 Red neuronal convolucional	72
4.4.1 Etapas de convolución y rectificación.....	73
4.4.2 Capa completamente conectada	75
4.4.3 Capa exponencial normalizada	76
4.4.4 Capa de clasificación.....	76
4.4.5 Entrenamiento de la red	77
4.5 Predicción del carácter deseado.....	78
4.6 Resultados y discusión.....	79
4.6.1 Resultados por filas ^T y columnas.....	79
4.6.2 Resultados para la predicción del carácter	82
CAPÍTULO 5. DETECCIÓN DE LA SEÑAL P300 PARA LA SELECCIÓN DE OBJETOS	86

5.1 Terapia virtual y P300.....	86
5.1.1 Escenario	87
5.1.2 Generación de intensificaciones.....	88
5.2 Detección de la señal P300 mediante CNN300	91
5.3 Resultados y discusión.....	92
5.3.1 Resultados para la detección de la señal P300 en cada objeto	93
5.3.2 Resultados para la detección del objeto deseado	98
CAPÍTULO 6. COMPARACIÓN DE LOS RESULTADOS	100
6.1 Estado del arte de los delectreadores basados en P300	100
6.1.1 Métodos basados en máquinas de vectores de soporte	100
6.1.2 Análisis de componentes independientes para la detección de P300.....	101
6.1.3 Redes neuronales convolucionales para detección de P300	102
6.1.4 Técnicas de promediado para el análisis de ERP	103
6.2 Análisis y comparación de los métodos.....	104
6.2.1 Pesos de cada criterio de evaluación	107
6.2.2 Cálculo de puntuación para cada criterio de evaluación	109
CAPÍTULO 7. RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO	113
REFERENCIAS	116

LISTA DE FIGURAS

Figura 1.1 Métodos de adquisición de señales cerebrales [1].	2
Figura 1.2 Espectro de frecuencia de una señal EEG [2].	3
Figura 1.3 Ejemplos de las señales clasificadas por banda de frecuencias [3].	5
Figura 1.4 Estructura general de un sistema BCI.	5
Figura 1.5 Sistema internacional 10-20, a) Vista lateral, b) Vista superior, c) Electrodo adicionales [2].	7
Figura 1.6 Campos de aplicación de sistemas BCI [1].	9
Figura 2.1 Interfaz del deletreador de Donchin [25].	19
Figura 2.2 Electrodo utilizado en la base de datos II de BCI Competition III [25].	21
Figura 2.3 OpenBCI [®] Cyton [10].	23
Figura 2.4 Dispositivo USB para la conexión bluetooth con OpenBCI [®] [10].	23
Figura 2.5 Casco Ultracortex Mark IV [10].	24
Figura 2.6 Electrodo que pueden ser utilizados mediante el casco Ultracortex Mark IV.	24
Figura 2.7 a) electrodo para zonas con cabello, b) electrodo para zonas sin cabello [10].	25
Figura 2.8 Electrodo de copa para el dispositivo OpenBCI [®] y gel conductor [10].	26
Figura 2.9 Sujeto de pruebas con el dispositivo Ultracortex Mark IV.	26
Figura 2.10 Banda elástica utilizada para el soporte de los electrodo de copa.	27
Figura 2.11 Electrodo de copa con el gel conductor.	27
Figura 2.12 Sujeto de pruebas con los electrodo de copa colocados mediante la banda elástica.	28
Figura 2.13 Flujo de datos del OpenBCI [®] en Matlab [®] .	29
Figura 2.14 Señales en banda de frecuencia <i>Alpha</i> en el canal Oz cuando un sujeto de pruebas cierra los ojos.	30
Figura 2.15 Encuesta realizada a los sujetos de pruebas antes de cada sesión.	31
Figura 2.16 Forma aproximada de la señal P300.	32
Figura 2.17 Interfaz del deletreador para la toma de muestras.	33
Figura 2.18 Periodo de intensificaciones en una época de deletreo.	34
Figura 2.19 Composición de la señal $s_j(n)$ de una época de deletreo con OpenBCI [®] a 250 Hz.	36

Figura 2.20 Respuesta a la frecuencia del filtro representada en dB.....	37
Figura 2.21 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) en los canales C3, Cz, C4 y P3.	39
Figura 2.22 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) en los canales Pz, P4, O1 y O2.	40
Figura 2.23 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) por filas y columnas en los canales O1 y O2.....	41
Figura 2.24 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) por filas ^T y columnas en los canales O1 y O2, mediante el deletreador de Donchin modificado.....	43
Figura 3.1 Esquema del método FFp300 basado en la extracción de características con base en la amplitud de la señal P300.....	46
Figura 3.2 Promedio de vectores $g_{Tj}(w)$ (P300) y $g_{Fj}(w)$ (No P300), para un sujeto de pruebas en los canales O1, Oz y O2.	48
Figura 3.3 Selección de características con base en la amplitud de la señal cerca de los 300 ms.	49
Figura 3.4 Índices k para columnas y filas ^T de la matriz del deletreador.....	50
Figura 3.5 Vectores de características en los canales O1, Oz y O2, para uno de los sujetos de pruebas.....	51
Figura 3.6 Arquitectura del método FFp300.	52
Figura 3.7 Topología de la red neuronal artificial utilizada para la clasificación de las características en un canal.	53
Figura 3.8 Topología de la red neuronal artificial utilizada para la combinación de la información que generan las redes neuronales por canal.	55
Figura 3.9 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 8.....	63
Figura 3.10 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 6.....	63
Figura 3.11 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 10.....	65
Figura 4.1 Esquema del método para la detección de la señal P300 basado en una CNN.....	70
Figura 4.2 a) Acondicionamiento de las señales EEG a formato de imagen, b) Representación gráfica de $I_k(j,m)$	71

Figura 4.3 ejemplo de una CNN utilizada para extracción de características y clasificación [42].	72
Figura 4.4 Esquema general de la red neuronal convolucional utilizada por CNN300.	73
Figura 4.5 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), a) sujeto 2, b) sujeto 4, c) sujeto 5, d) sujeto 6, e) sujeto 8.	83
Figura 4.6 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 7.	84
Figura 5.1 Escenario de frutas utilizado para la selección de objetos.	88
Figura 5.2 Escenario de frutas con la manzana seleccionada.	88
Figura 5.3 Escenario de frutas con uno de los objetos intensificados.	89

LISTA DE TABLAS

Tabla 3.1 Resultados de las matrices de confusión para FFp300 en entrenamiento y validación.	59
Tabla 3.2 Resultados de las métricas de evaluación.....	61
Tabla 3.3 Caracteres deletreados por cada uno de los sujetos y porcentaje de aciertos.	66
Tabla 4.1 Resultados de las matrices de confusión mediante CNN300 en entrenamiento y validación.	80
Tabla 4.2 Resultados con base en las métricas de evaluación para CNN300.	81
Tabla 4.3 Caracteres deletreados por cada uno de los sujetos y porcentaje de aciertos.	82
Tabla 5.1 Resultados de las matrices de confusión tomando en cuenta 15 intensificaciones. ..	93
Tabla 5.2 Resultados de las matrices de confusión tomando en cuenta 10 intensificaciones. ..	94
Tabla 5.3 Resultados de las matrices de confusión tomando en cuenta 5 intensificaciones.	94
Tabla 5.4 Resultados con base en las métricas de evaluación para 15 intensificaciones.	95
Tabla 5.5 Resultados con base en las métricas de evaluación para 10 intensificaciones.	96
Tabla 5.6 Resultados con base en las métricas de evaluación para 5 intensificaciones.	96
Tabla 5.7 Porcentaje de épocas de selección de objetos reconocidas correctamente.....	98
Tabla 6.1 Desempeño y características de los métodos propuestos en la literatura y este trabajo de tesis. El ranking solo toma en cuenta el porcentaje de reconocimiento.	105
Tabla 6.2 Puntuación máxima de los criterios para la evaluación.....	108
Tabla 6.3 Puntuación para cada uno de los criterios de evaluación por método.....	111

CAPÍTULO 1. INTRODUCCIÓN

La comunicación es una habilidad que permite el intercambio de información con el fin de transmitir mensajes. Esta habilidad es fundamental en las actividades diarias que realizan los seres humanos para ser independientes. Sin embargo, existen personas con discapacidad motora y del habla que tienen impedimentos en las habilidades de comunicación. Para contribuir en solucionar esta problemática, se han presentado diversas soluciones, entre las que destaca el desarrollo de tecnología para medir e interpretar la actividad cerebral. Entre las formas de realizar esta medición con propósitos de comunicación, se encuentran los potenciales relacionados a eventos (ERP por sus siglas en inglés), los cuales son evocados por estímulos externos aleatorios. Un ERP popular es la señal P300, la cual se puede obtener mediante señales EEG y se puede evocar mediante interfaces visuales.

En este trabajo de tesis se recurre al análisis de la señal P300 para desarrollar dos metodologías novedosas, que permitan codificar mensajes de personas mediante deletreadores y terapia virtual. El objetivo es contribuir en generar nuevas estrategias de comunicación para personas con discapacidad motora y del habla.

En este primer capítulo se presentan los conceptos elementales relacionados al trabajo realizado y se compone de la siguiente manera: las secciones 1.1 y 1.2 describen los métodos de adquisición de señales cerebrales y la electroencefalografía. La sección 1.3 detalla los elementos de una interfaz cerebro computadora que obtiene EEG. La sección 1.4 presenta los principales conceptos sobre potenciales relacionados a eventos y P300. La sección 1.5 presenta el estado del arte de los métodos basados en potenciales evocados. Finalmente se abordan la justificación y objetivos de la tesis en la sección 1.6.

1.1 Métodos de adquisición de señales cerebrales

La medición de las señales eléctricas generadas por la actividad de las neuronas ha sido de gran interés en los últimos años, ya que ha propiciado un gran avance tecnológico en áreas como: rehabilitación, comunicación, diagnóstico médico, farmacología, psicología, psiquiatría, seguridad, domótica, entre otros [1]. Las mediciones de la actividad cerebral se realizan con

diversos métodos de adquisición que se pueden clasificar en invasivos o no invasivos, tal como se muestra en la figura 1.1.

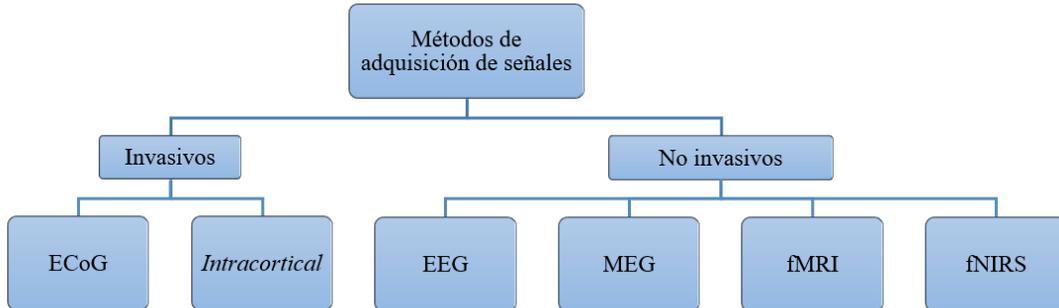


Figura 1.1 Métodos de adquisición de señales cerebrales [1].

Los métodos invasivos son aquellos en los que se requiere de una intervención quirúrgica para colocar los sensores y medios de adquisición. Los más comunes son:

- Electrocorticografía (ECoG): las señales eléctricas del cerebro se miden mediante electrodos que se colocan sobre la superficie del cerebro.
- *Intracortical*: consiste en la colocación de los electrodos en capas corticales que se encuentran dentro del cerebro.

Los métodos no invasivos son aquellos en los cuales no se requiere de intervención quirúrgica para la colocación de electrodos [1]. Los métodos más populares son:

- Electroencefalografía (EEG): se refiere al registro de señales con electrodos que se colocan en el cuero cabelludo.
- Magnetoencefalografía (MEG): se basa en medir los campos magnéticos producidos por las señales eléctricas del cerebro.
- Resonancia magnética funcional (fMRI): este método permite obtener imágenes que representan la actividad cerebral mediante la detección del cambio de flujo en la sangre del cerebro.
- Espectroscopia cercana infrarroja funcional (fNIRS): este método también mide el movimiento de la sangre en el cerebro, pero con el uso de luz infrarroja.

1.2 Electroencefalografía

La electroencefalografía es un método que se basa en la colocación de electrodos sobre el cuero cabelludo para medir y registrar la actividad eléctrica del cerebro. Permite conocer la actividad eléctrica de las neuronas cuando se presenta un estado mental o el cerebro reacciona a ciertos estímulos.

Las ventajas de la EEG respecto a los demás métodos de adquisición de señales cerebrales es que son de bajo costo y fácil implementación, ya que para colocar los electrodos no se requiere una intervención quirúrgica. De esta manera, la tecnología basada en EEG es muy flexible y se puede aplicar en diversos usos comerciales, lo cual es poco común con los demás métodos de adquisición de actividad cerebral. No obstante, la medición basada en EEG tiene mayores niveles de ruido que el resto de los métodos debido a que las señales no solo tienen información derivada de la actividad cerebral, sino que pueden contener ruido derivado de otras actividades del cuerpo humano, conocido como artefactos.

1.2.1 Clasificación de señales EEG con base en frecuencia

Las señales EEG se registran generalmente con una amplitud entre 0.5 a 100 μV , y los electrodos utilizados tienen una alta impedancia que ayuda a comprobar que tengan contacto con el cuero cabelludo. El rango de frecuencia de las señales EEG es de 1 a 50 Hz [2], pero las componentes de estas señales se concentran en bandas muy específicas, como lo muestra la figura 1.2, donde la energía se concentra más en el rango de 0 a 13 Hz.

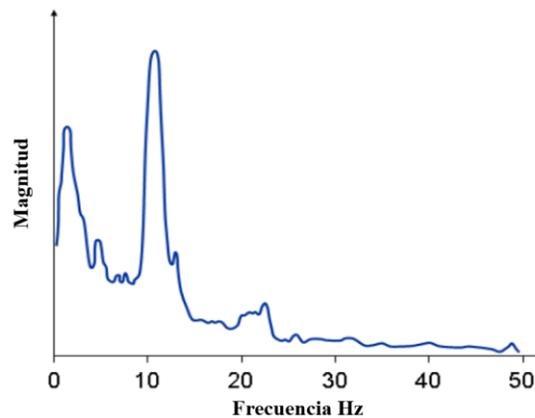


Figura 1.2 Espectro de frecuencia de una señal EEG [2].

Una forma común de clasificar las señales EEG es con base en el ancho de banda donde se concentra la mayor parte de los componentes de frecuencia, los cuales se pueden obtener mediante la transformada discreta de Fourier [3]. A continuación, se mencionan algunas de las bandas de frecuencia básicas:

- $0.5 \text{ Hz} < \textit{Delta} < 4 \text{ Hz}$
- $4 \text{ Hz} < \textit{Theta} < 8 \text{ Hz}$
- $8 \text{ Hz} < \textit{Alpha} < 13 \text{ Hz}$
- $13 \text{ Hz} < \textit{Beta}$

La banda *Delta* presenta gran actividad en personas adultas cuando están en un estado de sueño profundo. También se presenta en bebés, pero van disminuyendo su magnitud mientras incrementa la edad. La banda *Theta* se asocia a estados de meditación, cálculo mental y concentración. La banda *Alpha* se relaciona al procesamiento visual generado en la parte occipital del cerebro y su amplitud se incrementa cuando las personas cierran los ojos. La banda *Beta* se asocia a las actividades motoras y se presentan en la parte frontal y central del cerebro [4].

Existen dos bandas de frecuencia adicionales que tienen rangos similares a las bandas *Alpha* y *Beta*, las cuales son:

- $8 \text{ Hz} < \textit{Mu} < 13 \text{ Hz}$
- $30 \text{ Hz} < \textit{Gamma}$

La banda *Mu* tiene el mismo rango que las *Alpha*, pero las producen las neuronas de la corteza sensoriomotora en actividades de movimiento. La banda *Gamma* se presenta en personas adultas saludables con ciertas funciones motoras o perceptuales [4]. En la figura 1.3 se muestran ejemplos de la forma de onda de señales EEG de distintas bandas de frecuencia.

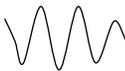
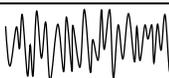
<i>Delta</i>	
<i>Theta</i>	
<i>Alpha y Mu</i>	
<i>Beta</i>	
<i>Gamma</i>	

Figura 1.3 Ejemplos de las señales clasificadas por banda de frecuencias [3].

1.3 Interfaces cerebro computadora

Debido a las ventajas que presentan las señales EEG, han sido utilizadas exitosamente para medir la actividad del sistema nervioso y el cerebro cuando se realizan actividades como pensar, observar, moverse, comunicarse, entre otras [5]. La tecnología que genera el medio de adquisición de estas señales y las envía a un sistema de cómputo, se denomina interfaz cerebro computadora (BCI por sus siglas en inglés). A los elementos de procesamiento que interpretan las señales que obtiene una BCI para una aplicación, se les denomina sistemas de interfaz cerebro computadora (sistemas BCI). Generalmente, estos sistemas se componen de cinco etapas: adquisición de señales, preprocesamiento, extracción de características, clasificación y finalmente la etapa de aplicación. En la figura 1.4 se observa el esquema general de las etapas en un sistema BCI y en las siguientes subsecciones se detallan cada una de las etapas.

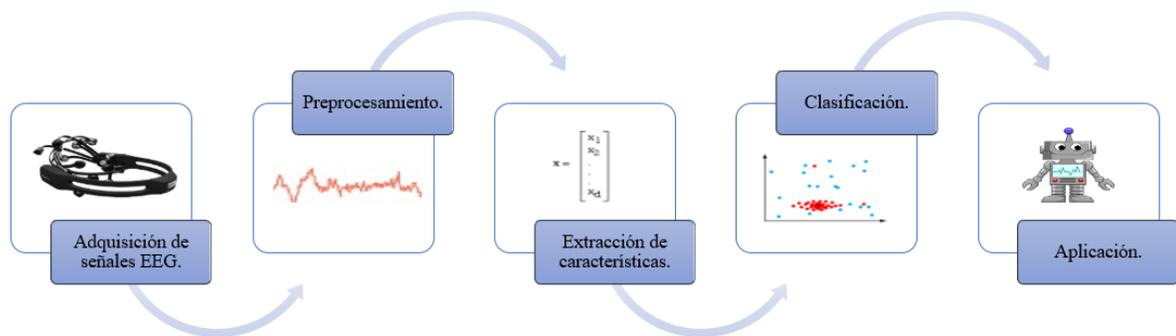


Figura 1.4 Estructura general de un sistema BCI.

1.3.1 Adquisición

En esta etapa se utilizan electrodos como sensores para medir la actividad cerebral [6]. En el caso de los sistemas BCI basados en EEG, se utiliza el sistema internacional 10-20 como referencia para la colocación de electrodos. Este sistema es un estándar que especifica 21 localizaciones básicas de electrodos sobre el cuero cabelludo en el perímetro de la cabeza por intervalos de 10% y 20% (figura 1.5). Para generar el perímetro de la cabeza se tienen dos puntos de referencia los cuales son *Nasion* e *Inion*, posicionados en la parte superior de la nariz y en la parte central de la nuca respectivamente. Los demás electrodos son posicionados de forma equidistante formando una cuadrícula con los puntos de referencia, tal como se ilustra en la figura 1.5a y 1.5b. Cada una de las localizaciones de los electrodos se nombra con dos caracteres. El primero se refiere a la zona del cráneo donde se encuentran ya sea la frontal (F), central (C), temporal (T), parietal (P) u occipital (O). El segundo carácter es un subíndice que depende de lo retirado que se encuentre el electrodo de la zona media del cráneo. Para las localizaciones en la sección media del cráneo, se utiliza el subíndice z, y para las localizaciones que se encuentran del lado izquierdo del cráneo los subíndices son números nones y del lado derecho números pares. Estos números incrementan conforme se van alejando de la parte media como se observa en las figuras 1.5a y 1.5b. Los lóbulos de las orejas son utilizados normalmente como referencias, los cuales se nombran A_1 para el lóbulo izquierdo y A_2 para el lóbulo derecho.

El sistema internacional 10-20 tiene una segunda configuración de electrodos extras, como se ve en la figura 1.5c. En esta configuración, se definen localizaciones intermedias entre las zonas del cráneo mencionadas anteriormente. Estas localizaciones se nombran de la siguiente forma: parietal occipital (PO), central parietal (CP), temporal parietal (TP) y frontal central (FC). La zona frontal es más extensa que las demás, por lo cual las localizaciones que se agregan en esta zona son nombrados AF [2]. El subíndice que los acompaña indica qué tan alejado está el electrodo de la zona media del cráneo, lo cual se puede ver en la figura 1.5c.

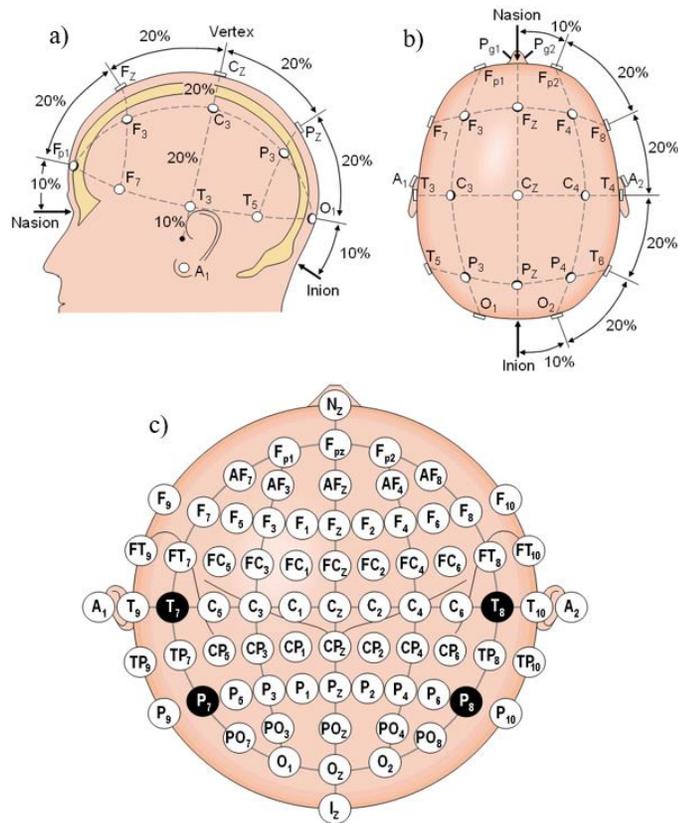


Figura 1.5 Sistema internacional 10-20, a) Vista lateral, b) Vista superior, c) Electrodo adicionales [2].

El sistema 10-20 ha servido de base para la configuración de electrodos en dispositivos de adquisición portátiles y no portátiles que ofrecen diversas compañías. Entre las aplicaciones de estos dispositivos se encuentran videojuegos, domótica, investigación, entre otras [7]–[10]. Algunos de estos dispositivos son mencionados a continuación:

- La compañía EMOTIV[®] se dedica a la fabricación de diademas con sensores capaces de medir la actividad cerebral por medio de la electroencefalografía. Tiene dos productos principales los cuales son la EMOTIV[®] Epoc+ y la EMOTIV[®] Insight. La Epoc+ es una diadema con 14 electrodos y cuenta con comunicación bluetooth. La EMOTIV[®] Insight cuenta con 5 electrodos y bluetooth. Estas diademas son comúnmente utilizadas en videojuegos o control de ciertos dispositivos electrónicos. Además, en la literatura se han utilizado para obtener potenciales evocados [7].
- Neurosky[®] es una compañía dedicada a la fabricación de sensores para señales EEG y cuenta con la diadema Mindwave. Esta diadema tiene el electrodo FPz y utiliza

comunicación bluetooth. Es capaz de medir estados mentales como el estrés, la meditación, la atención, entre otras [8].

- La compañía g.tec[®] se dedica a la ingeniería médica y tiene diferentes productos para la adquisición de señales EEG, dentro de los cuales el más conocido es el g.Nautilus. Este tipo de sistema de alta gama soporta más de 64 canales y tiene versiones alámbricas e inalámbricas [9].
- OpenBCI[®] se dedica a la creación de interfaces cerebro computadora con código abierto, sus productos cuentan con un software de desarrollo que permite crear diferentes algoritmos para la interpretación de las señales EEG [10].

1.3.2 Procesamiento de las señales EEG

Las señales EEG adquiridas se acondicionan mediante métodos basados en procesamiento digital de señales y se interpretan con modelos basados en reconocimiento de patrones [11]. El procesamiento de señales permite que la señal sea transformada de diversas formas para que puedan interpretarse más fácilmente. El reconocimiento de patrones es una ciencia que tiene como objetivo la clasificación de objetos en un número determinado de categorías o clases. Normalmente, se utiliza para la clasificación de señales de una o dos dimensiones [12]. Los sistemas BCI se diseñan como sistemas de reconocimiento de patrones y cuentan con las siguientes etapas: adquisición de señales, preprocesamiento, extracción de características y clasificación. Además, requieren de un conjunto de muestras para su entrenamiento y validación [13].

1.3.2.1 Preprocesamiento de señales

En esta etapa las señales son procesadas para reducir el ruido y obtener datos consistentes para análisis posteriores. Comúnmente se utilizan técnicas como filtrado, suavizado y normalización [14]. Existen diferentes transformadas que facilitan el procesamiento de las señales, algunas de las más populares son la transformada rápida de Fourier y la transformada wavelet.

1.3.2.2 Extracción de características

En esta parte se extrae la información relevante de las señales. Es decir, se transforma la entrada de datos en un vector de características para reducir el tamaño de los datos obteniendo información no redundante que sea lo más descriptiva posible. Existen muchos métodos para facilitar la extracción de características como lo son: análisis de componentes principales, patrones espaciales comunes, algoritmos genéticos, entre otros [4].

1.3.2.3 Clasificación

En esta etapa se utilizan las características extraídas para conocer a qué clase pertenece cada uno de los patrones de entrada. Existen clasificadores supervisados y no supervisados. Los supervisados ajustan sus parámetros mediante un conjunto de datos de entrenamiento. Dichos datos consisten en patrones de entrada y la clase a la que pertenecen cada uno de ellos. Los clasificadores no supervisados forman conjuntos basados en la similitud de los patrones de entrada. Algunos de los métodos para clasificadores son las redes neuronales artificiales, sistemas difusos, máquinas de vectores de soporte, entre otros [13].

1.3.3 Aplicaciones de sistemas BCI

Los sistemas BCI pueden ser utilizados en personas sanas o que presentan algún tipo de discapacidad ya que se basan en la interpretación de las señales eléctricas del cerebro y no requieren de la manipulación de dispositivos como un teclado, ratón, entre otros. Como se muestra en la figura 1.6, las aplicaciones más comunes de los sistemas BCI son entretenimiento, medicina, seguridad, educación, entornos inteligentes, etc.

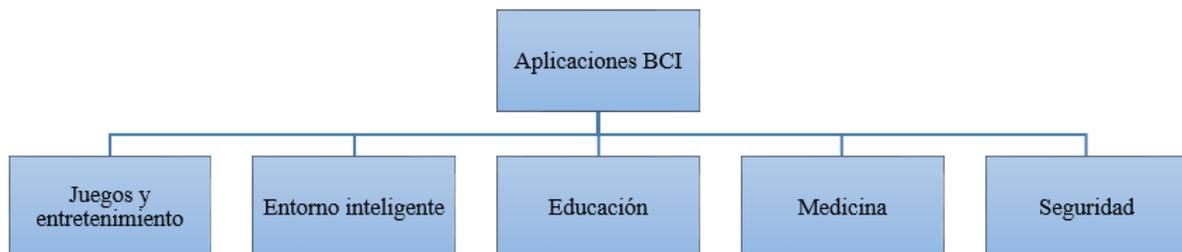


Figura 1.6 Campos de aplicación de sistemas BCI [1].

- Videojuegos: son juegos electrónicos donde las personas interactúan con las computadoras. Su funcionamiento es por medio de una retroalimentación visual y auditiva. Los videojuegos tienen fines de entretenimiento y de simulación de situaciones reales [5].
- Entorno inteligente: existen diversas aplicaciones de sistemas BCI para crear un entorno inteligente. Una de las más comunes es la domótica, que consiste en generar comandos mentales para controlar los dispositivos eléctricos y electrónicos que hay en el hogar.
- Aplicaciones médicas: se han encontrado diversos campos de la medicina en los cuales es posible aplicar un sistema BCI. Algunas de las aplicaciones son control de prótesis robóticas, manejo de sillas de ruedas, detección de anomalías en la actividad cerebral, rehabilitación mediante terapias virtuales, entre otras.
- Educación: los sistemas BCI se pueden utilizar en este campo para realizar aplicaciones que ayuden a las personas a aprender diversas cosas, un ejemplo podría ser el relacionar objetos con palabras. Además, se puede medir la concentración de los alumnos al momento de estar estudiando con el propósito de modificar las técnicas de estudio o del ambiente que los rodea.
- Seguridad: se refiere a la implementación de sistemas BCI para dar acceso a ciertos lugares o software restringidos por medio de las señales que se generan en el cerebro, por ejemplo, mediante algún pensamiento en específico.

1.4 Potenciales relacionados a eventos (ERP)

Los sistemas BCI basados en EEG utilizan diferentes tipos de estrategias mentales para generar las señales eléctricas que pueden describir fenómenos en la actividad cerebral [15]. La selección de la estrategia mental es determinante para la aplicación del sistema BCI que se va a desarrollar. Estas estrategias se pueden clasificar en potenciales evocados (EP por sus siglas en inglés) y eventos relacionados a sincronización y desincronización (ERD/ERS por sus siglas en inglés). Los potenciales evocados se generan mediante estímulos externos y se pueden utilizar en diversas aplicaciones relacionadas a comunicación entre personas, o de una persona a un dispositivo. Los eventos relacionados a la sincronización y desincronización son producidos

mediante tareas mentales, imaginación motora o calculo mental sin necesidad de un estímulo externo, y su detección puede ser utilizada para ejecutar comandos en un dispositivo electrónico.

El enfoque de este trabajo se centrará en los potenciales evocados, para desarrollar una aplicación de comunicación de una persona con un sistema de terapia virtual. Dentro de los potenciales evocados se encuentran; los potenciales evocados de estado estable (SSEP) y los potenciales relacionados a eventos (ERP). Los SSEP son evocados por un estímulo modulado a una frecuencia fija y se presentan en la señal EEG como un incremento en la magnitud de dicha frecuencia. Los ERP son cambios en la actividad cerebral que se miden el dominio del tiempo cuando se es expuesto a un evento aleatorio. Estos eventos pueden ser estímulos sensoriales, cognitivos o una respuesta motora [1]. Los potenciales relacionados a eventos pueden dividirse en dos categorías; exógenos y endógenos. El término exógeno se refiere a las ondas que se producen dentro de los primeros 100 ms de iniciado el estímulo, también llamadas ondas sensoriales. Por otra parte el término endógeno se refiere a las ondas que se producen después de los 100 ms de provocado el estímulo, estas ondas reflejan cómo el cerebro evalúa el estímulo [15].

Al codificarse en el tiempo, los ERP presentan más ruido que los SSEP pero presentan varias ventajas sobre los SSEP, entre las cuales destacan las siguientes:

- Permiten medir la respuesta endógena derivada de un estímulo externo.
- El estímulo se puede presentar en cualquier momento (no tiene que ser periódico como en SSEP) [15].

Por lo tanto, el uso de señales ERP es adecuado para personas con distintos tipos de enfermedades que propician discapacidades en comunicación y problemas motores. A continuación, se presenta una descripción de las etapas de los estímulos ERP.

Las ondas producidas por los eventos se pueden describir con base en su tiempo de ocurrencia y amplitud. Para la nomenclatura de las ondas se utilizará el prefijo P para las ondas positivas y N para las ondas negativas, seguidos del número en milisegundos relacionado con el tiempo en el que las ondas se producen después del estímulo. Por ejemplo la señal P300 es una

onda con amplitud positiva que ocurre a los 300 ms después del estímulo [15]. A continuación, se da una breve explicación de las ondas:

- P50: es una deflexión positiva que se da entre los 40-75 ms después del estímulo. La amplitud de la onda es la diferencia absoluta entre el valor más bajo y más alto de la señal en este rango de tiempo. Esta onda indica que existe un evento nuevo. En caso de ser un evento repetitivo éste no genera el estímulo para evitar la sobrecarga de información.
- N100: es una deflexión negativa que aparece cuando un estímulo inesperado se presenta y se encuentra entre los 90-200 ms después del estímulo. La N100 tiene una función de proceso de búsqueda y es cuando el cerebro trata de identificar qué evento ocurrió.
- P200: es una deflexión positiva que se encuentra entre los 100-250 ms después del estímulo. De acuerdo con la literatura, esta deflexión se asocia a cuando el cerebro tiene la sensación de que el evento fue reconocido.
- N200: es una deflexión negativa que se da alrededor de los 200 ms. Existen 3 componentes para esta onda como se muestra a continuación:
 - N2a: es un componente negativo provocado por un cambio en algún estímulo auditivo y representa el proceso automático del cerebro para reconocer el estímulo.
 - N2b: este se produce después de que la N2a aparece cuando se da un cambio físico en alguna tarea importante.
 - N2c: clasifica a la N200 y aparece cuando se requiere una clasificación de estímulos dispares.
- N300: esta onda se presenta cuando el cerebro asimila el estímulo.
- P300: esta deflexión positiva se da alrededor de los 250-400 ms después del estímulo. El tiempo de ocurrencia de esta señal depende del tipo de estímulo, ya que el cerebro no los reconoce a todos a la misma velocidad. Se dice que mientras sea menor el tiempo de ocurrencia de estas ondas, mayor es la capacidad mental de las personas. La amplitud en estas ondas se relaciona con la atención al evento.

- N400: esta onda negativa se produce entre los 300-600 ms después del estímulo. Está inversamente asociada con la esperanza de relacionar el evento a alguna palabra o término de una frase.
- P600: esta onda se produce cuando hay algún procesamiento de lenguaje y se presenta cuando se cumple cualquiera de las siguientes 3 condiciones:
 - Se encuentra un error de sintaxis.
 - Presencia de alguna sintaxis indeseada.
 - Tener una sintaxis compleja.

1.5 Trabajos realizados relacionados a la detección de ERP

Aunque las señales que contienen los ERP presentan varias ventajas sobre otros potenciales evocados, tienen una relación señal ruido baja. En consecuencia, diversos trabajos de investigación se han enfocado en realizar algoritmos que extraigan características y clasifiquen estas señales. El ERP más utilizado es la señal P300, la cual, como se mencionó anteriormente es una deflexión positiva que se encuentra alrededor de los 300 ms después del estímulo y se localiza mediante una serie de estímulos repetitivos debido a que su detección es complicada. A continuación, se muestran algunos trabajos realizados enfocados a la extracción de características y clasificación de las señales P300.

A. Rakotomamonjy y V. Guigue proponen en [16] un algoritmo para el deletreador de Donchin basado en la señal P300, este deletreador genera estímulos visuales mediante la intensificación de las filas y columnas en una matriz alfanumérica. Se utiliza la base de datos II de BCI Competition III que cuenta con señales EEG adquiridas por medio de un dispositivo g.TEC[®]. Se utilizaron 64 canales y las señales fueron procesadas con un filtro pasa bandas Chevyshev de orden ocho con frecuencias de 0.1 a 10 Hz. Para la extracción de características se toma una ventana de 667 ms después de cada estímulo visual y se submuestra para obtener solo 14 muestras. Estos datos son clasificados mediante un ensamble de 17 máquinas de vectores de soporte. Los resultados obtenidos fueron entre el 73.5% y 96.5% en la detección de la señal P300 para los dos sujetos de pruebas.

En [17] H. Cecotti y A. Gr proponen un método para la detección y clasificación de la señal P300 en el deletreador de Donchin, en el cual utilizan también la base de datos II de BCI Competition III. Se menciona que en la detección de las señales P300 se tienen dos problemas principales: la detección de esta señal normalmente tiene la misma amplitud que el ruido, y la combinación de las señales que se generan al estimular a los caracteres vecinos de la matriz. El método para extracción de características y clasificación se basa en una Red Neuronal Convolutiva (CNN por sus siglas en inglés) y los resultados obtenidos fueron en promedio alrededor del 95.5% del reconocimiento utilizando 64 canales del sistema internacional 10-20.

D. J. Krusienski et al. Proponen en [18] un estudio para analizar los efectos del clasificador de discriminante lineal paso a paso (SWLDA por sus siglas en inglés) con el deletreador de Donchin. En este trabajo se utilizan las señales EEG obtenidas mediante un gorro de 64 canales (*Electro-Cap International Inc.*) utilizando el software BCI2000® [19]. La extracción de características se basaba en analizar las señales promediando la respuesta de cada carácter y el vector de características fue formado separando las señales en bloques y calculando su media. Se obtuvieron resultados de 90% de reconocimiento en tres de las cinco personas a las que se les realizó el experimento.

N. Xu, X. Gao et al. en [20] trabajaron con un algoritmo de análisis de componentes independientes (ICA por sus siglas en inglés) para la detección del potencial relacionado a eventos P300. Sus pruebas se realizan con la base de datos IIB de BCI Competition 2003 [21]. Para la extracción de características se utiliza un análisis de componentes principales (PCA por sus siglas en inglés), donde se obtiene la covarianza de las señales como características. Los resultados obtenidos con este método fueron del 89.5% en reconocimiento de caracteres.

A. Arasteh et al. Utilizan en [22] la señal P300 para la detección de inocencia o culpabilidad. El método para la extracción de características propuesto es el modo de descomposición empírico (EMD por sus siglas en inglés). Este método descompone las señales en componentes mediante una función de modo intrínseco (IMF por sus siglas en inglés), el cual se basa en que cada IMF obtiene el número de veces que la señal cruza por cero y se encuentran simétricamente localizadas respecto al cero. Gracias a este método de extracción de características se obtuvieron

resultados altos en el clasificador, con una precisión del 92.73%, mejor que otros métodos antes utilizados.

En [23] A. F. Perez Vidal et al., obtienen las señales EEG para extraer la señal P300 por medio del dispositivo EMOTIV[®] Epoc+. Para la implementación del proyecto se recurrió al software TestBench de EMOTIV[®], OpenVibe[®] y MagicCard. TestBench es utilizado para la adquisición de señales crudas, posteriormente estas señales son transmitidas al software OpenVibe[®], el cual contiene una herramienta llamada MagicCard para la selección de objetos por medio la señal P300. La etapa del reconocimiento de las señales está basada en un algoritmo de análisis discriminante lineal. El objetivo principal es que un sujeto controle un robot con 5 grados de libertad utilizando dos servomotores por medio de estímulos visuales. Para ello, el sujeto se concentrará en una imagen particular de una matriz de imágenes mostradas en una pantalla de computadora. Los resultados que se obtuvieron fueron de 100% de reconocimiento en un ambiente controlado en el cual, no existía ruido visual ni auditivo. En ambientes con ruido auditivo el resultado fue de 90% de reconocimiento. No obstante, el resultado ante ruido visual fue de 50%.

En [24] H. Wang, et al., proponen un método para la extracción de características en señales EEG enfocado al reconocimiento de los potenciales evocados visuales de estado estable (SSVEP), donde se utiliza el método de análisis de correlación canónica (CCA por sus siglas en inglés). Este análisis utiliza las señales seno y coseno como funciones base para representar las frecuencias correspondientes a los estímulos visuales. En este trabajo, se utiliza una extracción de características espaciotemporal con regresión lineal multivariante (MLR por sus siglas en inglés). Este método es implementado para la reducción dimensional de las señales EEG en un vector de características óptimo para la discriminación de las señales. Los resultados obtenidos en este trabajo demuestran una gran confiabilidad en el reconocimiento de estrategias mentales SSVEP.

1.6 Justificación y objetivos de la tesis

En la literatura existen diversas investigaciones referentes a la detección de la señal P300. No obstante, los métodos enfocados a potenciales relacionados a eventos visuales requieren de

una gran cantidad de electrodos para su correcto funcionamiento además de que los dispositivos utilizados en la mayoría de los trabajos reportados en la literatura son costosos, incómodos, no son portables y tienen protocolos muy estrictos para la colocación de los electrodos. Derivado de esto, la investigación y generación de nuevas metodologías para el análisis de la señal P300 resulta costosa y complicada, razón por lo cual, se ha limitado el análisis de esta señal a entornos de laboratorio de investigación. Adicional a esto, cuando se tiene una gran cantidad de electrodos, se requiere de un proceso largo y complicado para colocarlos a los sujetos de prueba. Asimismo, el tener una gran cantidad de electrodos incrementa la complejidad computacional por lo que se requiere de equipos que soporten altos niveles de procesamiento. En consecuencia, aún se requiere hacer más trabajo para el desarrollo de algoritmos y sistemas, que ayuden a tener altos desempeños en el análisis de la señal P300, reduciendo considerablemente la cantidad de electrodos y el costo de los dispositivos BCI, además de que sean cómodos para los usuarios y tengan protocolos simples para la colocación de los electrodos. Esto motiva al desarrollo de nuevos esquemas de análisis de señales EEG que funcionen de forma más eficiente, utilizando las tecnologías más recientes en BCI.

Aunado a esto, el surgimiento de nuevos dispositivos para la adquisición de señales EEG ha permitido el desarrollo de sistemas menos costosos que pueden obtener una medición de las señales con mejor calidad. Estos dispositivos han propiciado el desarrollo de más investigación en la generación de aplicaciones basadas en el procesamiento de las señales EEG y al alcance de cualquier persona. Entre los dispositivos de nueva creación se encuentra el OpenBCI® [10], el cual es un sistema que entrega la medición de la señal EEG en 16 canales. A diferencia de otros dispositivos BCI, la configuración de las posiciones del sistema 10-20 es flexible y es un dispositivo de código abierto, lo que permite el acceso a los datos crudos y de esta manera se puede trabajar en tiempo real con diferentes plataformas.

Sin embargo, el tener un dispositivo de bajo costo no implica tener un sistema más funcional. Para ello, es necesario desarrollar nuevos métodos que analicen de una manera eficiente la señal P300 para mantener un alto desempeño fuera de los laboratorios. En este sentido, la extracción de características es un punto clave para el correcto análisis de la señal P300, debido a que las características utilizadas tienen un gran impacto en la eficiencia de la clasificación. Adicional a

ello, la señal P300 no se genera de la misma manera en todas las personas, por lo cual se requiere de algoritmos que sean capaces de modelar de forma automática el comportamiento de la señal en cada sujeto y extraer las características más descriptivas. Esto último es importante, ya que los algoritmos que se reportan en la literatura tienen desempeños muy buenos, pero con la desventaja de que funcionan para pocos sujetos. Es por esto que se requiere de algoritmos que interpreten las señales derivadas de sistemas de medición de tecnología reciente que ayuden a extraer las características y clasificar las señales EEG con P300, adaptándose a las señales cerebrales de cada persona sin intervención humana.

Por lo tanto, el propósito principal de este trabajo de tesis es contribuir con el desarrollo de nuevos métodos para la detección de las señales P300, utilizando dispositivos de tecnología reciente y de bajo costo. Adicional a ello, para reforzar la contribución de este trabajo, se va a tomar como paradigma un esquema bio inspirado. Como aplicación, se pretende impactar en el desarrollo de nuevos esquemas de terapia y sistemas de comunicación para personas con capacidades diferentes. De esta manera, el objetivo es diseñar algoritmos de extracción de características que encuentren la información más descriptiva de las señales P300. Para llevar a cabo esto, el diseño de los métodos propuestos se va a utilizar el deletreador de Donchin, el medio más popular en la literatura para evocar la señal P300. Para la implementación, se va a utilizar uno de los escenarios del sistema Terapiatec[®], el cual es un sistema de terapias virtuales para rehabilitación motora, enfocado a personas con discapacidades que se encuentra implementado en los Centros de Atención Múltiple del estado de Chihuahua.

Otro aspecto importante en el desarrollo de esta investigación es que forma parte del proyecto “Adquisición, procesamiento y análisis de señales EEG dirigida a generar nuevo conocimiento y tecnología para interfaz cerebro-computadora orientada a rehabilitación virtual” que actualmente está siendo desarrollado en el Laboratorio de Sistemas PVR[®] (Percepción Visual con Aplicaciones en Robótica) de Instituto Tecnológico de Chihuahua.

Este trabajo se organiza de tal manera que en el capítulo 2 se detallará la metodología de obtención y generación de las señales EEG con P300. El capítulo 3 describe un método de extracción de características y clasificación de la señal P300 mediante redes neuronales

artificiales. El capítulo 4 muestra el desarrollo de un método para la detección de la señal P300 mediante una red neuronal convolucional. En el capítulo 5 se describe el desarrollo de una interfaz para la selección de objetos basada en la señal P300. El capítulo 6 contiene el análisis y comparación de los métodos propuestos. Por último, en el capítulo 7 se muestran las conclusiones obtenidas y trabajo futuro.

CAPÍTULO 2. METODOLOGÍA DE OBTENCIÓN Y GENERACIÓN DE SEÑALES EEG CON P300

La señal P300 se puede evocar mediante diversos estímulos visuales y auditivos tales como deletreadores, menús, escenarios, sonidos, entre otros. Los más populares en la literatura para evocar la señal P300 con propósitos de comunicación son los deletreadores, los cuales consisten en generar los estímulos mediante una interfaz visual, mostrando un conjunto de caracteres alfanuméricos a un sujeto de prueba. Para registrar la señal, se debe contar con un dispositivo BCI que registre las señales EEG que contengan el P300.

El deletreador de Donchin es un estándar en la literatura para evocar la señal P300. Se trata de una interfaz visual para la selección de caracteres alfanuméricos como la que se muestra en la figura 2.1. La señal P300 se evoca mediante intensificaciones aleatorias (destello en cierta fila o columna como se muestra en la tercera fila de la figura 2.1). Estas intensificaciones son de corta duración y se basan en el paradigma Oddball que consiste en presentar estímulos aleatorios repetidamente [21].



Figura 2.1 Interfaz del deletreador de Donchin [25].

Las señales estudiadas en este capítulo se obtuvieron mediante el deletreador de Donchin de dos maneras. La primera fue mediante una base de datos disponible en la red, que se utiliza en la literatura como estándar para realizar pruebas. La segunda fue generarlas y adquirirlas

mediante el dispositivo OpenBCI[®], utilizando una metodología de adquisición desarrollada en este capítulo. Todos los experimentos mostrados en este capítulo se enfocan en obtener adecuadamente las señales P300 para detectarlas de una manera eficiente y para presentarlos, la estructura de este capítulo se compone de cuatro secciones. En la sección 2.1 se presenta la base de datos II de BCI Competition III, que contiene un conjunto de señales EEG obtenidas con el deletreador de Donchin. En la sección 2.2 se muestra el dispositivo OpenBCI[®], el cual es el encargado de medir y registrar las señales EEG. En la sección 2.3 se detalla la metodología utilizada para la adquisición de las señales EEG mediante el dispositivo OpenBCI[®]. Por último, en la sección 2.4 se describe la forma de generar la señal P300 con el deletreador de Donchin y el OpenBCI[®].

2.1 Base de datos de señales EEG

Existen diversas bases de datos que contienen registros de señales P300 derivadas del deletreador de Donchin. La más popular en la literatura es la base de datos II de BCI Competition III [25]. Esta base de datos se creó en el 2003 para evaluar los algoritmos de procesamiento de señales y clasificación en sistemas BCI basados en P300. Esta evaluación se hizo en una competencia que fue muy popular y muchos trabajos recientes siguen utilizándola para el diseño de sus algoritmos.

La base de datos II se compone de las señales de dos sujetos que contienen los potenciales relacionados a eventos P300 [25]. Las muestras fueron tomadas bajo las condiciones descritas en [21], por medio del software BCI2000[®] [19], el deletreador de Donchin y el dispositivo g.tec[®] [9]. El objetivo principal de esta base de datos es predecir los caracteres que los dos sujetos deseaban escribir fijando la vista en ellos. Las señales de esta base de datos fueron adquiridas a una frecuencia de 240 Hz y además fueron procesadas con un filtro pasa bandas de 0.1 a 60 Hz (no se especifica el tipo de filtro). Se utilizaron 64 electrodos del sistema internacional 10-20 con la configuración que se muestran en la figura 2.2. Los experimentos se basaron en que cada sujeto debía deletrear un conjunto de palabras teniendo un total de 185 caracteres. El deletreo de los caracteres se realiza de la siguiente forma [25]:

- Se le muestra al sujeto de pruebas la interfaz del deletreador (figura 2.1), para que identifique el carácter que desea deletrear.
- Las filas y columnas se intensifican aleatoriamente un total de 180 veces por carácter. Como se tienen seis filas y seis columnas (ver figura 2.1), cada fila y cada columna se intensifican 15 veces. El tiempo entre cada intensificación es de 175 ms: 100 ms se presenta la intensificación y 75 ms no hay intensificación.
- Se le pide al sujeto que realice un conteo mental de las intensificaciones que tiene cada uno de los caracteres que deletreará. Para cada carácter tiene que contar 30 intensificaciones, debido a que se intensifican 15 veces la fila y 15 veces la columna en donde está cada carácter. Esto es con el fin de asegurar que el sujeto este poniendo atención a cada una de las intensificaciones.
- El lapso entre el deletreo de cada carácter es de 2.5 s.

El deletreo de 185 caracteres que realizó cada sujeto de pruebas se hizo en cinco sesiones para evitar el cansancio durante las sesiones.

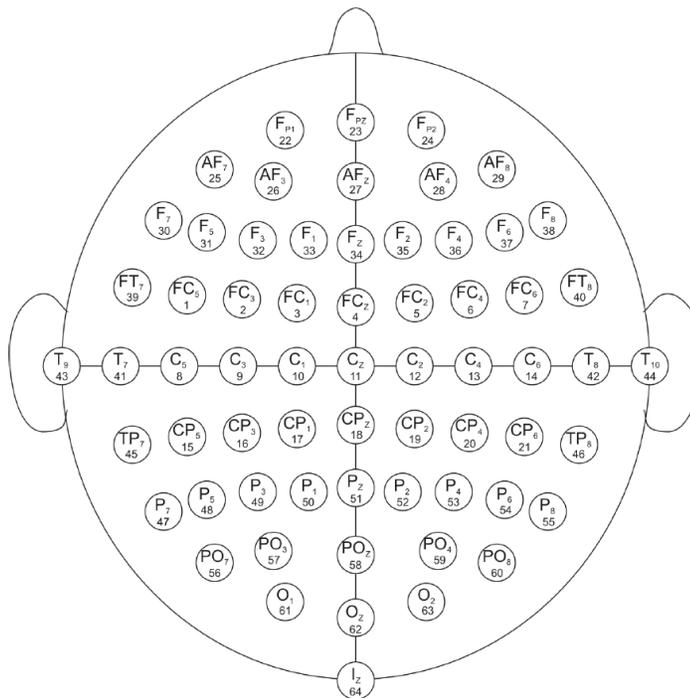


Figura 2.2 Electrodo utilizado en la base de datos II de BCI Competition III [25].

De los 185 caracteres, 85 se utilizan para el entrenamiento y 100 para la prueba. Debido a que estos caracteres fueron deletreados por palabras, estos se ordenan aleatoriamente, para no detectar aquellos caracteres erróneos infiriéndolos por las palabras, y de esta forma no se pueda hacer trampa en la medición de los desempeños. Dicha base de datos es un estándar en la literatura para que otros investigadores propongan sus métodos para procesar la señal P300.

Esta base de datos se utilizó para realizar pruebas que llevaran a entender el funcionamiento del deletreador de Donchin y el manejo del paradigma Oddball. Para las pruebas se realizó un trabajo que se publicó en [26], el cual consiste en el diseño de un algoritmo de extracción de características y clasificación de la señal P300 mediante un solo canal del sistema internacional 10-20 (en este caso fue el Cz). La extracción de características se realizó mediante el cálculo del promedio de los datos obtenidos en segmentos que van de los 300 a 500 ms después de las intensificaciones, para cada uno de los caracteres, y un índice de correlación con una señal base. La clasificación de la señal P300 se llevó a cabo por medio de una máquina de vectores de soporte, y para la detección del carácter se utiliza un método de inferencia probabilística. Con base en el trabajo realizado en este artículo, se dio inicio a la investigación de una primera propuesta para reducir la cantidad de canales necesarios para la detección de la señal P300. El análisis de los resultados se encuentra en [26].

Además de BCI Competition III, existen otras bases de datos con señales P300 que solo son utilizadas por los mismos investigadores que las generaron ya que tienen aplicaciones muy particulares. Una de ellas es LINI [27], la cual se utiliza para analizar diferentes factores que pueden afectar a la señal P300 evocada por medio del deletreador de Donchin. En [28] generan una base de datos para la selección de objetos denominada MMSPG. BNCI HORIZON [29] es otra base de datos que contiene la señales de sujetos con esclerosis lateral amiotrófica (ALS por sus siglas en inglés). *Visual P300 Speller* [30] también contiene señales adquiridas por medio del deletreador de Donchin para el deletreo de algunas palabras.

2.2 Dispositivo OpenBCI®

Aunque la base de datos II de BCI Competition III es un estándar en la literatura, se adquirió con un dispositivo que ahora se considera de tecnología antigua, se utilizaron solo dos sujetos

de prueba y fue diseñada originalmente para una competencia que se realizó en el año 2003. Sin embargo, ante el surgimiento de nuevos dispositivos para la adquisición de señales EEG, se decidió hacer una base de datos con señales EEG, esperando una mejor calidad debido al avance tecnológico tanto en electrodos como en amplificadores y además, con más sujetos para garantizar la reproducibilidad. Por lo tanto, para el diseño de los métodos fue seleccionado el dispositivo OpenBCI[®], el cual es un BCI de bajo costo, se basa en tecnología reciente y es de código abierto, es decir, no tiene candados ni licencias para el manejo de los datos crudos y el entorno de desarrollo. Los elementos principales que componen el OpenBCI[®] son la tarjeta de adquisición Cyton y la tarjeta para la comunicación USB. Además, los sensores que utiliza el dispositivo pueden ser secos o húmedos.

En la figura 2.3 se muestra la tarjeta OpenBCI[®] Cyton, la cual tiene la capacidad de obtener y digitalizar las señales EEG por medio de 16 canales, además cuenta con dos canales para referencia y dos más como tierra.

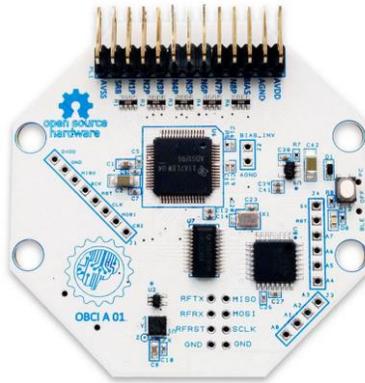


Figura 2.3 OpenBCI[®] Cyton [10].

La tarjeta Cyton se conecta hacia un dispositivo USB por medio de un módulo bluetooth 4.0 denominado RFDuino [31], este dispositivo se muestra en la figura 2.4.



Figura 2.4 Dispositivo USB para la conexión bluetooth con OpenBCI[®] [10].

2.2.1 Ultracortex Mark IV

El Ultracortex Mark IV es un casco para trabajar con el dispositivo OpenBCI®, diseñado para posicionar electrodos secos (no requieren de un gel conductor). Como se muestra en la figura 2.5, tiene predefinidos la posición de algunos de los electrodos más utilizados del sistema internacional 10-20 y cuenta con ajuste para diferentes tamaños de cabezas. Las localizaciones de electrodos que pueden ser utilizadas mediante este casco, son las que se encuentran marcadas con un círculo en la figura 2.6.

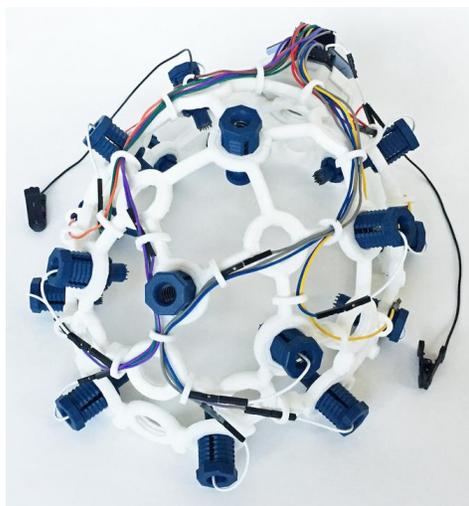


Figura 2.5 Casco Ultracortex Mark IV [10].

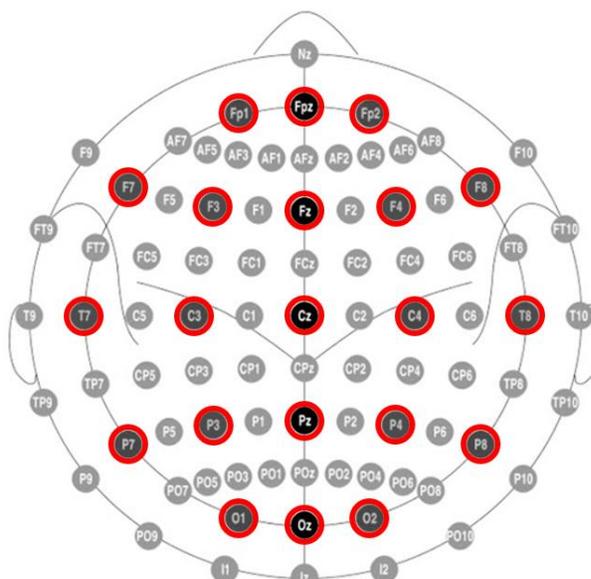


Figura 2.6 Electrodos que pueden ser utilizados mediante el casco Ultracortex Mark IV.

Existen dos tipos de electrodos compatibles con este casco, los cuales se muestran en la figura 2.7; electrodos para zonas donde hay cabello (figura 2.7a) y electrodos para zonas donde no hay cabello (figura 2.7b). Los electrodos para zonas donde hay cabello pueden resultar muy incomodos ya que cada uno tiene varias puntas que aseguran el contacto en la piel, pero pueden lastimar al sujeto de prueba.

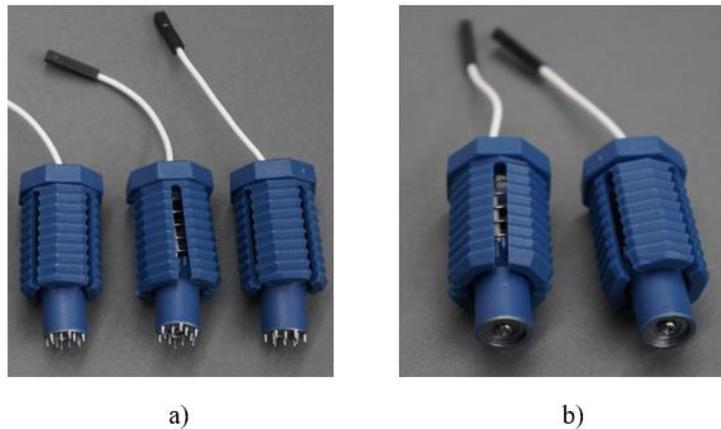


Figura 2.7 a) electrodos para zonas con cabello, b) electrodos para zonas sin cabello [10].

2.2.2 Electrodos de copa y gel conductor

Además de los electrodos secos, se cuenta también con electrodos de copa húmedos (requieren de un gel conductor). En la figura 2.8 se muestran estos electrodos, cuya forma es similar a una copa y se encuentran chapeados en oro para tener una mejor conducción. Estos electrodos se conectan a la tarjeta Cyton de OpenBCI® pero no son compatibles con el casco Ultracortex Mark IV. Es decir, se pueden colocar en cualquier parte de la cabeza, pero resulta complicado posicionarlos con el sistema internacional 10-20, ya que se tiene que encontrar la ubicación de cada una de las posiciones que se desea analizar. Para ello, se pueden utilizar diferentes formas ubicar estos electrodos, una de ellas adherirlos con cinta adhesiva o también se puede pegar los electrodos a una banda elástica.



Figura 2.8 Electrodo de copa para el dispositivo OpenBCI® y gel conductor [10].

2.2.3 Experimentos para la selección de electrodos

Se realizaron pruebas con el OpenBCI® para conocer la mejor forma de utilizar los electrodos. Las primeras pruebas se realizaron con los electrodos secos del casco Ultracortex Mark IV. En la figura 2.9 se muestra el casco colocado en un sujeto de prueba. Anteriormente, se mencionó que este casco ayuda a la fácil colocación de los electrodos de acuerdo con el sistema internacional 10-20. Sin embargo, se encontró que los electrodos con puntas para zonas con cabello pueden lastimar al sujeto de pruebas. Además, los sujetos de prueba se quejaban después de algunos minutos debido a la presión que el casco ejerce. Esta incomodidad causaba que no se terminara la sesión de toma de muestras.

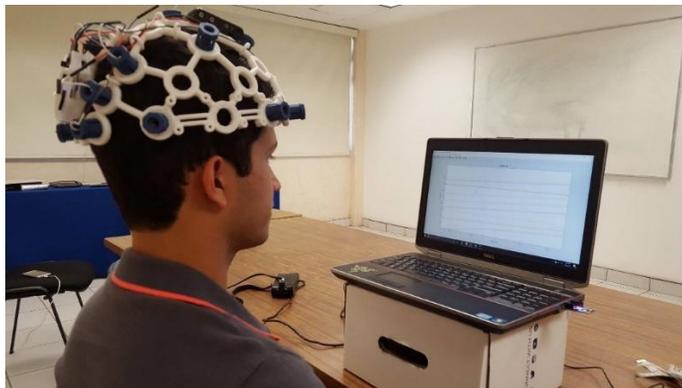


Figura 2.9 Sujeto de pruebas con el dispositivo Ultracortex Mark IV.

Otra opción es el uso de los electrodos de copa con el gel conductor. Estos electrodos no ejercen tanta presión en el sujeto de pruebas, pero, no es tan fácil colocarlos en el lugar adecuado de acuerdo con el sistema internacional 10-20. Para dar solución a esto, se utilizó una banda elástica para sostener los electrodos en la posición adecuada. En la figura 2.10 se muestra la banda que se utiliza para la colocación de los electrodos, la cual tiene indicado algunos de los electrodos de la parte occipital.

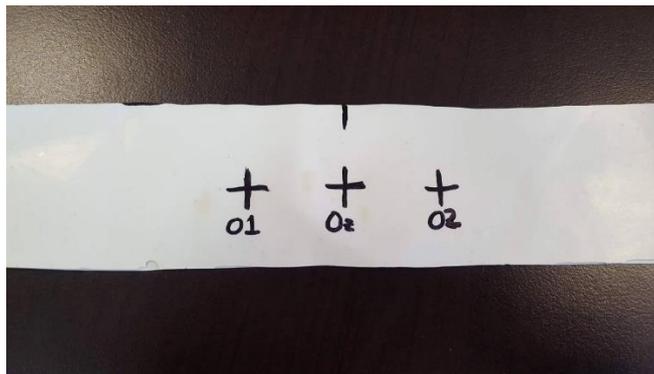


Figura 2.10 Banda elástica utilizada para el soporte de los electrodos de copa.

Para cada uno de los electrodos utilizados es necesario aplicar un gel especializado en neuro diagnósticos para mejorar la conductividad de los electrodos. La manera de colocar el gel conductor en los electrodos de copa es rellenándolos hasta que quede una capa uniforme en la parte cóncava del electrodo, como se muestra en la imagen de la figura 2.11.



Figura 2.11 Electrodo de copa con el gel conductor.

Al realizar pruebas con distintos sujetos, se les preguntó sobre su comodidad con el posicionamiento de los electrodos y se pudo corroborar que los electrodos de copa son más cómodos que los utilizados por el Ultracortex Mark IV.

Otra ventaja de utilizar los electrodos de copa fue que el contacto de los electrodos con el cuero cabelludo es más estable, es decir, los electrodos no se desconectan mientras se está realizando la sesión o cuando se realizan movimientos corporales. En cambio, los electrodos secos del Ultracortex Mark IV pierden contacto cuando se realiza algún movimiento brusco, como acomodarse en la silla o girar la cabeza.

Por último, en la figura 2.12 se muestra a un sujeto de pruebas con los electrodos de copa colocados mediante la banda elástica en la cual se tienen tres electrodos conectados en la parte occipital, además también tiene conectados la referencia y la tierra.

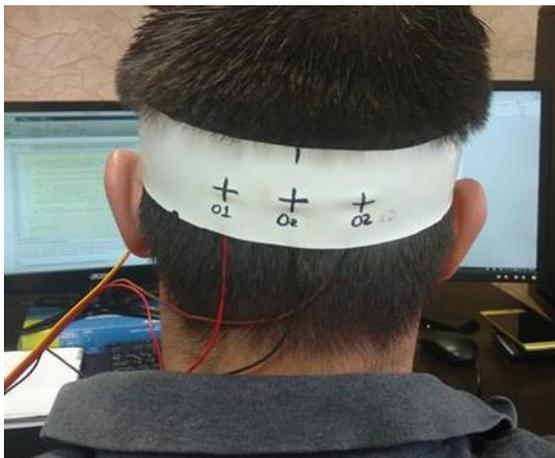


Figura 2.12 Sujeto de pruebas con los electrodos de copa colocados mediante la banda elástica.

Con base en estas pruebas, fueron seleccionados los electrodos de copa con banda elástica para la realización de los experimentos ya que resultaron ser más cómodos para los sujetos de pruebas, lo que permite realizar pruebas con mayor duración. Además, es poco probable que estos electrodos hagan mal contacto una vez que son colocados.

2.2.4 Conexión de OpenBCI® con Matlab®

OpenBCI® puede trabajar con distintos entornos que tengan compiladores como C, C++, Python, Java, C# y Matlab®. Entre estos, Matlab® es un entorno simple de manejar y puede acceder a los datos crudos del dispositivo OpenBCI® mediante la herramienta llamada *lab streaming layer*. Este acceso permite trabajar en tiempo real con datos crudos. En la figura 2.13 se muestran los datos crudos obtenidos en tiempo real en el software Matlab®, el cual ayuda a realizar el procesamiento de las señales de una manera sencilla y rápida ya que cuenta con diversas herramientas especializadas.

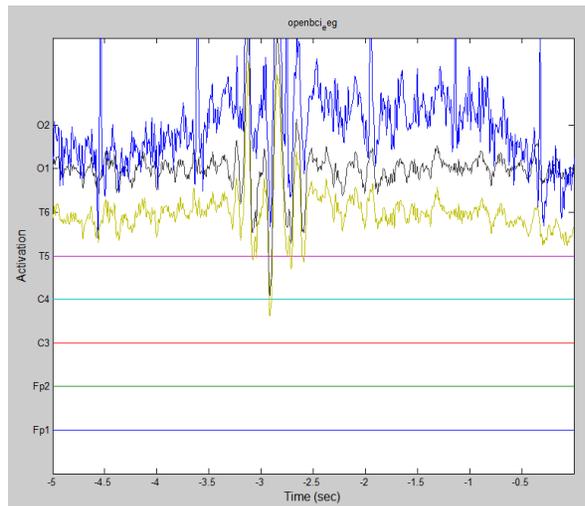


Figura 2.13 Flujo de datos del OpenBCI® en Matlab®.

Para la verificación de la correcta adquisición de los datos crudos en Matlab®, se realizó un análisis visual de la señal obtenida en el canal Oz. El análisis consistió en generar ondas *Alpha* y visualizarlas en una gráfica en Matlab®. Se sabe de [4] que cuando una persona cierra los ojos, la amplitud de las señales *Alpha* en la parte occipital se incrementa. De esta manera, se realizaron varias pruebas que consisten en que un individuo abra y cierre los ojos por lapsos de dos segundos. Esta prueba tiene la finalidad de poder observar si efectivamente la amplitud en las señales *Alpha* se incrementa y si la señal cambia en tiempo real. El resultado se puede observar en la figura 2.14, donde las señales *Alpha* que se encuentran dentro del rectángulo indican cuando el sujeto de pruebas tenía los ojos cerrados. El poder observar este fenómeno indica que los datos se adquieren correctamente en Matlab®, ya que es uno de los comportamientos

naturales de las señales EEG. Cabe mencionar que no existe retardo en el incremento de la amplitud de las señales *Alpha* cuando el sujeto cierra los ojos. Por último, para corroborar que la frecuencia de muestreo fuera correcta, se midió el tiempo en el que llegaba cada una de las muestras por medio de Matlab®, el cual fue de 4 ms, que corresponde al periodo de la frecuencia de muestreo indicada de 250 Hz.

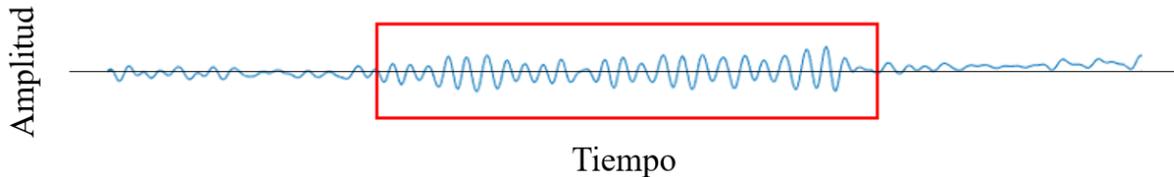


Figura 2.14 Señales en banda de frecuencia *Alpha* en el canal Oz cuando un sujeto de pruebas cierra los ojos.

En esta tesis se utilizó el software Matlab®, debido a que permite procesar las señales obtenidas por OpenBCI® en tiempo real y además, a diferencia de otros entornos de desarrollo con Matlab® no se requiere transferir la información entre distintos paquetes de software.

2.3 Protocolo para la adquisición de señales

Para utilizar correctamente el dispositivo OpenBCI®, se aplicó un protocolo de adquisición de señales que mantiene bajo control diversos aspectos que pudieran afectar una correcta lectura de los datos. En un principio se realizaron diversas pruebas para establecer el protocolo de adquisición que considere las características del sujeto de pruebas, el área o entorno donde se realizan las pruebas y el monitor que genera el estímulo externo. Después de varios experimentos se estableció lo siguiente.

En lo referente al sujeto de pruebas, se deben conocer los datos personales y condiciones de los individuos que están colaborando para la adquisición de las señales EEG. Por esta razón, se realiza una encuesta inicial para conocer al sujeto, tal como realiza en [23] [27] [32]. Con base en un análisis realizado con diversos trabajos en la literatura, se observó que los aspectos más relevantes a considerar en una encuesta para la correcta caracterización de los sujetos de pruebas al obtener potenciales evocados son los que se muestran en la figura 2.15, el cual es un formato que se llena antes de realizar una sesión y permite conocer la opinión de los sujetos de prueba,

generando información para realizar ajustes en aspectos tales como comodidad o la cantidad de muestras.

Nombre	
Edad	
Sexo	
Visión (normal o corregida)	
Extremidad predominante	
Ojo predominante	
Estado mental	
Ocupación	
Opinión sobre la sesión.	
Prefijo de muestras	

Figura 2.15 Encuesta realizada a los sujetos de pruebas antes de cada sesión.

En relación con el área o entorno para la toma de muestras se realizó en un ambiente libre de ruido visual y auditivo para evitar las distracciones en los sujetos de pruebas y obtener mediciones con el menor ruido posible. Las muestras tomadas en este entorno son utilizadas para el diseño y validación de los métodos propuestos en este trabajo de tesis. No fue necesario controlar la iluminación, ya que en un entorno real donde pueda ser utilizado este tipo de sistemas es poco probable tener un control total de la iluminación. Además, la temperatura de la habitación se mantiene en 24 grados centígrados para que el sujeto este cómodo. Los sujetos de pruebas son colocados en una silla que no tenga ruedas, con respaldo fijo y descansabrazos, para que el sujeto se mantenga en una posición fija y de frente al monitor. El monitor donde se muestran los estímulos visuales para la generación de la señal P300, es colocado a la altura de la cabeza y a una distancia aproximada de 50 cm como se realiza en [33].

2.4 Obtener señal P300 con deletreador de Donchin

Una vez conocida la forma de funcionamiento del OpenBCI[®], el entorno a desarrollar (en este caso fue Matlab[®]) y teniendo establecido el protocolo de adquisición, el siguiente paso es generar y registrar las señales EEG con P300. Para ello, se implementó el deletreador de Donchin y se realizaron diversas pruebas. Esta sección presenta los detalles sobre la implementación del deletreador, así como los experimentos realizados para determinar la configuración del OpenBCI[®] para una obtención efectiva de la señal P300.

2.4.1 Implementación del deletreador de Donchin Matlab®

Como se mencionó previamente, el deletreador de Donchin ha sido el estándar en la comunidad científica, para estudiar la señal P300. Los algoritmos que se desarrollan con este deletreador pueden ser aplicados en algunas otras interfaces como seleccionadores de objetos o detectores de mentiras [22].

Donchin y L. A. Farwell desarrollaron en [21] este deletreador basado en el paradigma Oddball. Los estímulos son generados por medio de la intensificación de las columnas y filas de una matriz de 6x6 que contiene las letras del alfabeto y algunos otros símbolos o comandos; este deletreador es capaz de generar la señal P300 para que posteriormente sea captada por algún método de adquisición de señales cerebrales. Durante el experimento se les pide a los sujetos de pruebas que fijen su vista en el carácter deseado y posteriormente las filas y columnas de la matriz comienzan a intensificarse aleatoriamente. Cada intensificación del carácter que el sujeto está observando, deberá generar una deflexión positiva a los 300 ms similar a la mostrada en figura 2.16.

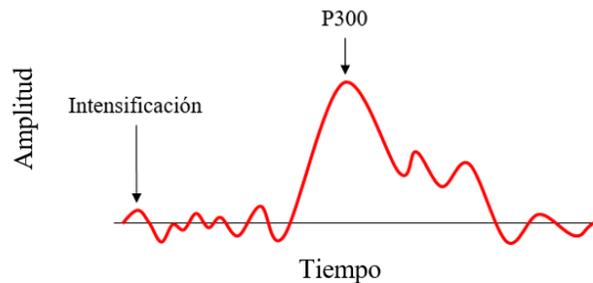


Figura 2.16 Forma aproximada de la señal P300.

Cabe mencionar que la amplitud de la señal P300 es similar a la del ruido obtenido, por lo tanto, las intensificaciones para las filas y columnas se deben realizar varias veces para que sea posible detectar cuál es el carácter deseado mediante diversas técnicas que incrementan la relación señal ruido [21]. Las intensificaciones de las filas y columnas se deben generar aleatoriamente, ya que estudios demuestran que entre más inesperado sea el estímulo mayor será la amplitud de la señal P300 [34].

La herramienta que permitió el diseño de deletreador fue GUIDE, la cual genera automáticamente el código a partir del diseño gráfico. La interfaz resultante se muestra en la figura 2.17. Esta interfaz para la toma de muestras incluye los siguientes elementos:

- Cuadro de texto para ingresar el carácter que el sujeto va a deletrear para que quede guardada en la base de datos su etiqueta para entrenamiento y validación.
- Botón para iniciar la toma de muestras, el cual marca una advertencia si no se ingresó el carácter que se deletreará o es algún carácter que no se encuentra en la matriz.
- Indicador con el número de muestras que se han tomado.
- Cuadro de texto para ingresar el nombre del archivo donde se guardarán las muestras, en caso de que el archivo no exista se creará uno nuevo. Si el archivo existe, éste no será reemplazado, sino que se agregará una muestra más al archivo.
- Matriz de seis filas y seis columnas con caracteres alfanuméricos donde se generan los estímulos visuales.

Las intensificaciones consisten en cambiar el color de las filas y columnas de gris a blanco, por ejemplo, la columna dos de la figura 2.17 se encuentra intensificada.

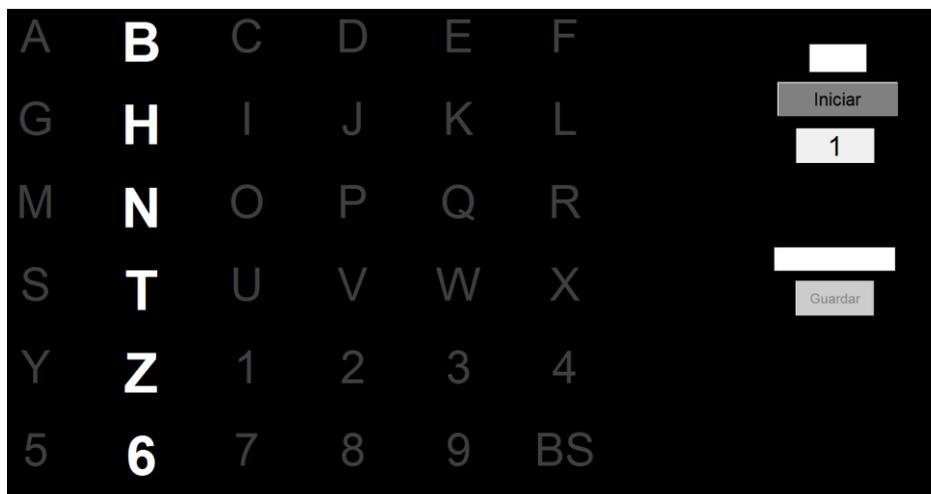


Figura 2.17 Interfaz del deletreador para la toma de muestras.

Para el funcionamiento del deletreador de Donchin, se utilizó un diseño similar al reportado en [25] donde cada lapso para deletrear un carácter se denomina época de deletreo, misma que

inicia con 4 s sin presentar cambios en las filas y columnas para que el sujeto identifique la posición del carácter que desea deletrear y luego de este lapso, inician los periodos de intensificación de manera aleatoria en las filas y columnas. Cada periodo de intensificación consta de un lapso de 100 ms en el que los caracteres de la fila o columna son presentados en color blanco y agrandados respecto a su tamaño inicial, y otro lapso de 72 ms en el que los caracteres permanecen en color gris y sin agrandamiento. El lapso de 100 ms con caracteres blancos y agrandados será llamado en lo futuro intensificación. Durante una época, cada fila y cada columna presentarán el periodo de intensificación 15 veces y el orden en que se vayan presentando será almacenado en una base de datos para posteriormente realizar el procesamiento de las señales registradas.

El sujeto deberá llevar un conteo mental de las intensificaciones que perciba en el carácter observado y para facilitar esto, fue que se incorporó el agrandamiento de letras ya que experimentalmente se corroboró con 5 sujetos de prueba, que cuando no había agrandamiento se perdía el conteo fácilmente y cuando se agrandaban las letras no se perdía. Este agrandamiento no provocó cambios significativos en la generación de la señal P300. Debido a que la interfaz muestra seis filas y seis columnas y cada una de ellas tendrá 15 periodos de intensificación, la época consta de 180 periodos de intensificación en total; por lo tanto, una época tendrá una duración de 35.788 s: 4 s sin cambio, 30.96 s de periodos de intensificación y 0.828 s sin cambio al final. En la figura 2.18 se muestra gráficamente el periodo de intensificaciones en una época de deletreo.

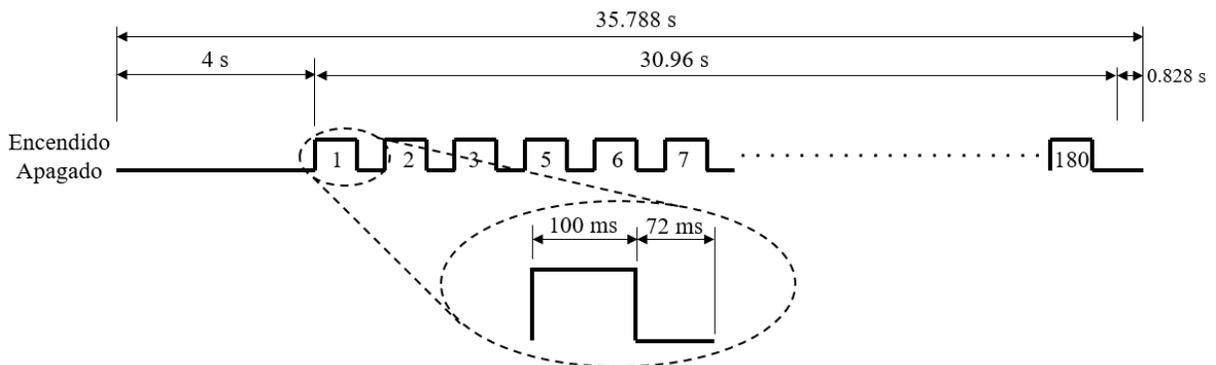


Figura 2.18 Periodo de intensificaciones en una época de deletreo.

2.4.2 Experimentación para la asegurar la generación de la señal P300

Al inicio de la etapa de experimentación se realizaron pruebas para decidir entre utilizar el Ultracortex Mark IV o los electrodos de copa del dispositivo OpenBCI[®], para transmitir los datos crudos por medio de la herramienta *lab streaming layer* a Matlab[®] con una frecuencia de muestreo de 250 Hz. Se encontró que la adquisición de las señales mediante los dos dispositivos tiene la misma calidad, sin embargo, el utilizar los electrodos de copa hacen que la sesión sea más cómoda y no se corre el riesgo de que los electrodos dejen de hacer contacto, por lo que estos últimos fueron seleccionados para todos los casos.

Una vez que se tiene la interfaz del deletreador de Donchin, es posible realizar experimentación para asegurar que la señal P300 se esté generando mediante los estímulos visuales. En esta sección se describe el procesamiento de las señales obtenidas con el deletreador de Donchin para poder visualizar en amplitud y latencia la respuesta de las intensificaciones que deben contener o no la señal P300. Para los experimentos realizados en el análisis de este deletreador, se utilizaron 5 sujetos de pruebas entre 20 y 25 años con visión normal o corregida. Para llevar a cabo la toma de muestras se aplicó el protocolo descrito en la sección 2.3 y se registraron 20 épocas de delecteo para cada sujeto. Se utiliza un total de 20 épocas de delecteo debido a que tras el primer experimento realizado para un sujeto, se encontró que eran suficientes para poder visualizar la señal P300 y no es una cantidad que provoque agotamiento en los sujetos. Con fines prácticos se muestran solo los resultados de uno de los sujetos de pruebas ya que el experimento se reprodujo para cada uno de ellos.

Las señales de las épocas de delecteo registradas consisten en datos crudos que se representan por $s_j(n)$, donde j representa los canales utilizados y n es el índice de tiempo. Los canales que se utilizaron fueron $j=\{C3, Cz, C4, P3, Pz, P4, O1, O2\}$ debido a que son las posiciones donde comúnmente se presenta la señal P300 [16]. El índice n representa tiempo, en el cual se tienen muestras con el intervalo de $1 \leq n \leq 8947$, ya que se graban 35.788 s a una frecuencia de muestreo de 250 Hz. En las muestras $1 \leq n \leq 1000$ no se genera ninguna intensificación, ya que estas mil muestras equivalen a 4 s y se recolectan con el fin de que el sujeto se prepare para el comienzo de las intensificaciones, además, ayudan con la estabilización

de un filtro que se explicará posteriormente. Como cada una de las intensificaciones se muestra durante 25 muestras y dura 18 muestras en aparecer la siguiente intensificación, se puede concluir que las intensificaciones se generan cada 43 muestras. Cada época de deletreo tiene un total de 180 intensificaciones, por lo tanto, se requieren 7740 muestras para representarlas. En las muestras $1001 \leq n \leq 8740$ es donde se contienen las señales producidas por las intensificaciones que pueden contener o no las señales P300. Finalmente, las 207 muestras $8741 \leq n \leq 8947$ se utilizan para completar un segundo después de la última intensificación, esto para en caso de que se requiera analizar la señal en este lapso. La descripción grafica de la señal $s_j(n)$ se muestra en la figura 2.19.

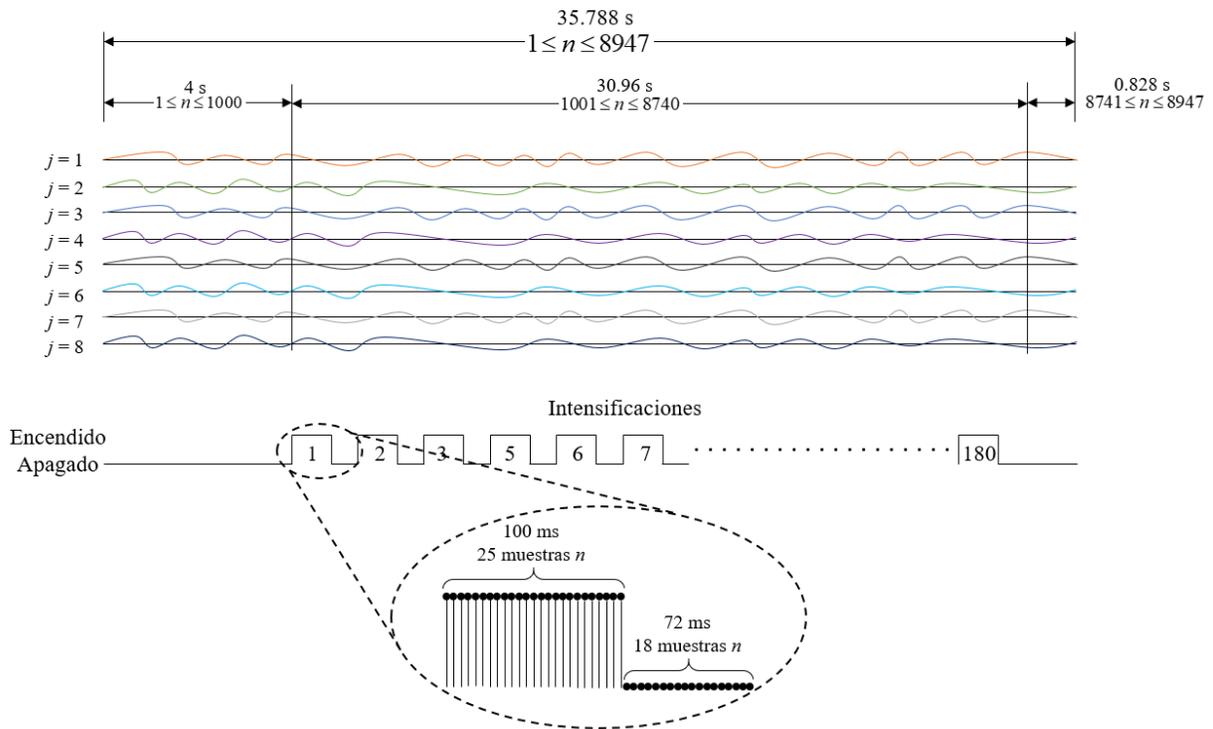


Figura 2.19 Composición de la señal $s_j(n)$ de una época de deletreo con OpenBCI® a 250 Hz.

El experimento consiste en obtener la respuesta promedio para cada intensificación del deletreador por filas y columnas, ya que se conoce cuáles de éstas deben contener o no la señal P300 y de esta manera poder realizar un análisis visual de las señales. Como se observó en la figura 2.16, las señales P300 tienen una deflexión positiva alrededor de los 300 ms, por lo tanto, cuando existe la P300 se espera ver esta deflexión.

Inicialmente, la señal $s_j(n)$ se filtra con un Butterworth pasa bandas de orden 3, con frecuencias de corte $\omega_{cb}=15$ Hz para el pasa bajas y $\omega_{ca}=1$ Hz para el pasa altas. El rango de frecuencias se eligió para conservar la información que contiene a la señal P300 y eliminar ruido [16]. La función Butterworth se eligió debido a que su respuesta a la frecuencia es suave y no tiene rizos en la banda de transición como se muestra en la figura 2.20. En la literatura se encontraron diversos tipos de filtro, pero no se abordan las razones para elegir el tipo.

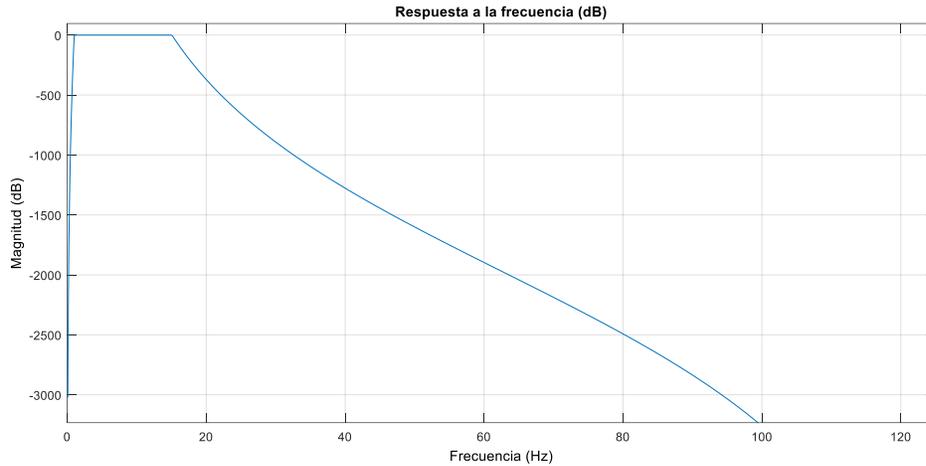


Figura 2.20 Respuesta a la frecuencia del filtro representada en dB.

La señal filtrada $s_{ff}(n)$ está dada por:

$$H(\omega): s_j \rightarrow s_{ff} \quad 2.1$$

donde

$$|H(\omega)|^2 = \frac{1}{1+(\omega/\omega_{cb})^6} \cdot \left(1 - \frac{1}{1+(\omega/\omega_{ca})^6} \right) \quad 2.2$$

$H(\omega)$ es la respuesta a la frecuencia del filtro, ω es la frecuencia que se evalúa, $\omega_{cb}=15$ Hz es la frecuencia de corte para el pasa bajas y $\omega_{ca}=1$ Hz la frecuencia de corte del pasa altas [35]. El filtro presenta distorsión de fase, razón por la cual tarda 3 s en estabilizarse. Por lo tanto, se aprovecha el periodo de latencia $1 \leq n \leq 1000$ para esperar a que el filtro tenga la respuesta esperada. Una vez pasadas las primeras 1000 muestras, el filtro se estabiliza y funciona adecuadamente cuando se presentan las intensificaciones.

Seguidamente, se genera un arreglo de segmentos de la señal $s_{ff}(n)$ para $1001 \leq n \leq 8947$, que consiste en segmentos de 600 ms por cada intensificación, lo que corresponde a 150 muestras (la frecuencia de muestreo es de 250 Hz). Este segmento es suficiente para detectar y visualizar la señal P300 debido a que la deflexión positiva de la señal P300 ocurre alrededor de los 300 ms con una latencia aproximada de 100 ms (el rango donde se puede presentar la señal P300 es de 250 a 400 ms después del estímulo [15]). De esta manera, la señal $s_{ff}(n)$ se descompone en segmentos que son representadas por $x_{ij}(m)$, donde las intensificaciones son numeradas con un índice i que va de 1 hasta 180 que corresponde a 15 intensificaciones por cada una de las 12 filas y columnas, en los canales j . Se utiliza un índice auxiliar m de muestras para cada uno de los segmentos.

Por último, se realiza un promedio respecto a i para todas aquellas señales $x_{ij}(m)$ que deberían contener la señal P300 y otro promedio para las que no. Esto es posible debido a que en estos experimentos se tiene conocimiento de los caracteres que los usuarios observan. Cuando $x_{ij}(m)$ debe contener la señal P300, se refiere a si la fila o columna que se intensificó contiene el carácter que el sujeto estaba observando. Para este caso particular solo 30 intensificaciones producen el estímulo para la generación de la señal P300, ya que son 15 de la fila en la que se encontraba el carácter deseado y 15 para la columna. Cuando $x_{ij}(m)$ no contiene la señal P300, es cuando las filas y columnas que se intensificaron no correspondían al carácter que el sujeto estaba observando, aquí son 150 intensificaciones las que no contienen P300. De esta manera se generan los siguientes vectores de promedios:

$$x_{Tj}(m) = \frac{1}{30} \sum_{i \in CC} x_{ij}(m) \quad CC : \text{Carácter correcto} \quad 2.3$$

$$x_{Fj}(m) = \frac{1}{150} \sum_{i \in NC} x_{ij}(m) \quad NC : \text{No es carácter correcto} \quad 2.4$$

donde $x_{Tj}(m)$ es el promedio de las señales $x_{ij}(m)$ para i con el carácter correcto y $x_{Fj}(m)$ es el promedio de las señales $x_{ij}(m)$ para i sin el carácter correcto, en cada época de delecteo para el

canal j . Las señales obtenidas en $x_{Tj}(m)$ y $x_{Fj}(m)$ serán utilizadas para verificar visualmente las diferencias entre una señal sin P300 y otra señal con P300.

2.4.3 Análisis de P300 en el deletreador de Donchin original

El análisis mostrado en esta sección consiste en visualizar el promedio de los vectores $x_{Tj}(m)$ que contienen la respuesta de aquellas intensificaciones que deben evocar las señales P300, y $x_{Fj}(m)$ el cual no se debe contener la señal P300, de las 20 épocas de deletreo del experimento realizado en la sección anterior.

Este análisis se realizó con la finalidad de saber si la señal P300 se está generando correctamente mediante la interfaz del deletreador de Donchin implementada en Matlab®. Además, el poder visualizar la amplitud y latencia de la señal P300 ayuda a definir cuál será el método que se utilizará para la extracción de características y clasificación. Los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ de un sujeto de ejemplo se muestran como P300 y no P300 respectivamente en la figura 2.21 y figura 2.22 para los ocho canales j . Estas figuras ayudan a realizar una comparación visual del promedio de aquellas intensificaciones que deben contener la señal P300 de las que no y de esta manera es posible conocer si la señal se está generando correctamente.

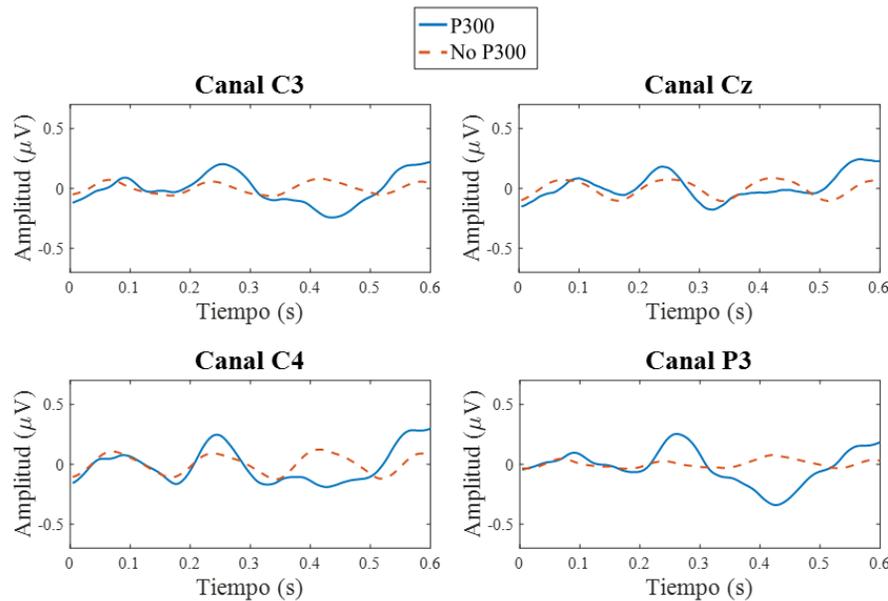


Figura 2.21 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) en los canales C3, Cz, C4 y P3.

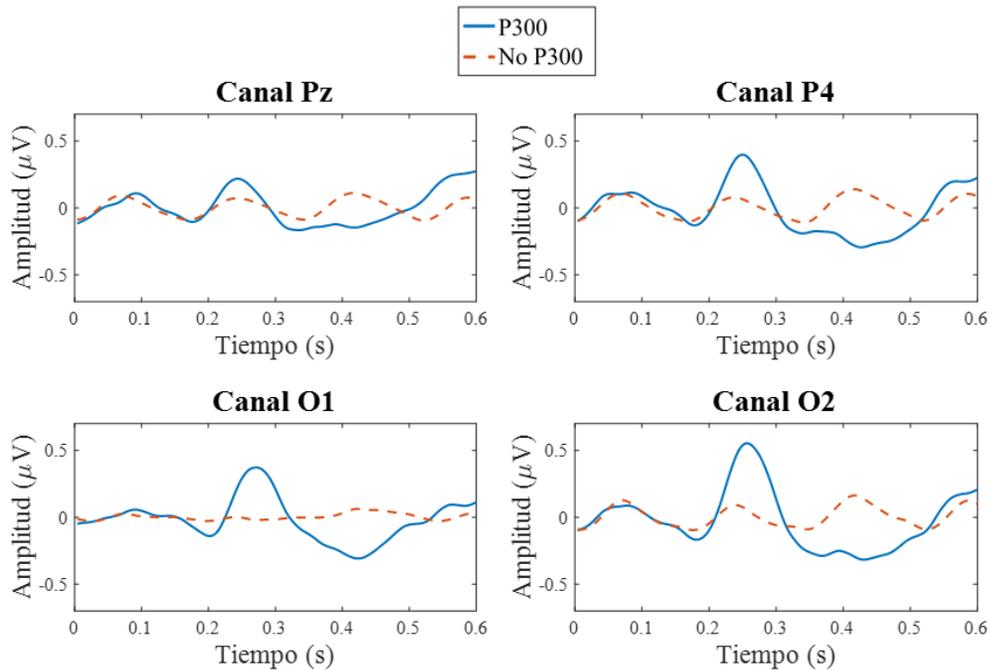


Figura 2.22 Promedio de los vectores $x_{Tf}(m)$ (P300) y $x_{Ff}(m)$ (No P300) en los canales Pz, P4, O1 y O2.

En la figura 2.21 y 2.22, se alcanza a observar que la señal que debe contener P300 en los canales C3, Cz, C4 y Pz, si tiene una deflexión positiva cerca de los 300 ms. Sin embargo, la amplitud de dicha señal es muy similar a la que no debe contener P300. En canales P3 y P4 las señales con P300 tienen una diferencia notoria contra aquellas señales que no. En el caso de los canales O1 y O2 es donde se presenta con mayor amplitud la señal P300, además tiene una mayor diferencia contra aquellas señales que no contienen P300. Este es un resultado esperado, ya que se trata de estímulos visuales y la parte occipital es donde se encuentra la corteza visual humana [36], encargada de procesar la visión.

Con base en los resultados obtenidos por los ocho canales mostrados, se llegó a la conclusión de seguir los experimentos tan solo con los canales occipitales O1 y O2 ya que es donde mejor se detectó la señal P300, en los 5 sujetos de pruebas. Los canales P3 y P4 fueron descartados debido a que la amplitud de la señal P300 no es tan notoria como en los canales occipitales, y el propósito de este trabajo es reducir la cantidad de canales utilizados para la detección de la señal P300. Otro motivo de comenzar a utilizar solo los canales occipitales es que se deje de utilizar

el casco Ultracortex Mark IV y se comienzan a utilizar los electrodos de copa mediante la banda elástica, para que las pruebas se realicen de una manera más cómoda para los sujetos de pruebas.

Una vez que se definió con cuales canales analizar la generación de las señales P300, se propone un experimento para corroborar si las intensificaciones en filas y columnas producen la señal P300 con la misma amplitud. El experimento consiste en verificar el comportamiento de la señal P300 con las intensificaciones por filas y columnas para los mismos 5 sujetos de pruebas. En la figura 2.23 se muestran los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ en los canales O1 y O2 por filas y columnas, para un sujeto, ya que el resultado se reprodujo. Donde, se puede observar que la señal P300 se produce con mayor amplitud cuando se trata de las columnas.

La generación de menor amplitud en la señal P300 por las líneas de caracteres horizontales, puede deberse a que entre más inesperado sea el estímulo, mayor es la amplitud de la señal P300 [21]. De esta manera, debido a que los sujetos de pruebas que participaron en el experimento están acostumbrados a leer de manera horizontal, se tiene la hipótesis de que la señal P300 se generó con menor amplitud en las filas que en las columnas, ya que no es tan inesperado ver caracteres de manera horizontal que vertical.

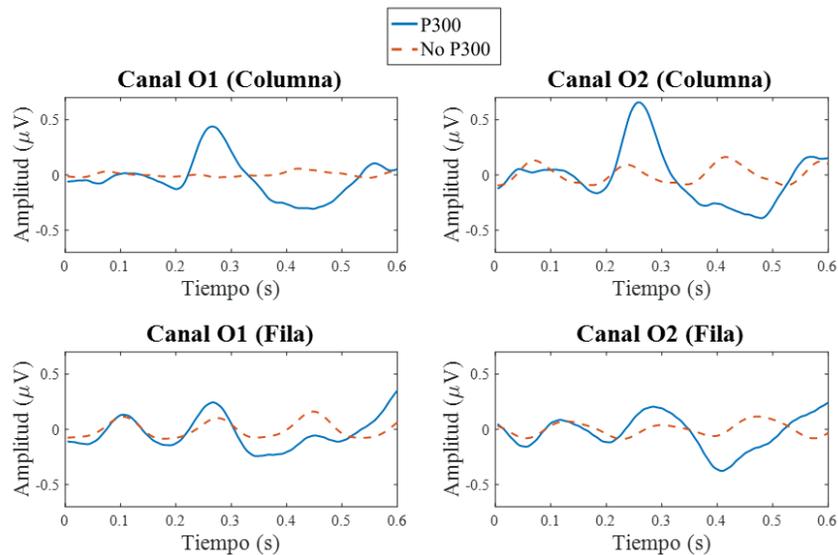


Figura 2.23 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) por filas y columnas en los canales O1 y O2.

2.4.4 Paradigma de Donchin modificado

Como se concluyó en las pruebas del deletreador de Donchin, la señal P300 tiene mayor amplitud cuando el estímulo se da en las columnas que en las filas. Por lo tanto, es más complicado encontrar cual es la fila donde se encuentra el carácter deseado que la columna. Es por esto que se decidió realizar una modificación al paradigma propuesto por Donchin, la cual consiste en encontrar el carácter deseado por medio de la intensificación de columnas solamente. El nuevo paradigma para la generación de las intensificaciones se lleva a cabo de la siguiente manera:

1. Se intensifican solamente las columnas: esto generará la información suficiente para poder detectar la columna donde se encuentra el carácter deseado.
2. Se transpone la matriz: las filas de la matriz ahora se vuelven columnas (de ahora en adelante serán llamadas filas^T) y se le da al sujeto un tiempo de 2 s para poder encontrar la nueva posición del carácter deseado.
3. Se intensifican las filas^T: de esta manera se permite obtener la información que corresponde a la fila del carácter deseado en la matriz original.

Cuando la matriz es transpuesta, el carácter deseado cambiará de posición a menos que este se encuentre en la diagonal que al transponerse no cambia. Con la práctica y conocimiento sobre la transposición de la matriz, se vuelve muy fácil detectar y predecir donde será la nueva posición de cada uno de los caracteres.

2.4.5 Análisis de P300 en el deletreador de Donchin modificado

El análisis mostrado en esta sección es similar al mostrado en la sección 2.4.3. Nuevamente se utilizaron cinco sujetos y los experimentos consisten en observar la respuesta promedio de las intensificaciones por columnas y filas^T. Por fines prácticos, se muestran las señales para el mismo sujeto de pruebas con 20 épocas de deletreo, pero ahora por medio del deletreador de Donchin modificado, el cual se caracteriza por generar los estímulos visuales mediante la intensificación de las columnas solamente.

A continuación, en la figura 2.24 se muestran los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$, en los canales O1 y O2 por columnas y filas^T.

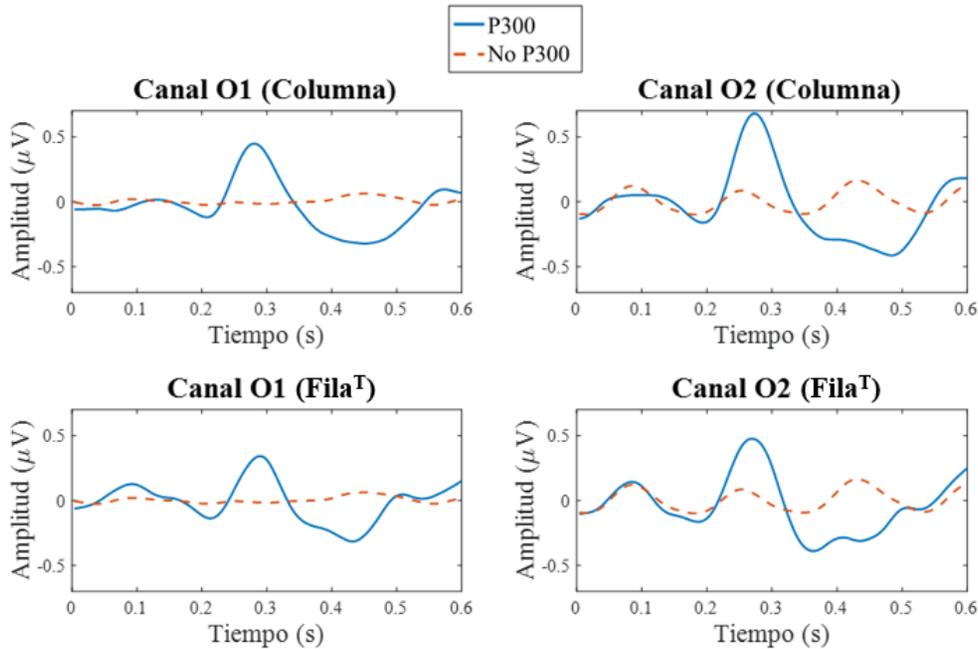


Figura 2.24 Promedio de los vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300) por filas^T y columnas en los canales O1 y O2, mediante el deletreador de Donchin modificado.

Mediante la figura 2.24 se puede observar como la señal P300 tiene una amplitud similar para las columnas y filas^T, esto es debido a que las intensificaciones se realizaron de la misma manera, vertical. Dado esto, será posible desarrollar algoritmos que detecten la señal P300 con el mismo método tanto para filas como para columnas.

2.4.6 Conclusiones de los experimentos de generación de P300

Mediante el análisis propuesto en la sección 2.4 se logró observar que la señal P300 se encontraba con mayor amplitud en los canales O1 y O2 del sistema internacional 10-20. Dados estos resultados, se optó por utilizar solo la parte occipital del cerebro, además, se agregará el electrodo Oz, con la finalidad de tener más resolución en dicha área. La ventaja de utilizar estos tres electrodos solo en la parte occipital es que la cantidad de canales se reduce considerablemente, con respecto a los trabajos reportados en la literatura, para la detección de

la señal P300 en el delectreador de Donchin. Esto conlleva a que la colocación de los electrodos en los sujetos de pruebas se pueda realizar fácilmente por medio de la banda elástica mostrada en la figura 2.12. Con base en los resultados de este capítulo, se concluyó que se van a utilizar los electrodos de copa, situados en la parte occipital en los canales O1, Oz y O2. Este esquema se seguirá durante los próximos capítulos de la tesis, cada vez que se hable de la adquisición de señales EEG con P300.

Una vez caracterizado el delectreador y que electrodos utilizar, el siguiente paso es desarrollar los algoritmos para detección de la señal P300 y el carácter.

CAPÍTULO 3. EXTRACCIÓN DE CARACTERÍSTICAS BASADA EN LA AMPLITUD DE LA SEÑAL P300 Y CLASIFICACIÓN MEDIANTE UNA RED *FEEDFORWARD*

Una vez generado el esquema adecuado para obtener las señales P300, se procede a diseñar los algoritmos de procesamiento para la detección de dichas señales. En el caso de esta tesis se proponen dos métodos para realizar la detección, en el primer método se especifican las características con base en la amplitud de las señales P300 y para la clasificación se utiliza una red neuronal *feedforward*. El segundo método utiliza un modelo de aprendizaje profundo. En ambos métodos, se utilizó como medio para evocar la señal el deletreador de Donchin modificado.

En este capítulo, se presenta el primer método de procesamiento que detecta la señal P300 mediante un análisis de la amplitud de la señal y se utiliza una red neuronal para su clasificación. A partir de la detección de la señal P300, se propone un algoritmo para la detección de carácter que el sujeto está observando en el deletreador. Para presentar el método propuesto, este capítulo se estructura de la siguiente manera: en la sección 3.1 se muestra un esquema general del método propuesto, la sección 3.2 da a conocer la etapa de preprocesamiento de las señales EEG y se documenta la manera en la que se extraen las características, la sección 3.4 presenta la red neuronal utilizada para clasificar, y finalmente en la sección 3.5 se muestran los resultados y discusión del método propuesto.

3.1 Esquema general del método propuesto

El método propuesto en este capítulo se denominó *FeedForwad* P300 (FFp300) y como se puede observar en la figura 3.1, se compone de cuatro etapas. La primera etapa es el preprocesamiento de las señales EEG, en la cual se filtran las señales para eliminar ruido. Posteriormente, en la segunda etapa se realiza la extracción de características, mediante el agrupamiento de las señales obtenidas después de cada intensificación. La tercera etapa es una red neuronal artificial cuya entrada es el vector de amplitudes, y la salida de la red indica si se generó la señal P300 en cierta columna o fila^T. Por último, con la información de la salida de la red se indica cual es el carácter correcto.

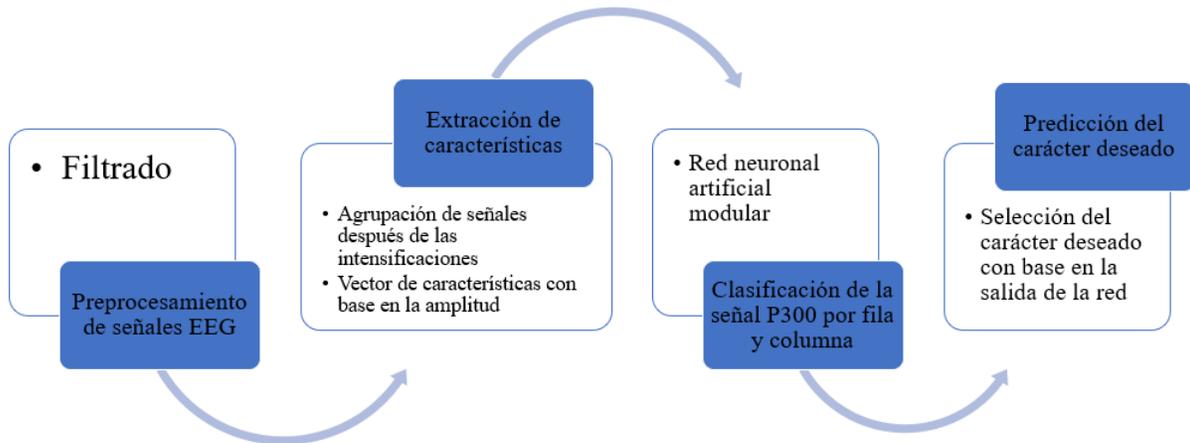


Figura 3.1 Esquema del método FFp300 basado en la extracción de características con base en la amplitud de la señal P300.

3.2 Preprocesamiento y extracción de características

El preprocesamiento de las señales consiste en un filtrado pasa bandas con el mismo modelo descrito en la sección 2.4.2. Es decir, se obtiene la señal $s_{ff}(n)$ mediante la ecuación 2.1, con la diferencia de que ahora se utilizan solo los canales $j=\{O1, Oz, O2\}$ (estos fueron seleccionados en la sección 2.4).

Una vez obtenida $s_{ff}(n)$, se genera el vector de características que alimenta a la red neuronal. La generación del vector de características se divide en dos etapas. La primera corresponde al agrupamiento de las señales obtenidas después de cada intensificación en columnas y filas^T, la cual es mostrada en la sección 3.2.1. La segunda consiste en tomar la amplitud de ciertas muestras de las señales agrupadas para formar las características, la cual se define en la sección 3.2.2.

3.2.1 Agrupamiento de señales por intensificaciones

El agrupamiento de las señales por intensificaciones consiste en tomar la señal $s_{ff}(n)$ y descomponerla por segmentos de 400 ms, que van de los 100 a 500 ms después de cada intensificación. Cada uno de los segmentos de la señal $s_{ff}(n)$ es representado por la señal $x_{ij}(m)$ al igual que se realizó en la sección 2.4.2, pero ahora con un tamaño distinto. Debido a que la frecuencia de muestreo es de 250 Hz, los segmentos de 400 ms contienen un total de 100

muestras, el tamaño fue elegido ya que toda la onda de la señal P300 se puede presentar en ese intervalo de tiempo después del estímulo [16].

A continuación, se propone una reducción dimensional de las señales $x_{ij}(m)$ mediante un submuestreo por un factor de 10. La reducción se realiza con la finalidad de representar a $x_{ij}(m)$ con una cantidad menor de muestras, ya que experimentalmente se encontró que la señal P300 se puede observar aun con esta cantidad de muestras, y esto evita el exceso de procesamiento. Dicha reducción de dimensión de $x_{ij}(m)$ se representa mediante $\lambda_{ij}(w)$, la cual se define en la ecuación 3.1.

$$\lambda_{ij}(w) = x_{ij} \downarrow (m) \quad 1 \leq w \leq 10 \quad 3.1$$

donde λ_{ij} representa a la señal $x_{ij}(m)$ submuestreada en la intensificación i en el canal j , y w es el índice de muestras decimado obtenido de las muestras m . Cada una de las señales $\lambda_{ij}(w)$ se normalizan para cada intensificación por canal para obtener una media cero y varianza unitaria.

$$\mathcal{G}_{ij}(w) = (\lambda_{ij}(w) - \mu_{ij}) / \sigma_{ij} \quad 3.2$$

donde $\mathcal{G}_{ij}(w)$ representa a $\lambda_{ij}(w)$ normalizada, μ_{ij} es la media de las muestras de cada señal $\lambda_{ij}(w)$ y σ_{ij} la desviación estándar. La normalización se realiza con la finalidad de disminuir las alteraciones con grandes amplitudes en la señal $s_{ij}(n)$, provocadas por movimientos de los sujetos de pruebas.

Con el fin de poder observar el comportamiento de las intensificaciones que contienen o no la señal P300, se realiza un análisis visual similar al mostrado en la sección 2.4.3, que consiste en promediar las señales $\mathcal{G}_{ij}(w)$ en las intensificaciones i que deberían contener la señal P300 y otro promedio para las que no. Un proceso similar al que se da en las ecuaciones 2.3 y 2.4 se muestra a continuación en las ecuaciones 3.3 y 3.4 respectivamente.

$$\mathcal{G}_{Tj}(w) = \frac{1}{30} \sum_{i \in CC} \mathcal{G}_{ij}(w) \quad CC : \text{Carácter correcto} \quad 3.3$$

$$\mathcal{G}_{F_j}(w) = \frac{1}{150} \sum_{i \in NC} \mathcal{G}_{i_j}(w) \quad NC : \text{No es carácter correcto} \quad 3.4$$

donde $\mathcal{G}_{T_j}(w)$ es el promedio de las señales $\mathcal{G}_{i_j}(w)$ para i con el carácter correcto y $\mathcal{G}_{F_j}(w)$ es el promedio de las señales $\mathcal{G}_{i_j}(w)$ para i sin el carácter correcto, en cada época de deletreo.

En la figura 3.2 se muestran los promedios de los vectores submuestreados $\mathcal{G}_{T_j}(w)$ y $\mathcal{G}_{F_j}(w)$ como P300 y no P300 respectivamente. Se puede observar que cuando se realiza el submuestreo, aún se sigue manteniendo la forma de la señal P300 en amplitud y latencia. También, se puede observar en la figura 3.2 que solo en ciertos elementos w de $\mathcal{G}_{T_j}(w)$ y $\mathcal{G}_{F_j}(w)$ se puede ver la señal P300. Es decir, los elementos $w \neq 5$ y $w \neq 6$ de los vectores $\mathcal{G}_{T_j}(w)$ y $\mathcal{G}_{F_j}(w)$, tienen valores muy similares en las señales con P300 y No P300, pero aquellos elementos $w=5$ y $w=6$ de los vectores $\mathcal{G}_{T_j}(w)$ y $\mathcal{G}_{F_j}(w)$ presentan diferencia cuando existe P300 y cuando no. Por lo tanto, las amplitudes de los elementos $w=5$ y $w=6$ son los que se tomarán para definir las características que representaran si existe o no P300. Esto se explica en la siguiente sección.

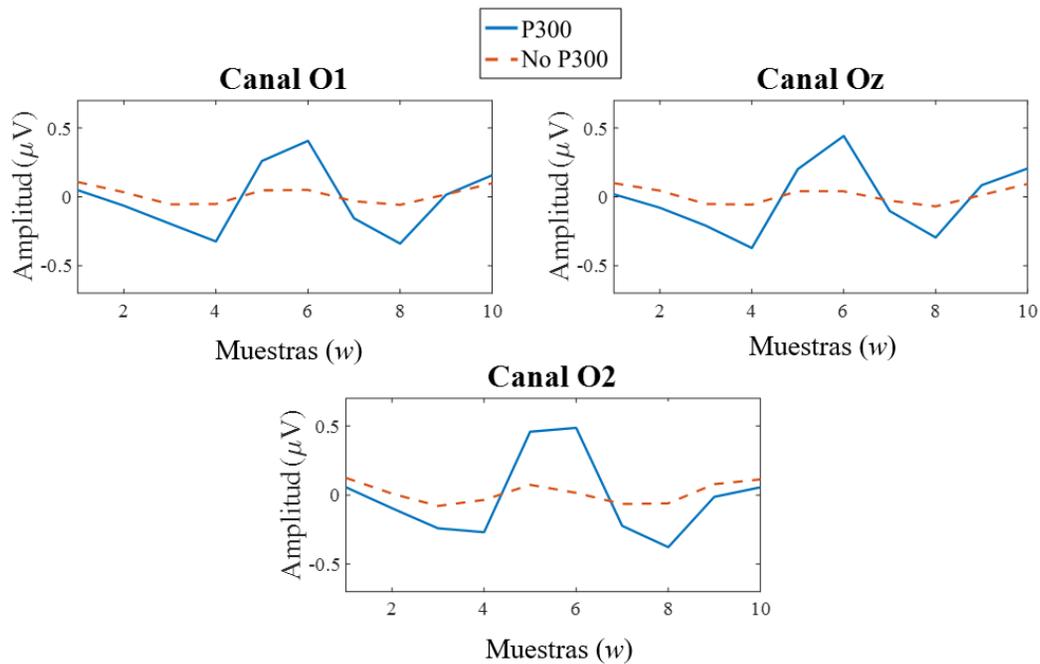


Figura 3.2 Promedio de vectores $\mathcal{G}_{T_j}(w)$ (P300) y $\mathcal{G}_{F_j}(w)$ (No P300), para un sujeto de pruebas en los canales O1, Oz y O2.

3.2.2 Características con base en la amplitud

Como se sabe, la señal P300 tiene una mayor amplitud alrededor de los 300 ms después del estímulo visual, y en caso de no existir un estímulo visual no se generará la señal P300, y solo será ruido aleatorio. Al realizar el promedio de las señales que solo contienen ruido, se obtuvo una respuesta con tendencia a cero. Si observamos la figura 3.3, para $\mathcal{G}_{Tj}(w)$ $w=5$ y $w=6$ se presenta una mayor amplitud que en $\mathcal{G}_{Fj}(w)$. Para las demás muestras diferentes de $w=5$ y $w=6$ no se encontró una diferencia significativa. Debido a que esto es reproducible para los 5 sujetos de pruebas con los que se experimentó en el capítulo anterior, estos valores son tomados en cuenta para formar los vectores de características.

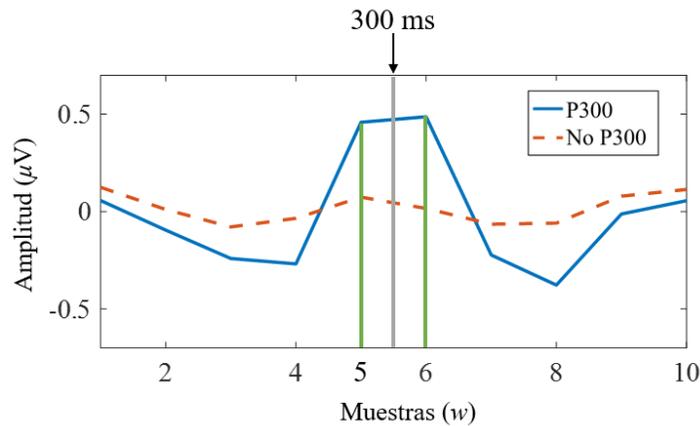


Figura 3.3 Selección de características con base en la amplitud de la señal cerca de los 300 ms.

Por lo tanto, el siguiente paso para generar los vectores de características consiste en promediar cada una de las señales $\mathcal{G}_{ij}(w)$ por columnas y filas^T, de cada época de delecteo. Esto se realiza ya que el método FFp300 consiste en la detección de la señal P300 por columnas y filas^T, no por intensificación independiente como algunos trabajos en la literatura.

Como la matriz del delecteador tiene un total de 36 caracteres y corresponde a una matriz de 6x6, se tiene un total de 12 columnas y filas^T. Para indicar cual columna o fila^T se está seleccionando, se genera un índice con el rango de $1 \leq k \leq 12$, tal como se ve en la figura 3.4.

1	2	3	4	5	6	7	8	9	10	11	12
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
A	B	C	D	E	F	A	G	M	S	Y	5
G	H	I	J	K	L	B	H	N	T	Z	6
M	N	O	P	Q	R	C	I	O	U	1	7
S	T	U	V	W	X	D	J	P	V	2	8
Y	Z	1	2	3	4	E	K	Q	W	3	9
5	6	7	8	9	_	F	L	R	X	4	_

Figura 3.4 Índices k para columnas y filas^T de la matriz del deletreador.

Para la detección de un carácter, se tiene que identificar en que columna y fila^T k se encuentra. Cada época de deletreo tiene un total de 180 intensificaciones, de las cuales son 15 para cada columna y fila^T. Por lo tanto, se realiza el promediado de las señales $\mathcal{G}_{ij}(w)$ para k representado por $\overline{\mathcal{G}_{kj}}(w)$.

$$\overline{\mathcal{G}_{kj}}(w) = \frac{1}{15} \sum_i \mathcal{G}_{ij}(w) \quad 3.5$$

Donde i se agrupa por columnas y filas^T. $1 \leq k \leq 6$ se refiere a las columnas y $7 \leq k \leq 12$ se refiere a las filas^T. De esta manera, $\overline{\mathcal{G}_{kj}}(w)$ contiene los promedios por columnas y filas^T de cada intensificación. Para cada época de deletreo se seleccionan únicamente los datos de $w=5$ y $w=6$, los cuales pueden contener la señal P300. De esta manera, se genera el vector de características de la siguiente forma:

$$F_{kj} = (\overline{\mathcal{G}_{kj}}(5), \overline{\mathcal{G}_{kj}}(6)), \quad 1 \leq k \leq 12, \quad j = \{O1, Oz, O2\} \quad 3.6$$

En la figura 3.5 se muestra el espacio de características para 34 épocas (conjunto intensificaciones para deletrear un carácter) de deletreo de uno de los sujetos en el canal O1, Oz y O2. Por lo tanto, la figura 3.5 muestra 68 vectores de características con P300, ya que por cada época de deletreo existen dos vectores con P300, uno para las columnas y uno para las filas^T.

También se muestran 340 vectores de características sin P300, ya que en cada época existen diez columnas y filas^T que no contienen P300. Se puede observar cómo se forman dos grupos en las características, los cuales son factibles de clasificar por el método que se describe en la siguiente sección.

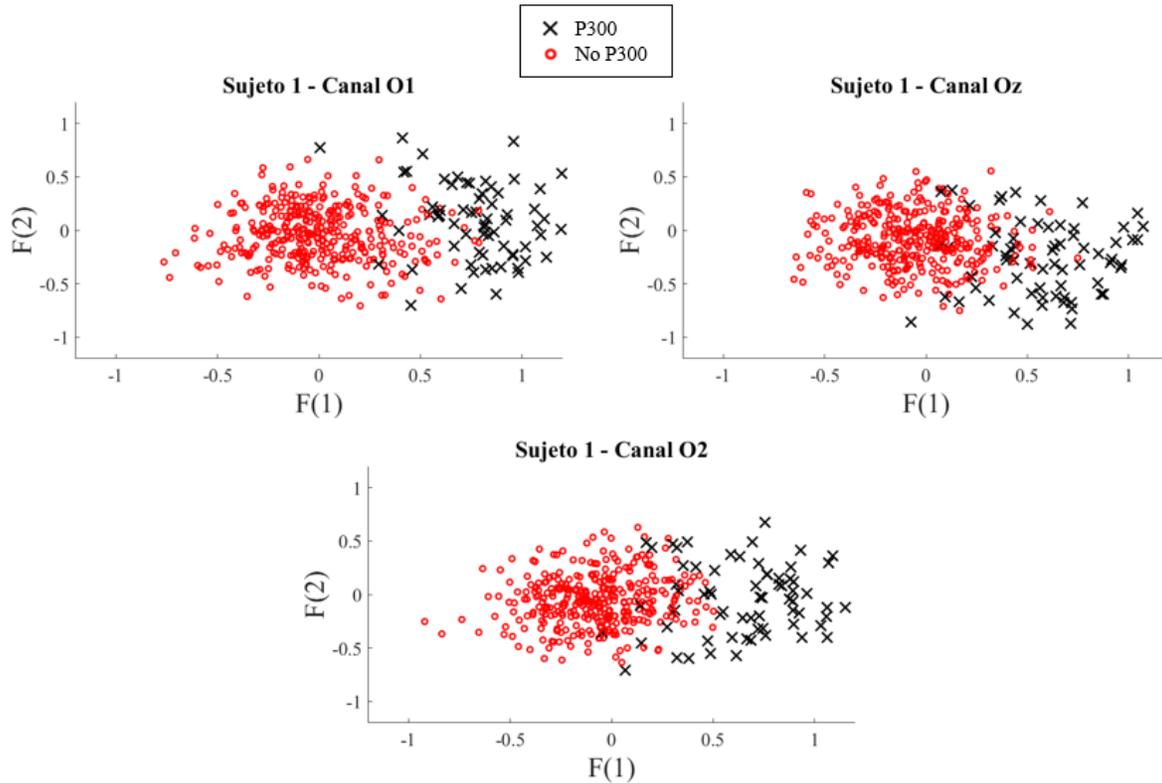


Figura 3.5 Vectores de características en los canales O1, Oz y O2, para uno de los sujetos de pruebas.

3.3 Clasificación

Una vez obtenido el vector de características $F_{k,j}$, el siguiente paso es definir si se tiene o no la señal P300 con un clasificador y posteriormente detectar el carácter. Este clasificador se compone de dos etapas. La primera es la detección de la señal P300 con una red neuronal artificial que predice si cada columna o fila^T contiene la señal P300. La segunda etapa se basa en detectar cual es la columna y fila^T que contiene la señal P300 con mayor magnitud para encontrar es el carácter deseado.

3.3.1 Clasificación para la detección de la señal P300

Para la detección de la señal P300 por columnas y filas^T se propuso en FFp300 un conjunto de redes neuronales artificiales tipo *feedforward* [37]. Este conjunto forma una arquitectura modular como se muestra en la figura 3.6 y consta de dos capas, la primera capa son tres redes, cada una tiene como entrada las características de amplitud de cada canal y la segunda capa es una red que toma como entrada la salida de las redes anteriores y sirve para detectar si se tiene o no el P300. La red tipo *feedforward* fue elegida debido a que en cada canal existe traslape entre clases con incertidumbre, como se ve en la figura 3.5. Es decir, dentro del espacio de características de cada canal, existe una región donde se pueden mezclar las clases P300 y No P300. Esta mezcla de clases genera un traslape que en cada canal es muy diferente y, además, no se sabe si en las siguientes muestras se presentará un traslape similar al anterior. Esto se comprobó mediante un experimento, el cual consistió en un análisis visual de las características en algunas épocas de delecteo, donde, en la mayoría de los casos se generó un traslape muy diferente en cada clase. Por esta razón, se propuso un esquema modular de una red por canal con la finalidad de obtener tres respuestas para cada columna y fila^T indicando si contienen o no la señal P300. Posteriormente, la predicción que se obtiene para cada uno de los tres canales alimenta a otra red neuronal, dando así una salida que combina la información de los tres canales y, de esta manera predecir si la señal P300 se encuentra en cierta columna y fila^T.

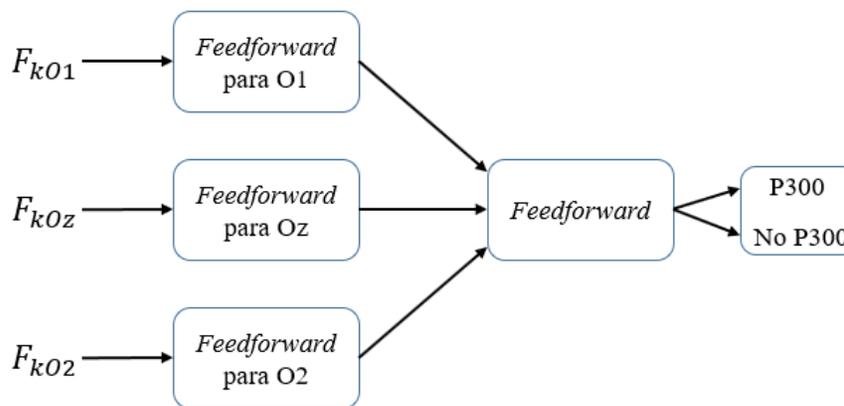


Figura 3.6 Arquitectura del método FFp300.

Cada una de las tres redes neuronales en la primera capa de FFp300 se alimenta con el vector de características F_{kj} , como se muestra en la figura 3.6. La capa oculta de dichas redes tiene un total de cinco neuronas con su respectivo factor de polarización. Esta cantidad de neuronas fue elegida tratando de evitar el sobreajuste mediante una buena generalización de los datos, pero sin caer en la sobre-generalización. El experimento para definir el número de neuronas en la capa oculta consistió en colocar un número aleatorio de neuronas en la red, el cual fue 10, y se analizaron los resultados obtenidos en el entrenamiento. Posteriormente, se fue reduciendo el número de neuronas en la capa oculta hasta que existiera un cambio significativo en la precisión del entrenamiento, siendo cinco el número mínimo de neuronas que no provocó un cambio significativo. En la capa de salida se cuenta con solo una neurona, debido a que se trata de una clasificación binaria, donde la salida de la red indica si cada columna y fila^T contiene o no a la señal P300. En la figura 3.7 se muestra la topología de la red para cada uno de los canales.

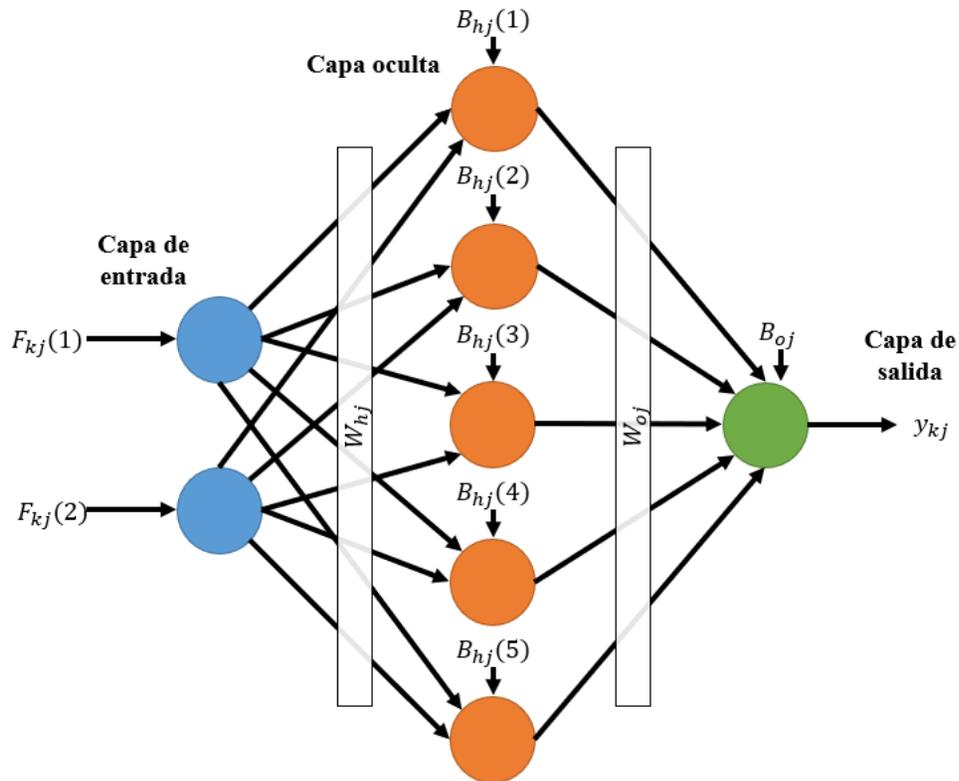


Figura 3.7 Topología de la red neuronal artificial utilizada para la clasificación de las características en un canal.

La actividad en la capa oculta para la red mostrada en la figura 3.7 es representada por:

$$V_{kj} = F_{kj} \cdot W_{hj} + B_{hj} \quad 3.7$$

donde W_{hj} es el vector de pesos de la red en la capa oculta y B_{hj} es el vector de pesos para la polarización en la capa oculta, siendo h el subíndice para indicar que pertenecen a la capa oculta. Cada una de las neuronas en la capa oculta utilizan una función de activación sigmoideal $\Phi(V_{kj})$, la cual se muestra en la ecuación 3.8.

$$\Phi(V_{kj}) = \frac{1}{1 + e^{-V_{kj}}} \quad 3.8$$

Entonces, la salida para cada una de las redes por canal se expresa como:

$$y_{kj} = W_{oj} \cdot \Phi(V_{kj}) + B_{oj} \quad 3.9$$

donde W_{oj} es el vector de pesos de la red en la capa de salida y B_{oj} es el valor del peso de la polarización en la capa de salida, con el subíndice o para indicar que se trata de la capa de salida. Una vez que se calcula la salida de las tres redes neuronales artificiales para los canales O1, Oz y O2, se crea el vector:

$$Y_k = \{y_{kO1}, y_{kOz}, y_{kO2}\} \quad 3.10$$

El vector Y_k se utiliza como entrada de la red de la segunda capa, la cual combina la información de los tres canales, para posteriormente dar como salida la predicción de la red neuronal para encontrar la señal P300 en cierta columna o fila^T. La red utilizada para combinar la información se muestra en la figura 3.8. La cual tiene la misma topología de la red que se utiliza por canal, debido a que se diseñó de la misma manera, con la diferencia de que tiene tres entradas.

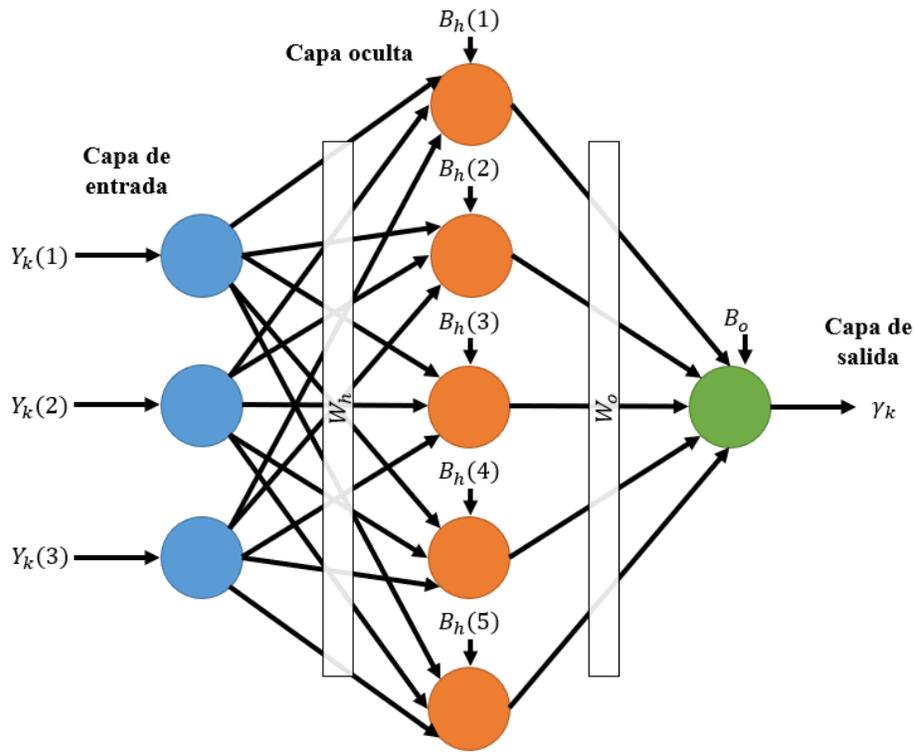


Figura 3.8 Topología de la red neuronal artificial utilizada para la combinación de la información que generan las redes neuronales por canal.

Una vez calculada la salida final de la red neuronal modular para una época de delectreo, se obtiene un vector de salida, el cual contiene la predicción para cada una de las columnas y filas^T de contener la señal P300 de dicha época, la salida es representada por:

$$\gamma_k = \{\gamma_1, \gamma_2, \dots, \gamma_{12}\} \quad 3.11$$

donde γ_k representa la salida de la red neuronal modular para las columnas y filas^T.

El entrenamiento de las redes se realiza por medio del algoritmo *backpropagation* de gradiente conjugado escalado [38]. Este algoritmo es una variación de los algoritmos de gradiente conjugado, en el cual, el valor del parámetro de aprendizaje cambia durante cada iteración para realizar más rápido la convergencia del algoritmo. Los algoritmos de gradiente conjugado normalmente modifican el gradiente de una forma lineal. No obstante, en el caso del algoritmo de gradiente conjugado escalado, la dirección en la que se busca el gradiente va

cambiando durante cada iteración mediante el cálculo de un vector de dirección, como se muestra a continuación. Este algoritmo sigue el esquema básico de los métodos de gradiente conjugado, donde la actualización de los pesos W , está dada por:

$$W(a+1) = W(a) + \eta(a)p(a) \quad 3.12$$

donde $\eta(a)$ es el parámetro de aprendizaje y $p(a)$ es el vector de dirección en la iteración a . El vector de dirección inicial es el negativo del vector gradiente cuando $a=0$. Por lo tanto:

$$p(0) = -g(0) \quad 3.13$$

donde $g(a)$ es el vector gradiente y cada uno de los vectores de dirección posteriores son calculados por medio de:

$$p(a+1) = -g(a+1) + \beta(a)p(a) \quad 3.14$$

donde $\beta(a)$ es un parámetro que varía mediante la ecuación de *Fletcher-Reeves* [39]. El cual utiliza un número máximo de iteraciones de 1000. Se detiene si no logra reducir el error a través de 6 iteraciones consecutivas. El parámetro inicial que determina el cambio en los pesos es de 5×10^{-5} y trata de llegar a un rendimiento de 0 para el cambio del valor del gradiente.

3.3.2 Predicción del carácter deseado

Una vez que la red neuronal de la segunda capa detecta si existe o no el P300 en una época de delecteo, es necesario predecir cuál es el carácter que el sujeto desea delectear. La predicción del carácter correcto se basa en encontrar la columna y fila^T con el valor máximo predicho por la red, ya que en esas columnas y filas^T el P300 se debe generar con mayor amplitud. La columna (C) predicha por la red está dada por el máximo argumento de γ_k para $1 \leq k \leq 6$, y la fila^T (R) por el máximo argumento de γ_k para $7 \leq k \leq 12$. Esto puede ser expresado como:

$$C = \operatorname{argmax}(\gamma_k) \quad 1 \leq k \leq 6 \quad 3.15$$

$$R = \operatorname{argmax}(\gamma_k) \quad 7 \leq k \leq 12 \quad 3.16$$

Ya que se conoce dónde está el valor máximo de salida de la red para las columnas y para las filas^T se toma la decisión de la siguiente manera. El valor máximo en la salida de la red que tiene de la columna C y fila^T R , se evalúa a través de una función de umbral u , para indicar si la magnitud con la que se detectó la señal P300 es lo suficientemente alta para ser considerada como P300, y no por la actividad del ruido. Este umbral evita errores de detección del carácter cuando se esté realizando una sesión de deletreo, y el sujeto de pruebas no esté poniendo atención en la matriz. Es decir, se establece un umbral basado en la detección del P300 para asegurarse de que los valores máximos de columnas y filas^T representen el P300 y no ruido. Si $\gamma(C)$ o $\gamma(R)$ no sobrepasan el umbral u , la respuesta del método será que no se pudo predecir el carácter correcto. El valor de u se colocó en 0.8, ya que tras un análisis de las columnas y filas^T con P300 que fueron correctamente predichas, sobrepasaban este valor. Esto se puede definir como:

$$L = \begin{cases} M(C, R) & \text{si } u < \gamma(C) \ \& \ u < \gamma(R) \\ \text{Indefinido} & \text{de otra forma} \end{cases} \quad 3.17$$

donde L es el carácter predicho, y M es la matriz con los caracteres definida como:

$$M = \begin{array}{|c|} \hline A \ B \ C \ D \ E \ F \\ \hline G \ H \ I \ J \ K \ L \\ \hline M \ N \ O \ P \ Q \ R \\ \hline S \ T \ U \ V \ W \ X \\ \hline Y \ Z \ 1 \ 2 \ 3 \ 4 \\ \hline 5 \ 6 \ 7 \ 8 \ 9 \ _ \\ \hline \end{array} \quad 3.18$$

3.4 Resultados y discusión

En esta sección se muestran los resultados y análisis del método FFp300 utilizando en el deletreador de Donchin modificado. Los resultados se basaron en medir el desempeño de FFp300 para la detección de la señal P300 y en la detección del carácter. En la sección 3.5.1 se

encuentra el análisis para la detección de la señal P300 por columnas y filas^T, mientras que en la sección 3.5.2 se muestra el análisis de la predicción del carácter correcto.

3.4.1 Análisis de resultados en detección por filas^T y columnas

Para realizar las pruebas, participaron 10 sujetos, los cuales deletrearon un total de 20 caracteres para el entrenamiento y 14 caracteres para la validación. La cantidad de muestras se definió de tal manera que fuera la cantidad máxima sin que los sujetos de pruebas se agotaran. Esto fue con la finalidad de que el cansancio o sueño en las personas no causara irregularidades en las muestras.

Debido a que se deletrearon 20 caracteres para el entrenamiento, cada uno con 12 vectores de características, seis para las columnas y seis para las filas^T, el entrenamiento de la red utiliza un total de 240 vectores para cada uno de los sujetos de pruebas. Algunos vectores representan la señal P300 y otros no, los que representan P300 son 40, 20 para las columnas y 20 para las filas^T, por lo tanto 200 vectores no deben representar la señal P300. Para la validación se utilizaron un total de 168 vectores de características (fueron deletreados 14 caracteres), de los cuales 14 vectores para las columnas y 14 para las filas^T son los que representan la señal P300, por lo tanto 140 vectores no deben representarla. La diferencia que existe entre la cantidad de vectores con y sin P300, se debe a que, en el deletreador por cada intensificación que evoca la señal P300 existen cinco intensificaciones que no lo hacen.

En la tabla 3.1 se muestra la clasificación del método FFp300 indicando si cada una de las columnas y filas^T contiene o no la señal P300 por época de deletreo. Estos resultados son mostrados tanto en el entrenamiento como en la validación para cada uno de los sujetos de pruebas. Los resultados son obtenidos mediante las matrices de confusión, donde se muestra la cantidad (#) y porcentaje de verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN). Debido a que se tiene conocimiento del carácter que las personas deseaban deletrear, los VP representan aquellas columnas y filas^T k donde se detectó la señal P300 y si la debían contener, los VN para k donde no se detectó la señal P300 y no la deberían contener. Por otro lado, los FP representan las columnas y filas^T k donde se detectó como señal P300 y no la contenían y para los FN es donde no se detectó la señal P300 y si se

debería haber detectado. Además, en la última fila de la tabla 3.1, se muestra el valor ideal para cada una de las evaluaciones, representando si en la clasificación no existieran errores, con la finalidad de poder tener una referencia de cuál es el valor que se hubiera deseado obtener.

Tabla 3.1 Resultados de las matrices de confusión para FFp300 en entrenamiento y validación.

Sujeto	Entrenamiento								Validación							
	VP		VN		FP		FN		VP		VN		FP		FN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	33	13.8	191	79.6	9	3.8	7	2.9	22	13.1	137	81.5	3	1.8	6	3.6
2	25	10.4	193	80.4	7	2.9	15	6.3	18	10.7	136	81.0	4	2.4	10	6.0
3	29	12.1	191	79.6	9	3.8	11	4.6	25	14.9	136	81.0	4	2.4	3	1.8
4	33	13.8	198	82.5	2	0.8	7	2.9	24	14.3	139	82.7	1	0.6	4	2.4
5	31	12.9	199	82.9	1	0.4	9	3.8	22	13.1	140	83.3	0	0.0	6	3.6
6	33	13.8	197	82.1	3	1.3	7	2.9	25	14.9	139	82.7	1	0.6	3	1.8
7	24	10.0	194	80.8	6	2.5	16	6.7	17	10.1	134	79.8	6	3.6	11	6.5
8	38	15.8	199	82.9	1	0.4	2	0.8	25	14.9	140	83.3	0	0.0	3	1.8
9	15	6.3	194	80.8	6	2.5	25	10.4	14	8.3	137	81.5	3	1.8	14	8.3
10	15	6.3	192	80.0	8	3.3	25	10.4	14	8.3	133	79.2	7	4.2	14	8.3
Ideal	40	16.7	200	83.3	0	0.0	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0

En la tabla 3.1 se puede observar como la detección de la señal P300 por columnas y filas^T para cada uno de los sujetos de pruebas no se da de la misma, es decir, los resultados varían dependiendo de los sujetos. Los sujetos 1, 4, 5, 6 y 8 tienen resultados cerca del valor ideal, sin embargo, los demás sujetos tienen resultados más bajos. Tras observar los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ de cada sujeto de pruebas, se concluyó que el desempeño de FFp300 está directamente relacionado al tiempo de respuesta y amplitud de las señales P300. Es decir, el bajo desempeño para los sujetos 2, 3, 7, 9 y 10, se da debido a que su señal P300 no se encuentra cercana a los 300 ms o tiene una amplitud similar a la del ruido, ya que las características se basan en que la señal P300 se encuentra en $w=5$ y $w=6$.

Una de las razones por la que la señal P300 se adelanta es porque probablemente los sujetos tienen reflejos rápidos. Esto significa lo siguiente: en [15], se dice que la señal P300 puede presentarse en diferentes tiempos después del estímulo, y que este tiempo depende de la capacidad que tienen las personas de percibirlos. Esta capacidad está relacionada al tipo de

actividades diarias que realiza cada uno de los sujetos de pruebas. Es decir, aquellos sujetos con altos reflejos que practican arte marcial, deportes o videojuegos, generan su señal P300 alrededor de los 200 ms. De esta manera, se decidió preguntarles a los sujetos 2, 3, 7, 9 y 10 si hacían una actividad similar y efectivamente, algunos de estos sujetos practicaban artes marciales o jugaban videojuegos, lo cual ocasionó que las muestras $w=5$ y $w=6$ no contuvieran información que pudiera discriminar de una manera eficiente aquellas señales que no contienen P300 de las que sí.

Adicional a la información proporcionada por las matrices de confusión de la tabla 3.1, se realizó también un análisis basado en las métricas de reconocimiento, precisión, sensibilidad y *F-measure* con la finalidad de poder observar detalladamente el desempeño de FFp300, al detectar la señal P300 por columnas y filas^T. Las métricas utilizadas son las siguientes:

- Reconocimiento: indica la tasa de muestras que se reconocen correctamente como clase P300 y No P300. Se calcula de la siguiente manera:

$$\text{reconocimiento} = \frac{VP + VN}{VP + VN + FP + FN} \quad 3.19$$

- Precisión: indica la tasa de muestras con señal P300 que se detectaron correctamente, en relación con todas aquellas muestras que se detectaron con señal P300. Esta métrica se calcula como se muestra en la siguiente ecuación:

$$\text{precisión} = \frac{VP}{VP + FP} \quad 3.20$$

- Sensibilidad: indica la tasa de detecciones correctas de la señal P300, en relación con todas las muestras que contienen la señal P300. La cual se calcula de la siguiente manera:

$$\text{sensibilidad} = \frac{VP}{VP + FN} \quad 3.21$$

- *F-measure*: es la media armónica que combina los valores de la precisión y la sensibilidad. La cual es una medida única tomando en cuenta estos dos valores, y se calcula de la siguiente manera:

$$F - measure = 2 \frac{\text{sensibilidad} \cdot \text{precisión}}{\text{precisión} + \text{sensibilidad}} \quad 3.22$$

A continuación, en la tabla 3.2 se muestran los resultados obtenidos de las métricas de reconocimiento, precisión, sensibilidad y *F-measure* para cada uno de los sujetos mediante el método FFp300.

Tabla 3.2 Resultados de las métricas de evaluación.

Sujeto	Entrenamiento				Validación			
	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>
1	0.93	0.79	0.83	0.80	0.95	0.88	0.79	0.83
2	0.91	0.78	0.63	0.69	0.92	0.82	0.64	0.72
3	0.92	0.76	0.73	0.74	0.96	0.86	0.89	0.88
4	0.96	0.94	0.83	0.88	0.97	0.96	0.86	0.91
5	0.96	0.97	0.78	0.86	0.96	1.00	0.79	0.88
6	0.96	0.92	0.83	0.87	0.98	0.96	0.89	0.93
7	0.91	0.80	0.60	0.69	0.90	0.74	0.61	0.67
8	0.99	0.97	0.95	0.96	0.98	1.00	0.89	0.94
9	0.87	0.71	0.38	0.49	0.90	0.82	0.50	0.62
10	0.86	0.65	0.38	0.48	0.88	0.67	0.50	0.57
Promedio	0.93	0.83	0.69	0.75	0.94	0.87	0.74	0.79

En lo referente a los resultados promedio de la tabla 3.2, la tasa de reconocimiento para el entrenamiento y validación es de 0.93 y 0.94 respectivamente. Con estos resultados se puede observar que la red neuronal artificial tiene una buena generalización de los datos por cada uno de los sujetos, debido a que los resultados para el entrenamiento y la validación son muy similares. Si los resultados en el entrenamiento hubieran sido más altos que los de la validación, se podría asumir que la red tiene un sobreajuste.

En la métrica de precisión, se tiene un 0.83 y 0.87 para entrenamiento y validación respectivamente. Esto se relaciona directamente a la cantidad de falsos positivos. Sin embargo, una mayor precisión no significa que se tenga un mejor desempeño. Por ejemplo, se puede observar que el sujeto 5 y 8 tienen una precisión de 1, esto quiere decir que no existió ningún falso positivo por medio de FFp300. Sin embargo, los verdaderos positivos pueden ser pocos, lo cual se ve reflejado en la métrica de sensibilidad, la cual tiene valores de 0.79 y 0.89 para el sujeto 5 y 8 respectivamente. Para el sujeto 6 se tiene una precisión de 0.96, sin embargo, tiene la misma tasa de reconocimiento que el sujeto 8 y mejor que el del sujeto 5, esto es debido a que la sensibilidad del sujeto 6 es 0.89, mayor al del sujeto 5.

Se puede observar que las métricas de precisión y sensibilidad ayudan a la medición del desempeño del algoritmo, pero resulta complicado medir el desempeño cuando se analizan por separado. En el caso de la precisión se puede tener valor de uno, si solo se detecta una vez una señal P300 y ésta es correcta (ver ecuación 3.20). En el caso de la sensibilidad, se puede tener un algoritmo con valor de uno, si en todas las muestras se detecta que hay P300 (ver ecuación 3.21). Por lo tanto, es recomendable analizar estas métricas en conjunto, lo cual se puede lograr mediante la métrica *F-measure*, como esta métrica calcula la media armónica entre la precisión y la sensibilidad, tiene valor de uno si precisión y sensibilidad son uno. Entre más cercano a uno sea *F-measure*, mejor es el desempeño en el algoritmo. De acuerdo con la tabla 3.2, los resultados promedio de *F-measure* en FFp300 con los diez sujetos de pruebas es de 0.75 en entrenamiento y 0.79 en validación. El máximo valor obtenido para validación se da en el sujeto 8 con un valor de 0.94, seguido del sujeto 6 con un valor de 0.93, y el más bajo desempeño es en el sujeto 10 con 0.57.

En lo referente a los sujetos con desempeños más altos se puede comentar lo siguiente: como se observa en la tabla 3.2, los sujetos 6 y 8 tuvieron el mismo resultado en la métrica de reconocimiento. Sin embargo, la métrica *F-measure* muestra un mejor desempeño en el sujeto 8 que en el sujeto 6. Los motivos por los cuales el sujeto 8 tiene mejor desempeño, es que la amplitud de la señal es mayor como se observa en la figura 3.9 y la señal P300 se encuentra muy cercana a los 300 ms. Para el caso del sujeto 6, su P300 se encuentra cercana a los 300 ms como

se muestra en la figura 3.10, pero la amplitud de la señal P300 es menor que en el caso del sujeto 8.

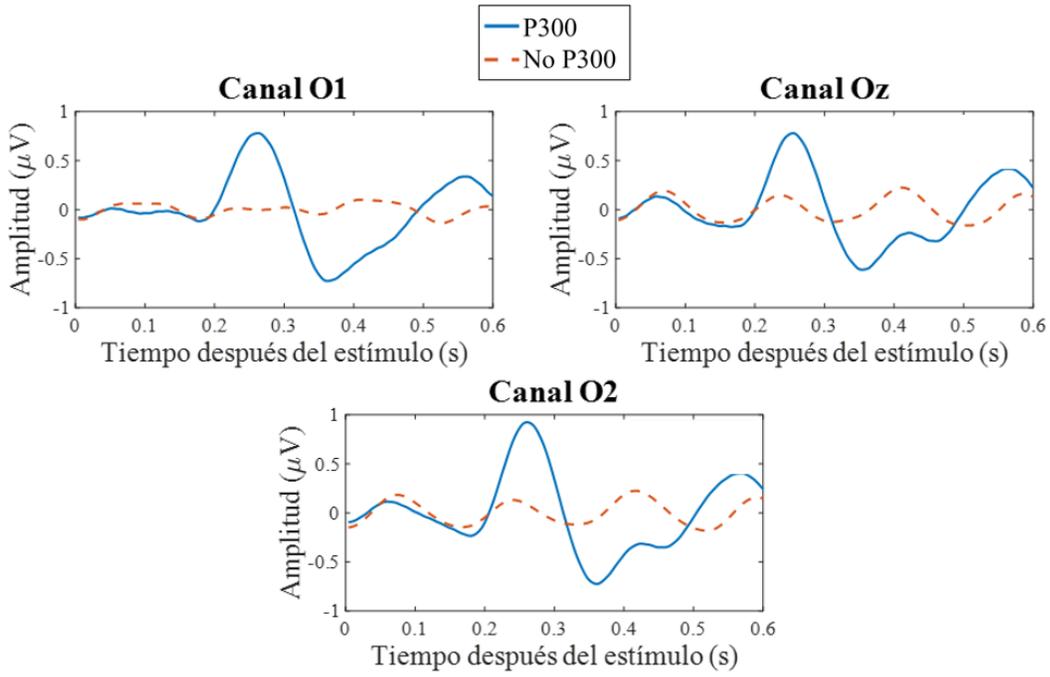


Figura 3.9 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 8.

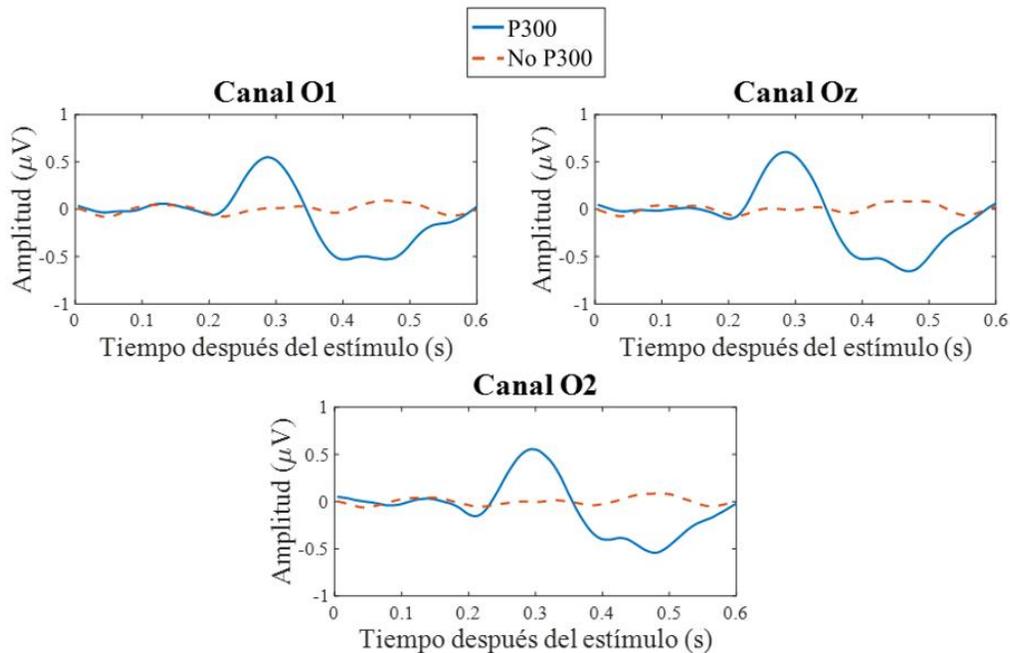


Figura 3.10 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 6.

La tasa de reconocimiento más alta obtenida en validación es de 0.98 para los sujetos 6 y 8, y la más baja es de 0.88 para el sujeto 10. Esta variación que existe entre los resultados motivó un análisis visual de las señales obtenidas con un sujeto que generó la más alta tasa de reconocimiento y otro la más baja, los cuales son los sujetos 6 y 10 respectivamente. En la figura 3.10 donde se muestran los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ para el sujeto 6, donde se puede observar que es muy notoria la presencia de la señal P300 ya que el promedio de los vectores $x_{Tj}(m)$ tiene una mayor amplitud que el ruido en el promedio de $x_{Fj}(m)$. Además, la amplitud máxima de la señal P300 se encuentra muy cercana a los 300 ms. Ambos factores, la amplitud de la señal y el tiempo en que se presenta la señal P300 son los factores más determinantes para la correcta detección en el método FFp300.

Para el caso del sujeto 10, la señal P300 se genera correctamente, pero se adelanta. En la figura 3.11 se pueden observar los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ de dicho sujeto. Se puede apreciar visualmente que la diferencia entre la señal con P300 y las que no tienen P300 es bastante notoria, tal como el sujeto 6. Sin embargo, la señal se encuentra un poco adelantada, ya que ocurre un instante después de los 200 ms. Para indagar a que se debió que se adelantará la señal P300, se le preguntó al sujeto que actividades diarias realizaba, y respondió que practicaba karate y juega videojuegos. Esto se correlaciona con la teoría de que las personas que realizan actividades que generan altos reflejos, la señal P300 se adelanta. Aunado a esto, se puede observar en la figura 3.11 que las señales que no contienen la señal P300 tienen mayor actividad que las mostradas en el sujeto 6.

En conclusión, debido a que la señal P300 se genera con diferencias en latencia y amplitud para cada uno de los sujetos de pruebas [40], se realiza entrenamiento del método FFp300 por sujeto, ya que este método no es capaz de generalizar las señales P300 para los diez sujetos de pruebas utilizados.

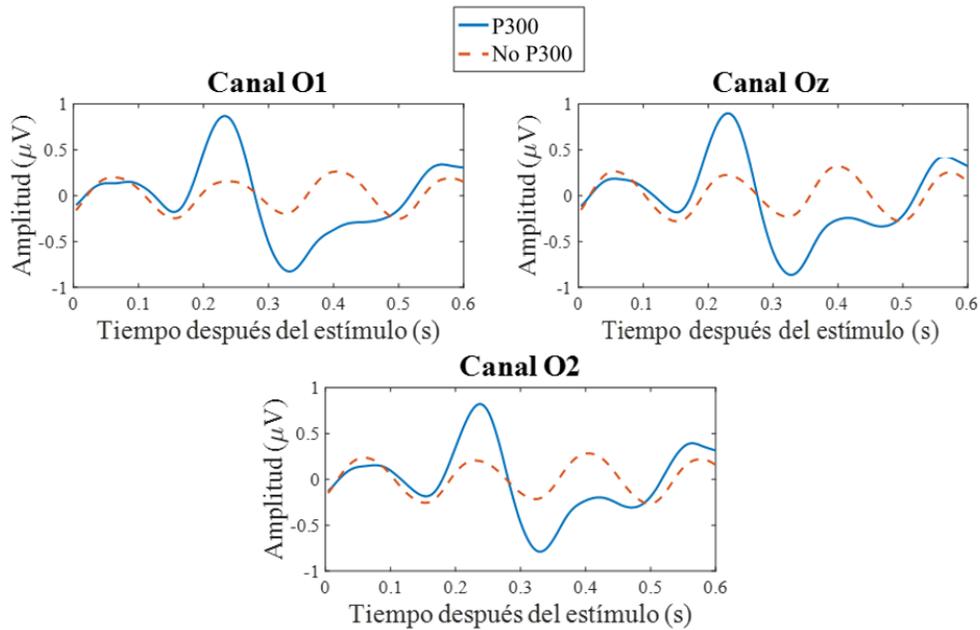


Figura 3.11 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), para el sujeto 10.

Como se mencionó anteriormente, el tiempo de ocurrencia de la señal P300 puede depender de las actividades diarias de las personas, y la amplitud varía en función de diversos factores. Para indagar porque varía la amplitud, se observaron los promedios de los vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ de todos los sujetos de pruebas, donde se encontró una correlación entre la estatura de los sujetos y la amplitud de la señal P300. Como en la mayoría de los sujetos se cumple que a mayor estatura, mayor es la amplitud de la señal P300, se tiene la hipótesis que la estatura de las personas es uno de los muchos factores que puede influir en la amplitud de esta señal, hablando de adultos entre 20 y 25 años.

Un comportamiento de FFp300 que causó inquietud fue que todas las métricas mostraban resultados mejores en la validación que en entrenamiento. Esto puede ser causado debido a que la mayoría de los sujetos de pruebas no había utilizado un sistema BCI enfocado a deletreadores. Por lo tanto, como las primeras muestras fueron tomadas para el entrenamiento, se puede deducir que también entrenaron a los sujetos de pruebas, y al momento de tomar las muestras de validación, éstos ya tenían experiencia de cómo realizar la sesión.

3.4.2 Análisis de resultados en la predicción del carácter

Para la validación en la predicción del carácter, se realizaron 14 épocas de deletreo, con el fin de deletrear la frase “TEC DE CHIHUAHUA”. Se eligió una frase en vez de utilizar caracteres aleatorios para simular un entorno de aplicación real, similar a cuando una persona trataría de comunicarse por medio del deletreador, de la misma manera que se realizó para generar la base de datos II de BCI Competition III [25]. Debido a que las señales son adquiridas por medio del deletreador de Donchin modificado, la matriz del deletreador es transpuesta a la mitad de la época de deletreo, por lo tanto, las señales P300 son adquiridas desde diversas posiciones en la matriz.

En la tabla 3.3 se muestran las detecciones para los caracteres deletreados por cada uno de los sujetos de pruebas y el porcentaje de aciertos resultantes de FFp300. Los caracteres mal predichos por FFp300 se encuentran resaltados. En la primera fila de la tabla, se muestran los caracteres que los sujetos intentaron deletrear.

Tabla 3.3 Caracteres deletreados por cada uno de los sujetos y porcentaje de aciertos.

Caracteres															
Frase	T	E	C	D	E	C	H	I	H	U	A	H	U	A	% de aciertos
Sujeto 1	T	E	C	D	E	C	6	O	H	U	A	H	U	A	85.71%
Sujeto 2	T	E	7	D	9	C	H	I	6	U	A	6	U	A	71.42%
Sujeto 3	Z	E	C	D	E	E	N	I	H	U	A	H	W	A	71.42%
Sujeto 4	T	F	C	D	E	C	H	I	H	U	A	H	U	A	92.85%
Sujeto 5	T	E	C	D	E	C	H	I	H	U	B	H	U	A	92.85%
Sujeto 6	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 7	T	E	F	D	E	C	6	I	H	U	A	N	O	E	64.28%
Sujeto 8	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 9	X	E	C	D	J	C	N	K	L	U	A	H	U	A	64.28%
Sujeto 10	T	D	F	C	4	C	P	I	H	Z	A	H	U	C	50%
Promedio															79.28%

En el caso de la detección de carácter, se tiene que identificar la columna y filas^T donde se encuentra éste, si solo se predice correctamente una de ellas, la clasificación del carácter será errónea. Por lo tanto, para encontrar el carácter correcto se tiene que decidir dos veces entre seis clases, lo que sería igual a una vez en 36 clases. De esta manera, el porcentaje de reconocimiento para los caracteres es de 79.28% a diferencia del reconocimiento por columnas y filas^T donde se obtuvo una tasa de 0.94, equivalente al 94% (mostrado en la tabla 3.2), ambos en validación.

Entre los 10 sujetos de pruebas, se deletrearon un total de 140 caracteres, de los cuales 29 caracteres no fueron predichos correctamente. En 25 de estos caracteres, se logró predecir correctamente solo la columna o fila^T donde se encontraba. Es decir, que el 86.2% de las veces que el algoritmo falló, fue al predecir solo la columna o fila^T, pero no las dos. Es por esto que el porcentaje de reconocimiento promedio decayó un 13.72% de la detección de columnas y filas^T a la detección del carácter.

Sin embargo, para el método FFp300 el porcentaje promedio en la detección del carácter decae en contra del porcentaje de reconocimiento de la señal P300 en esta base de datos, es posible que, con otras muestras, dicho porcentaje suba. Por ejemplo, algunos sujetos como el sujeto 6 y 8 obtuvieron el 100% en el reconocimiento de los caracteres de acuerdo con los resultados mostrados en la tabla 3.3, aun cuando no se obtuvo el 100% de predicción correcta por columnas y filas^T. Es posible obtener el 100% de reconocimiento en la predicción del carácter deseado, sin tener un reconocimiento perfecto de las columnas y filas^T. Esto es debido a que pueden existir casos particulares en los cuales la señal P300 se detecte en dos o más columnas o filas^T. Cuando la señal P300 se detecta dos o más veces por columna o filas^T, la decisión se toma por aquella que tenga la mayor magnitud en la salida de la red. Un ejemplo podría ser cuando en una época de deletreo se tienen dos columnas donde se detectó la señal P300, una detección corresponde a un verdadero positivo, y la otra a un falso positivo, si el valor de la columna en la salida de la red para el verdadero positivo es mayor que la del falso positivo, la columna donde se encuentra el carácter será predicha correctamente y viceversa. Por lo tanto, los falsos positivos en la detección de las señales P300 impactan indirectamente en la predicción del carácter correcto, de tal manera que los porcentajes de reconocimiento de carácter son diferentes a los porcentajes de reconocimiento de la señal P300.

El tiempo promedio que dura FFp300 para predecir cada carácter es de 146 ms, lo cual es suficientemente rápido para una aplicación en línea, ya que el procesamiento de los datos se realiza una vez que han sido recolectadas todas las muestras de la época de delectreo. FFp300 fue implementado en Matlab[®] por medio del paquete de herramientas de redes neuronales.

CAPÍTULO 4. EXTRACCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE LA SEÑAL P300 MEDIANTE UNA RED NEURONAL CONVOLUCIONAL

En este capítulo se presenta el método basado en aprendizaje profundo para la detección de la señal P300. El potencial se evoca con el delectador de Donchin modificado y como medio de adquisición se utiliza OpenBCI[®]. La extracción de características y clasificación se realizan con una red neuronal convolucional (CNN por sus siglas en inglés), la cual es un método popular en la literatura para la implementación del aprendizaje profundo. A diferencia del método propuesto en el capítulo anterior, las características no fueron especificadas, ya que el modelo basado en CNN se encarga de generarlas automáticamente utilizando solo los datos crudos preprocesados.

Para describir el método propuesto, este capítulo tiene la siguiente estructura. En la sección 4.1 se muestra el esquema general. Las secciones 4.2 y 4.3 presentan la etapa de preprocesamiento y organización de las señales que alimentarán a la CNN. La sección 4.4 documenta la arquitectura y funcionamiento de la CNN. En la sección 4.5 se presenta la detección del carácter en función de la salida de la CNN. Finalmente, la sección 4.6 presenta los resultados obtenidos y discusión.

4.1 Esquema general del método propuesto

El método propuesto en este capítulo se denomina detección de P300 con CNN (CNN300) y se compone de tres etapas principales. La primera etapa consiste en el preprocesamiento de las señales EEG por medio de un filtro pasa bandas y normalización. La segunda etapa consiste en acondicionar las señales filtradas, agrupándolas por cada una de las intensificaciones generadas. La tercera etapa es la red neuronal convolucional, ésta es la que se encarga de la extracción de características y clasificación para detectar si las señales contienen o no P300. Finalmente, la detección de la señal P300 se utiliza para la predicción del carácter deseado.

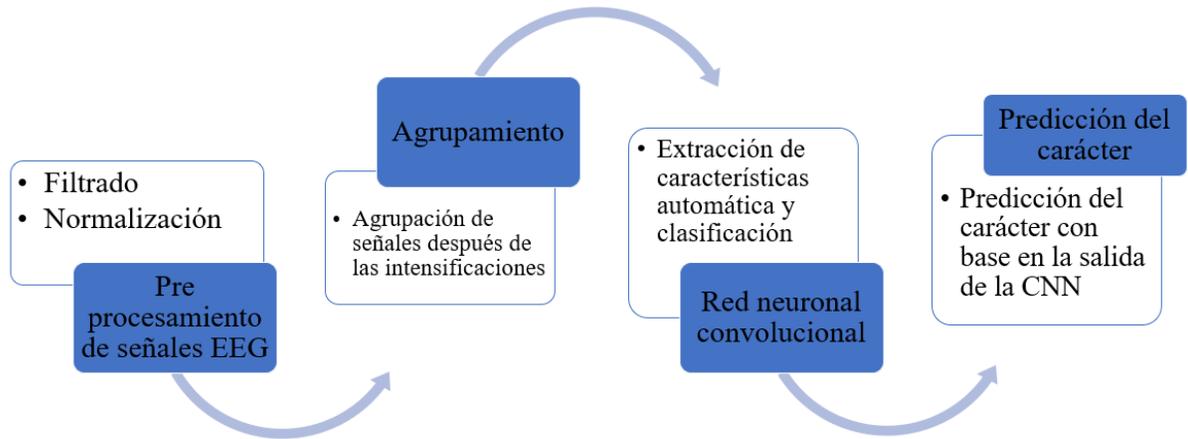


Figura 4.1 Esquema del método para la detección de la señal P300 basado en una CNN.

4.2 Preprocesamiento de las señales EEG

El preprocesamiento consiste en un filtrado pasa bandas y normalización por cada uno de los canales de las señales EEG. El filtro utilizado es el mismo que se define en la ecuación 2.1 (sección 2.4.2), obteniendo la señal $s_{ff}(n)$ para los canales $j=\{O1, Oz, O2\}$. De esta manera, se eliminan frecuencias que son consideradas ruido y se conservan los datos de la señal P300. Una vez que los datos son filtrados, se procede a realizar una normalización para obtener media cero y varianza unitaria en cada uno de los canales. La normalización se realiza con la finalidad de que todos los canales se encuentren dentro de la misma escala en amplitud. La normalización de la señal $s_{ff}(n)$ se define en la ecuación 4.1.

$$\varphi_{ff}(n) = (s_{ff}(n) - \mu_j) / \sigma_j \quad 4.1$$

donde $\varphi_{ff}(n)$ es la señal normalizada, μ_j es la media de los datos filtrados cada canal y σ_j la desviación estándar.

4.3 Agrupamiento de las señales por intensificación

En esta etapa se realiza la descomposición de la señal $\varphi_{ff}(n)$ por intensificaciones en segmentos de 400 ms. Al igual que en la sección 3.2.1, los segmentos se generan de los 100 a 500 ms después de la intensificación, y se representan por la señal $\chi_{ij}(m)$. Debido a que la

frecuencia de muestreo es de 250Hz, los 400 ms se representan con cien datos. Posteriormente, se realiza un promediado de las señales $\chi_{ij}(m)$ por columnas y filas^T k , de esta manera se obtiene la señal $\overline{\chi_{kj}}(m)$, la cual se define mediante la ecuación 4.2.

$$\overline{\chi_{kj}}(m) = \frac{1}{15} \sum_i \chi_{ij}(m) \quad 4.2$$

Debido a que las redes neuronales convolucionales se utilizan principalmente en el área de procesamiento digital de imágenes [41], la entrada de la red tiene una estructura similar a la de una imagen digital, para este caso es como una imagen en escala de grises, es decir, $\chi_{ij}(m)$ se dimensiona como un arreglo de datos de $J \times S$, donde $J=3$ indica el número de canales utilizados, y $S=100$ se refiere a los segmentos de 400 ms. Por lo tanto, para cada uno de los k segmentos de $\overline{\chi_{kj}}(m)$ se forma una matriz $I_k(j,m)$ donde $j=1, \dots, J$, $m=1, \dots, S$ y $k=1, \dots, 12$. En la figura 4.2a se muestra gráficamente la manera en la que las señales EEG son acondicionadas para tener el mismo formato que una imagen y en la figura 4.2b se muestra una representación gráfica de la señal $I_k(j,m)$ ya acondicionada.

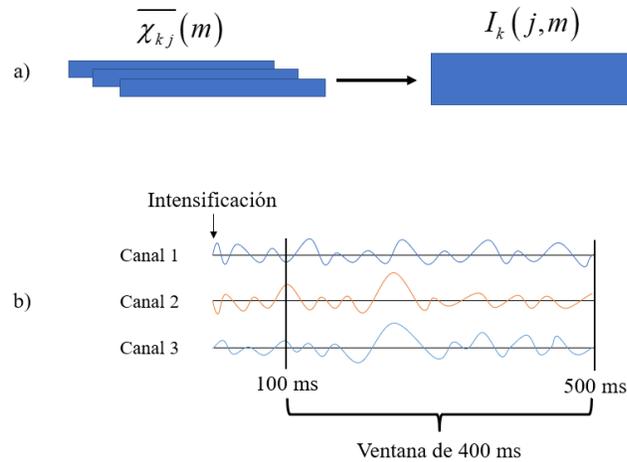


Figura 4.2 a) Acondicionamiento de las señales EEG a formato de imagen, b) Representación gráfica de $I_k(j,m)$.

De esta manera, $I_k(j,m)$ contiene las señales $\overline{\chi_{kj}}(m)$ agrupadas para cada una de las columnas y filas^T k en todos los canales j , y corresponde a la entrada de la red neuronal convolucional.

4.4 Red neuronal convolucional

Las redes neuronales convolucionales son un tipo de red neuronal artificial para el aprendizaje profundo [42]. Uno de los usos más comunes de las CNN es la extracción automática de características, es decir, que no se necesita especificar manualmente las características para la clasificación de la información. La arquitectura de las CNN es definida por medio de múltiples capas de filtros de convolución, de una o más dimensiones, y capas totalmente conectadas. Cuando son utilizadas para clasificación, la fase de extracción de características se encuentra en las primeras capas de la red. Estas están compuestas por capas convolucionales seguidas de algún método de rectificación lineal. Conforme los datos de entrada son propagados por la fase de extracción de características, las neuronas son menos sensibles a los cambios en la información que alimenta a la red, pero cada vez analizan características más complejas. Finalmente, se llega a la etapa de clasificación, donde las características extraídas por medio de las capas convolucionales alimentan a una red neuronal artificial de tipo *feedforward*, para etiquetar cada uno de los patrones de entrada. Para ejemplificar cómo funcionan las CNN, en la figura 4.3 se muestra una red que contiene dos capas de convolución para extracción de características y sus mapeos por cada filtro de convolución. Además, también se pueden ver dos capas de una red neuronal *feedforward* para la clasificación.

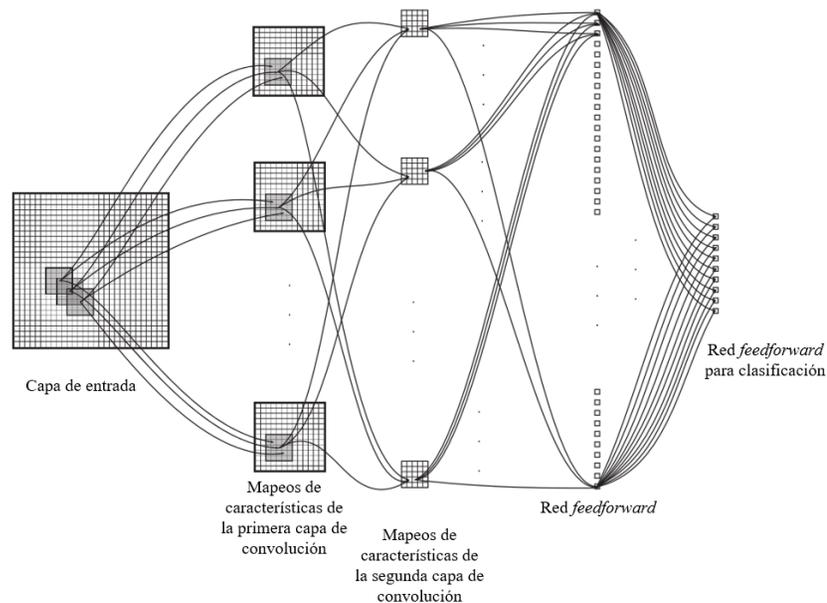


Figura 4.3 ejemplo de una CNN utilizada para extracción de características y clasificación [42].

El uso de las CNN es muy reciente y la mayoría de la investigación se enfoca a los algoritmos de visión por computadora. En lo referente a procesamiento de señales EEG, los resultados han sido muy prometedores, pero es poco el trabajo realizado y aún no se tienen métodos concretos para el diseño [42]. Es por esto que algunos de los parámetros del método propuesto en este capítulo son definidos empíricamente.

De esta manera, la CNN para el método CNN300 tiene cinco etapas y se muestra en la figura 4.4. La primera etapa es la entrada de la red neuronal que corresponde a $I_k(j,m)$. La segunda y tercera etapa consisten en los procesos de convolución y rectificación lineal, y realizan la extracción automática de características. En la cuarta etapa se clasifican las características extraídas y se compone de tres secciones: la red completamente conectada, una normalización exponencial, y la clasificación. La quinta y última etapa es la salida de la red neuronal donde se indica si la entrada contiene o no la señal P300. Cada una de las capas de la CNN son representadas por $0 \leq l \leq 7$.

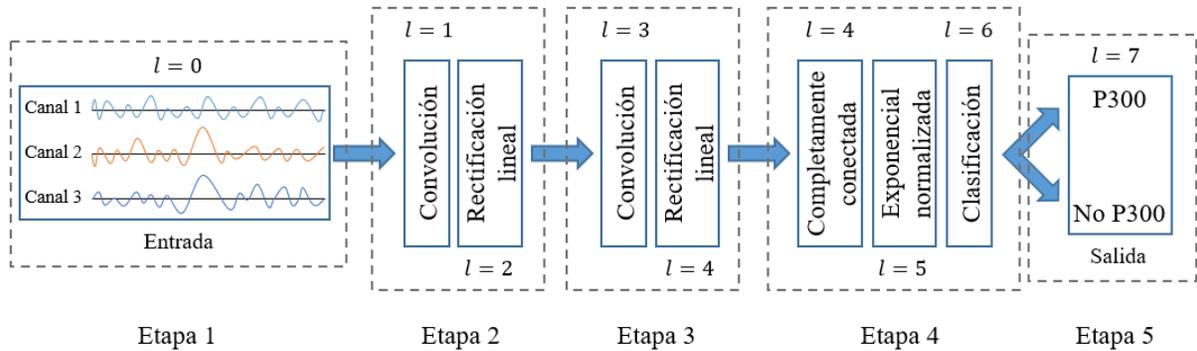


Figura 4.4 Esquema general de la red neuronal convolucional utilizada por CNN300.

4.4.1 Etapas de convolución y rectificación

Las etapas 2 y 3 de la CNN propuesta para CNN300 tienen dos capas, una de convolución y otra de rectificación. Cada capa de convolución se define de la siguiente manera [43]:

$$C_{l_f} = Q_{l-1} * W_{l_f} + B_{l_f} \quad l = \{1,3\} \quad 4.3$$

donde C_{lf} representa a los mapeos en la salida de la capa l de convolución para los filtros f , Q_{l-1} corresponde a la salida de la capa anterior donde $Q_0=I_k(j,m)$, W_{lf} son los pesos del filtro y B_{lf} la polarización.

Cada una de las capas de convolución modifica el tamaño de su respectiva entrada, generando mapeos de características para cada uno de los filtros utilizados. Los tamaños de los mapeos de características en las dimensiones a están dados por:

$$Y_{Ma} = (Y_{Ea} - Y_{Fa} + 2 \cdot Padding) / Stride + 1 \quad 4.4$$

donde Y_{Ma} representa el tamaño del mapeo, Y_{Ea} el tamaño de la entrada, Y_{Fa} el tamaño del filtro en la dimensión a , $Padding$ se refiere al relleno de ceros, y $Stride$ al número de saltos verticales y horizontales que dan los filtros cuando se realiza el proceso de convolución. En esta red el parámetro $Padding$ se colocó en cero y el parámetro $Stride$ se colocó en uno, debido a que en la clasificación de las señales EGG con P300 solo existen dos clases y no fue necesario modificarlos, en comparación al trabajar con clasificación de objetos en imágenes donde existe una mayor cantidad de clases y se requiere de manejar estos parámetros para poder extraer características de una manera diferente [17]. Por lo tanto, el tamaño de los mapeos Y_{Ma} para la CNN utilizada se puede definir mediante la siguiente ecuación:

$$Y_{Ma} = Y_{Ea} - Y_{Fa} + 1 \quad 4.5$$

Como se observa en la figura 4.4, cada capa de convolución tiene una capa de rectificación lineal para reducir la complejidad computacional del algoritmo, ya que se ha comprobado que éstas ayudan a que el entrenamiento se realice de una manera más rápida, sin afectar el desempeño de la red [14]. La salida de la capa de rectificación puede considerarse como una función de umbral, donde, todos los elementos menores que cero se colocan en cero. Esta función está representada por:

$$C_{lf}(d) = \begin{cases} d, & d \geq 0 \\ 0, & d < 0 \end{cases} \quad 4.6$$

donde d representa cada uno de los elementos de los mapeos obtenidos por las capas de convolución.

Las etapas de convolución y rectificación se utilizan para la extracción de características. La primera etapa extrae características por cada uno de los canales j de la entrada, sin combinar la información de éstos. La segunda etapa combina la información de los tres canales j utilizados.

Para esto, la primera capa de convolución $l=1$ tiene 30 filtros de tamaño 1×70 . Debido a que la entrada $I_k(j,m)$ es de 3×100 , se obtendrán 30 mapeos de tamaño Y_{Ma} 3×31 de acuerdo a la ecuación 4.5. Estas dimensiones se explican a continuación. La cantidad de 30 filtros fue especificada con base en obtener una buena generalización de los datos sin caer en el sobre ajuste. Las dimensiones de 1×70 de cada filtro, se especificaron de tal manera que en la primera capa de convolución no exista correlación entre los canales j de la entrada, y reducir los datos de 100 a 31. De esta manera, se generan características que reducen dimensionalmente la entrada de 3×100 a 3×31 , sin combinar la información entre canales. El tamaño en la segunda dimensión del filtro, que reduce la entrada de 100 a 31 fue definido para generalizar los patrones de entrada disminuyendo las dimensiones de $I_k(j,m)$. Luego de la capa convolucional, los mapeos se someten a la capa de rectificación $l=2$ definida en la ecuación 4.6.

La segunda capa de convolución C_{3f} combina la información de los canales sin modificar las características extraídas anteriormente. Por lo tanto, esta capa de convolución utiliza 30 filtros de 3×1 , donde la entrada de esta capa son los mapeos rectificados de la etapa anterior con un tamaño Y_{Ma} de 3×31 . En consecuencia, la salida de ésta serán mapeos de tamaño Y_{Ma} 1×31 . La salida de esta segunda capa de convolución también se somete a una capa de rectificación lineal $l=4$, representada por la ecuación 4.6.

4.4.2 Capa completamente conectada

Esta capa es la encargada de tomar la salida C_{4f} de la última capa de rectificación para decidir cuándo se tiene o no la señal P300. El número de neuronas utilizadas en esta capa es de 930, igual al tamaño de la salida de la capa precedente, donde, para la entrada de 1×31 se generaron

30 mapeos. Para esta capa se tiene un vector de pesos W_l , y la polarización B_l para cada una de las neuronas. En la ecuación 4.7 se expresa la salida T_k de esta capa.

$$T_k = Q_{l-1} \cdot W_l + B_l \quad l=5 \quad 4.7$$

donde T_k contiene valores numéricos en la salida de la capa completamente conectada, y estos datos son utilizados para alimentar la capa de exponencial normalizada.

4.4.3 Capa exponencial normalizada

Esta capa se conoce también como *softmax* y se refiere a la función de activación encargada de calcular en términos de probabilidad a que clase pertenece la entrada $I_k(j,m)$ de la red. El cálculo de la probabilidad se realiza de la siguiente manera:

$$P(c_b | T_k, \theta) = \frac{P(T_k, \theta | c_b) P(c_b)}{\sum_{r=1}^2 P(T_k, \theta | c_r) P(c_r)} \quad 4.8$$

donde $P(T_k, \theta | c_b)$ es la probabilidad condicional de la muestra dada la clase c_b , $P(c_b)$ es la probabilidad de la clase a priori [14] y θ corresponde a los parámetros W de los filtros y la polarización B , r indica que clase b que se está calculando. En este caso se tienen dos clases: No P300 y P300, por lo cual, $b=\{0,1\}$. Estas reglas de probabilidad cumplen con las reglas básicas

donde $0 \leq P(c_b | T_k, \theta) \leq 1$ y $\sum_{r=1}^2 P(c_r | T_k, \theta) = 1$.

4.4.4 Capa de clasificación

En esta capa se genera un vector que contiene las probabilidades de cada columna y fila^T de pertenecer a la clase $b=1$, dados los patrones de entrada $I_k(j,m)$. El vector de probabilidad para cada columna y fila^T por época de deletreo, se representa como:

$$P(c_1 | T_k, \theta) = \{P(c_1 | T_1, \theta), P(c_1 | T_2, \theta), \dots, P(c_1 | T_{12}, \theta)\} \quad 4.9$$

La salida de la red también se utiliza para calcular el error de clasificación, que es utilizado para entrenar la red. En este caso se utiliza la función de entropía cruzada para el cálculo del error mostrada en la ecuación 4.10.

$$E(\theta) = -\sum_q \sum_b t_{qb} \ln P_q(c_b | T, \theta) \quad 4.10$$

donde $0 \leq t_{qb} \leq 1$ indica que la q -ésima muestra pertenece a la b -ésima clase, y $P(c_b | T, \theta)$ es la probabilidad de que la red asocie a la q -ésima entrada con la clase b . El conjunto de muestras q está conformado por todas las entradas $I_k(j, m)$.

4.4.5 Entrenamiento de la red

El algoritmo utilizado para el entrenamiento de la red neuronal es el gradiente estocástico descendiente con momento [14], el cual calcula el gradiente sobre la función de error y se define como $\nabla E(\theta_a)$. Este algoritmo trata de minimizar el error mediante la actualización de los pesos en los filtros W_{lf} y polarización B_{lf} . El algoritmo para la actualización de los pesos se puede definir como:

$$\theta_{a+1} = \theta_a - \alpha \nabla E(\theta_a) + \psi (\theta_a - \theta_{a-1}) \quad 4.11$$

donde $\alpha=0.01$ es el factor de aprendizaje. Para evitar las oscilaciones drásticas en el descenso del gradiente se colocó el parámetro $\psi=0.9$, el cual es llamado momento, éste determina la contribución del gradiente anterior a la iteración actual. Estos valores fueron definidos experimentalmente hasta que se encontró una rápida convergencia en el entrenamiento de la red neuronal convolucional sin causar inestabilidad.

Para evitar el sobre entrenamiento de la red neuronal se utiliza el método de regularización L2 definido en [43], el cual actúa directamente sobre el error reduciendo la importancia de los parámetros θ , esto ayuda a que se realice una mejor generalización en el entrenamiento de la red, la función de regularización se puede definir como:

$$E_R(\theta) = E(\theta) + \delta \Omega(W) \quad 4.12$$

donde $\delta=0.0001$ es el parámetro de regularización el cual fue seleccionado experimentalmente, W es el vector de pesos, y la función de regularización utilizada es $\Omega(W)$, la cual está definida por [37]:

$$\Omega(W) = \frac{1}{2} W^T W \quad 4.13$$

Durante el entrenamiento el gradiente es calculado con subconjuntos de 128 muestras, y se realiza un total de 500 épocas, ya que se encontró que con una cantidad mayor de épocas se obtenían los mismos resultados. Los pesos son inicializados usando distribución Gaussiana con media cero y desviación estándar de 0.01, la polarización es inicializada en cero. Con el fin de evitar caer en ciclos, los datos son ordenados aleatoriamente antes de comenzar a entrenarse.

El entrenamiento se realizó con los mismos datos que la red neuronal en el capítulo 3, de las cuales se utilizan 240 muestras para el entrenamiento y 168 para la validación.

4.5 Predicción del carácter deseado

Una vez obtenida la salida de la red neuronal convolucional para una época de deletreo se genera el vector $P(c_I/T_k, \theta)$ de la ecuación 4.9, el cual indica la probabilidad de cada columna y fila^T de contener la señal P300. Este vector $P(c_I/T_k, \theta)$ es utilizado para predecir el carácter deseado de la misma manera que se mostró en la sección 3.3.2, pero ahora en términos de probabilidad, donde la columna C y fila^T R se definen de la siguiente manera:

$$C = \operatorname{argmax}(P(c_1 | T_k, \theta)) \quad 1 \leq k \leq 6 \quad 4.14$$

$$R = \operatorname{argmax}(P(c_1 | T_k, \theta)) \quad 7 \leq k \leq 12 \quad 4.15$$

Posteriormente se pasa a través de una función de umbral $u=0.8$ similar a la descrita en la ecuación 3.17, la cual se muestra en la siguiente ecuación:

$$L = \begin{cases} M(C, R) & \text{si } u < P(c_1 | T_C, \theta) \text{ \& } u < P(c_1 | T_R, \theta) \\ \text{Indefinido} & \text{de otra forma} \end{cases} \quad 4.16$$

Y de esta manera definir cuál es el carácter predicho L con base en la matriz M mostrada en la ecuación 3.18.

4.6 Resultados y discusión

Los resultados del método CNN300 se midieron en la detección de la señal P300 y el carácter al igual que en la sección 3.4. Se utilizó la misma base de datos que en el capítulo 3, la cual se generó por medio del deletreador de Donchin modificado con 10 sujetos de pruebas. Cada sujeto deletreó un total de 34 caracteres, 20 para el entrenamiento y 14 para la validación. Por lo tanto, se tienen 240 muestras para el entrenamiento, de las cuales solo 40 contienen la señal P300 y 168 muestras para la validación, donde 28 muestras contienen la señal P300. Los análisis de estos resultados se muestran en dos partes: la sección 4.6.1 se muestran los resultados de la red neuronal convolucional para la detección de la señal P300 por columnas y filas^T y en la sección 4.6.2 se encuentran los resultados de la predicción del carácter correcto.

4.6.1 Resultados por filas^T y columnas

A continuación, en la tabla 4.1 se muestran los resultados obtenidos para la detección de la señal P300 por columnas y filas^T, con base en las matrices de confusión del método CNN300. Se muestra la cantidad (#) y porcentaje de verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN) para entrenamiento y validación, como se realizó en la sección 3.4.1. En los resultados de las matrices de confusión mostrados, se puede observar como la detección de la señal P300 para las columnas y filas^T es similar para todos los sujetos de pruebas. Esto contrasta con los resultados obtenidos con el método FFp300, en el cual existía una mayor variación entre los resultados para cada uno de los sujetos, debido a que las señales tienen diferencias en amplitud y latencia. Por lo tanto, se asume que el algoritmo CNN300, se adapta mejor a las diferentes señales de cada uno de los sujetos de pruebas que FFp300. En CNN300 el tiempo de ocurrencia de la señal P300 no es determinante para el

desempeño del método, ya que la CNN es alimentada por segmentos de tiempo lo suficientemente grandes para contener la señal P300.

Tabla 4.1 Resultados de las matrices de confusión mediante CNN300 en entrenamiento y validación.

Sujeto	Entrenamiento								Validación							
	VP		VN		FP		FN		VP		VN		FP		FN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	38	15.8	199	82.9	1	0.4	2	0.8	23	13.7	139	82.7	1	0.6	5	3.0
2	35	14.6	198	82.5	2	0.8	5	2.1	24	14.3	139	82.7	1	0.6	4	2.4
3	37	15.4	192	80.0	8	3.3	3	1.3	24	14.3	135	80.4	5	3.0	4	2.4
4	38	15.8	200	83.3	0	0.0	2	0.8	26	15.5	140	83.3	0	0.0	2	1.2
5	36	15.0	200	83.3	0	0.0	4	1.7	25	14.9	139	82.7	1	0.6	3	1.8
6	40	16.7	197	82.1	3	1.3	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0
7	35	14.6	193	80.4	7	2.9	5	2.1	23	13.7	131	78.0	9	5.4	5	3.0
8	39	16.3	200	83.3	0	0.0	1	0.4	27	16.1	140	83.3	0	0.0	1	0.6
9	37	15.4	195	81.3	5	2.1	3	1.3	25	14.9	132	78.6	8	4.8	3	1.8
10	39	16.3	198	82.5	2	0.8	1	0.4	24	14.3	133	79.2	7	4.2	4	2.4
Ideal	40	16.7	200	83.3	0	0.0	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0

A continuación, al igual que en los resultados mostrados en el capítulo 3 también se tienen los resultados de CNN300 con las métricas de reconocimiento, precisión, sensibilidad y *F-measure*. Estos resultados son mostrados en la tabla 4.2, donde también se muestra el promedio de los diez sujetos de pruebas que conforman la base de datos.

De acuerdo con la tabla 4.2, la tasa de reconocimiento promedio para los diez sujetos de pruebas es de 0.98 en el entrenamiento y de 0.96 en la validación, las cuales son similares, y de esta manera, se puede deducir que la red está generalizando de manera adecuada la información que se encuentra en la base de datos. Si los resultados en el entrenamiento superaran con una cantidad considerable a los de validación, la generalización de los datos no sería buena, y se inferiría que el algoritmo está memorizando las muestras de entrenamiento. La tasa de reconocimiento más alta obtenida en el entrenamiento fue de 1 para el sujeto 8, y para la validación fue de 1 en el sujeto 6. Estos resultados se correlacionan con los del método FFp300, en el cual los sujetos 6 y 8 muestran los más altos desempeños.

Tabla 4.2 Resultados con base en las métricas de evaluación para CNN300.

Sujeto	Entrenamiento				Validación			
	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>
1	0.99	0.97	0.95	0.96	0.96	0.96	0.82	0.88
2	0.97	0.95	0.88	0.91	0.97	0.96	0.86	0.91
3	0.95	0.82	0.93	0.87	0.95	0.83	0.86	0.84
4	0.99	1.00	0.95	0.97	0.99	1.00	0.93	0.96
5	0.98	1.00	0.90	0.95	0.98	0.96	0.89	0.93
6	0.99	0.93	1.00	0.96	1.00	1.00	1.00	1.00
7	0.95	0.83	0.88	0.85	0.92	0.72	0.82	0.77
8	1.00	1.00	0.98	0.99	0.99	1.00	0.96	0.98
9	0.97	0.88	0.93	0.90	0.93	0.76	0.89	0.82
10	0.99	0.95	0.98	0.96	0.93	0.77	0.86	0.81
Promedio	0.98	0.93	0.94	0.93	0.96	0.90	0.89	0.89

Para FFp300, la extracción de características se realiza por medio de la amplitud de la señal cercana a los 300 ms. Sin embargo, se encontró que en algunos sujetos de pruebas la señal P300 se encontraba un poco adelantada, como lo fue con el sujeto 10, cuyos promedios de vectores $x_{Tj}(m)$ y $x_{Fj}(m)$ se muestran en la figura 3.11. El que la señal P300 se encontrara un poco adelantada o atrasada a los 300 ms fue uno de los problemas principales para el método FFp300. Sin embargo, ese no es un problema para el método CNN300, y esto se puede observar en que ahora el sujeto 10 tiene una tasa de reconocimiento de 0.93.

Algo que se puede observar en los resultados obtenidos por CNN300, fue que todos los resultados mostrados por las métricas de la tabla 4.2 tienen mejor desempeño con las muestras de entrenamiento, que con las de validación. Para CNN300, *F-measure* tuvo valores de 0.93 y 0.89 para el entrenamiento y validación respectivamente. Esto es un comportamiento natural en las redes neuronales artificiales, debido a que estas son las muestras con las que se diseña el algoritmo.

En conclusión, CNN300 es capaz de generalizar de manera adecuada los patrones de entrada. Lo cual genera que los resultados sean reproducibles entre los sujetos de pruebas que

participaron en el experimento, incluso cuando las señales P300 se generen con desfases en el tiempo.

4.6.2 Resultados para la predicción del carácter

En la tabla 4.3 se muestran los resultados de CNN300 para la predicción del carácter deseado, y, además se muestra el porcentaje de aciertos para cada uno de los sujetos de pruebas. Los caracteres que están resaltados son los que se predijeron incorrectamente por CNN300.

Tabla 4.3 Caracteres deletreados por cada uno de los sujetos y porcentaje de aciertos.

Caracteres															
Frase	T	E	C	D	E	C	H	I	H	U	A	H	U	A	% de aciertos
Sujeto 1	T	E	C	D	E	C	H	I	H	7	A	H	U	A	92.85%
Sujeto 2	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 3	T	9	C	D	K	C	H	I	H	U	A	H	U	A	85.71%
Sujeto 4	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 5	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 6	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 7	H	E	C	D	3	C	A	I	H	U	A	H	U	A	78.57%
Sujeto 8	T	E	C	D	E	C	H	I	H	U	A	H	U	A	100%
Sujeto 9	X	E	C	D	E	C	H	I	H	U	A	U	S	A	78.57%
Sujeto 10	T	E	C	D	E	C	H	I	H	U	A	H	O	A	92.85%
Promedio															92.85%

Como se mencionó anteriormente, para el reconocimiento de un carácter, es necesario detectar en que columna y en que fila^T se encuentra. Por lo tanto, la probabilidad de encontrar el carácter correcto es menor que el detectar la señal P300 por columnas y filas^T. Para el método CNN300 se obtuvo un 92.85% de reconocimiento en la predicción del carácter deseado. En este caso, cinco de los sujetos de pruebas obtuvieron un 100% (sujetos 2, 4, 5, 6 y 8) en el reconocimiento del carácter, y el porcentaje más bajo fue de 78.57% (sujetos 7 y 9).

Para analizar los casos en donde la detección de carácter fue de 100%, se presenta la figura 4.5. Para los sujetos 2, 4, 5, 6 y 8, se muestran los promedios de los vectores $x_{T_j}(m)$ y $x_{F_j}(m)$ en

la figura 4.5a, 4.5b, 4.5c, 4.5d y 4.5e respectivamente. Las señales de estos cinco sujetos de pruebas cumplen con un patrón en específico, en el cual aquellas señales que no contienen P300, tienen una amplitud menor a aquellas que si contienen P300. Los sujetos 2 y 5 son los que tienen una menor amplitud en la señal P300, sin embargo, las señales que no contienen P300 también tienen menor amplitud.

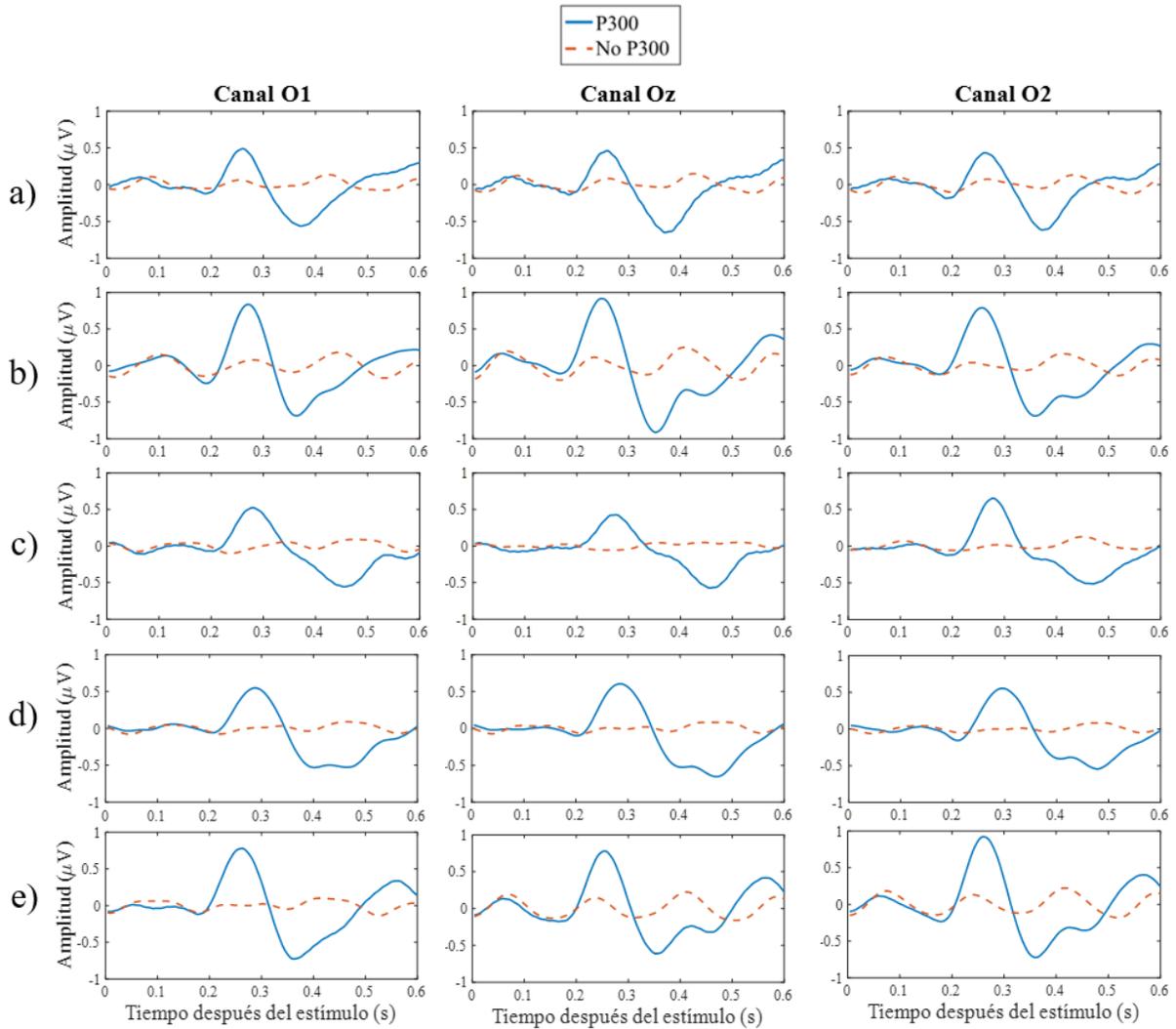


Figura 4.5 Promedio de vectores $x_{Tj}(m)$ (P300) y $x_{Fj}(m)$ (No P300), a) sujeto 2, b) sujeto 4, c) sujeto 5, d) sujeto 6, e) sujeto 8.

El caso con el desempeño más bajo fue el sujeto 7; para ver que sucedió, se muestra las señales obtenidas en la figura 4.6. Se puede observar que, la amplitud de la señal P300, no tiene

una diferencia significativa contra aquellas señales que no contienen la P300. Esto es causado debido a la fisiología de cada uno de los sujetos de pruebas, ya que no en todas las personas se puede detectar la señal P300 con la misma calidad.

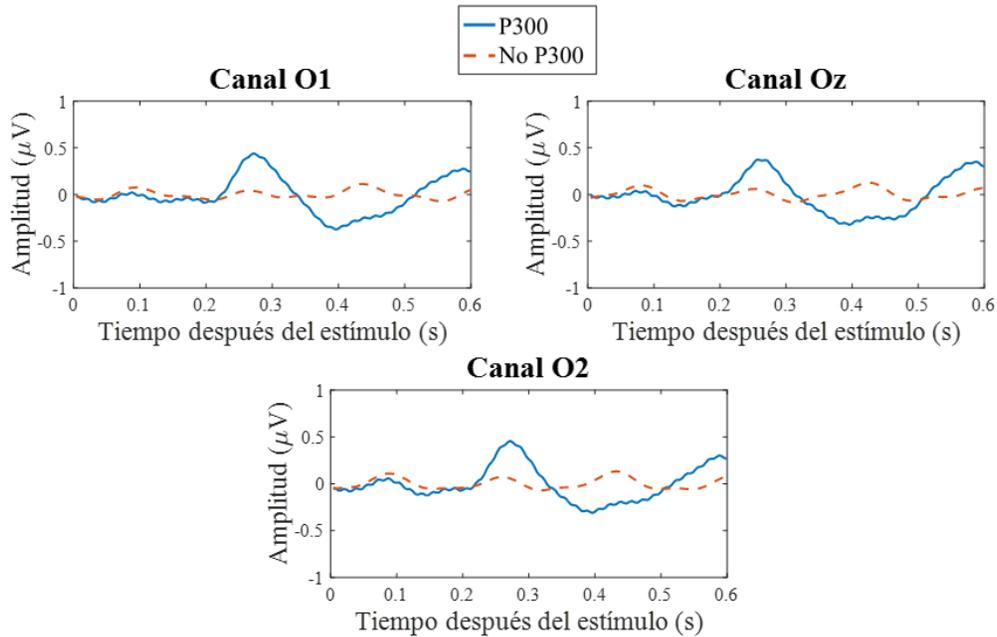


Figura 4.6 Promedio de vectores $x_{TJ}(m)$ (P300) y $x_{FJ}(m)$ (No P300), para el sujeto 7.

Una de las principales ventajas del uso de redes neuronales convolucionales, es que se encargan de la extracción de características, y de esta manera encuentran patrones en las señales en espacios multidimensionales, donde las operaciones que se realizan en estos espacios exceden a la capacidad humana.

Los resultados mostrados para CNN300 tienen un desempeño factible para aplicaciones en línea, que pueden ser aplicadas en situaciones reales. Sin embargo, debido a la variación de las señales de cada uno de los sujetos de pruebas, se concluyó que es necesario realizar un entrenamiento por sujeto, ya que resulta complicado generalizar las señales que se presentan en el cerebro humano.

El tiempo de procesamiento para la detección de un carácter es de 34.7 ms, el cual es menor que FFp300, por lo tanto, también es factible para una aplicación en línea. CNN300 también se implementó en Matlab[®] por medio del paquete de herramientas de redes neuronales.

CAPÍTULO 5. DETECCIÓN DE LA SEÑAL P300 PARA LA SELECCIÓN DE OBJETOS

Como se mencionó al inicio de este documento, la intención de analizar la señal P300 es generar un método de comunicación con propósitos de terapia. Aunado a esto, en el capítulo anterior se pudo concluir que el método con el mejor desempeño es CNN300, el cual utiliza una CNN para detectar la señal P300. En consecuencia, en este capítulo se describe el diseño de una interfaz de terapia virtual, en la cual, un sujeto puede seleccionar objetos mediante la detección de la señal P300. La extracción de características y clasificación se realiza mediante CNN300.

Para detallar el diseño y funcionamiento de la interfaz, este capítulo se compone de las siguientes secciones: en la sección 5.1 se describe el diseño de la interfaz, en la sección 5.2 se muestra cómo se implementó el método CNN300. Finalmente, en la sección 5.3 se muestran los resultados obtenidos y discusión.

5.1 Terapia virtual y P300

Cuando las señales P300 son evocadas por medio de estímulos visuales, es posible detectar el tiempo en el que ocurren, y de esta manera se pueden tomar decisiones, como se hace con el delectador de Donchin. Además de los delectadores, existen varias aplicaciones que pueden aprovechar la detección temporal de la señal P300, por ejemplo, el manejo de un control remoto [44], controlar brazos robóticos [23], entre otros. En el caso de este trabajo de tesis, se va a trabajar con terapias virtuales, las cuales consisten en el desarrollo de un esquema de rehabilitación basado en un escenario virtual, con el cual una persona puede interactuar directamente con objetos virtuales en un entorno que está dirigido a entrenar o fortalecer algún tipo de habilidad.

En el Laboratorio de Sistemas PVR[®] del Instituto Tecnológico de Chihuahua se cuenta con los escenarios virtuales del equipo Terapiatec[®], el cual se encuentra instalado en los Centros de Atención Múltiple (CAM) de la ciudad de Chihuahua. Estos escenarios se diseñaron para que personas con capacidades diferentes puedan realizar terapias de brazos y piernas. Para ello, el equipo de Terapiatec[®] utiliza un sensor de profundidad. No obstante, un trabajo que se pretende

hacer a futuro es anexar al equipo de Terapiatec® escenarios en los cuales las personas puedan interactuar con objetos virtuales mediante un sistema BCI en vez de un sensor de profundidad. De esta manera, personas que tienen discapacidades motoras y del habla podrán realizar diversos tipos de actividades. Por lo tanto, para contribuir con este esfuerzo de diseñar terapias virtuales con sistemas BCI, se implementó el método CNN300 en un escenario virtual que simula un ambiente real para entrenar a personas con capacidades diferentes para la toma de decisiones. Se eligió el método de CNN300 ya que tuvo mejor desempeño que FFp300.

Para que un escenario virtual pueda ser utilizado en la detección de P300, este debe generar intensificaciones similares a las que hace el deletreador de Donchin. En las siguientes subsecciones se muestran los detalles del escenario virtual y la manera de realizar las intensificaciones.

5.1.1 Escenario

El escenario que fue seleccionado para la terapia basada en P300 es el de “Frutas” del equipo Terapiatec®, el cual se muestra en la figura 5.1. La finalidad de utilizar este escenario es simular un ambiente real para que una persona que no pueda moverse seleccione objetos por medio de alguna computadora o algún dispositivo electrónico. Se puede observar que en este escenario consta de un cesto y cuatro cajas que contienen cuatro diferentes frutas: tomate, naranja, manzana y plátano. Este escenario fue seleccionado porque está diseñado con la finalidad de que las personas puedan seleccionar los objetos que se encuentran en las cajas. Además, es ideal para generar intensificaciones ya que los objetos se encuentran separados, y de esta manera las intensificaciones no se traslaparán. El escenario se implementó en Matlab® ya que este software permite realizar la interfaz visual, y procesar las señales con P300 sin la necesidad de transmitir la información entre diferentes paquetes de software, al igual que como se realizó con el deletreador de Donchin en el capítulo 2.

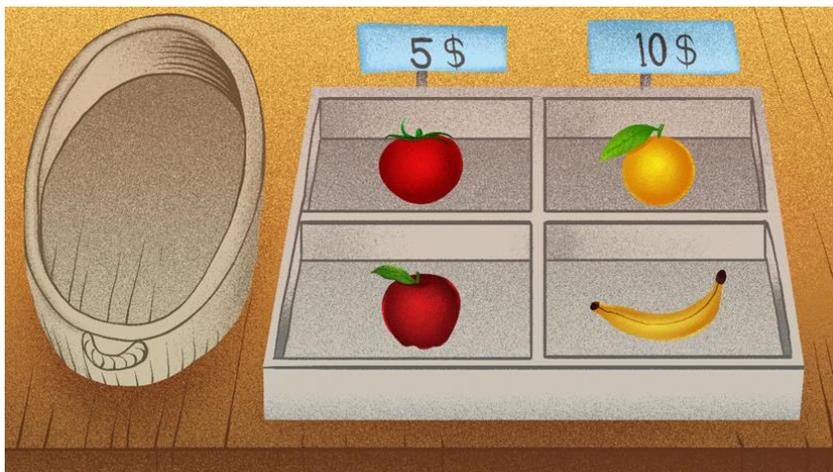


Figura 5.1 Escenario de frutas utilizado para la selección de objetos.

La interacción del sujeto con el escenario es la siguiente: el sujeto observa el escenario que se muestra en la figura 5.1, luego este piensa en una de las frutas y fija su vista en ella. Después de las intensificaciones para evocar el P300, la fruta que seleccionó aparece en la canasta que se encuentra a la izquierda, como lo muestra la figura 5.2 donde, se puede observar como la manzana fue seleccionada.

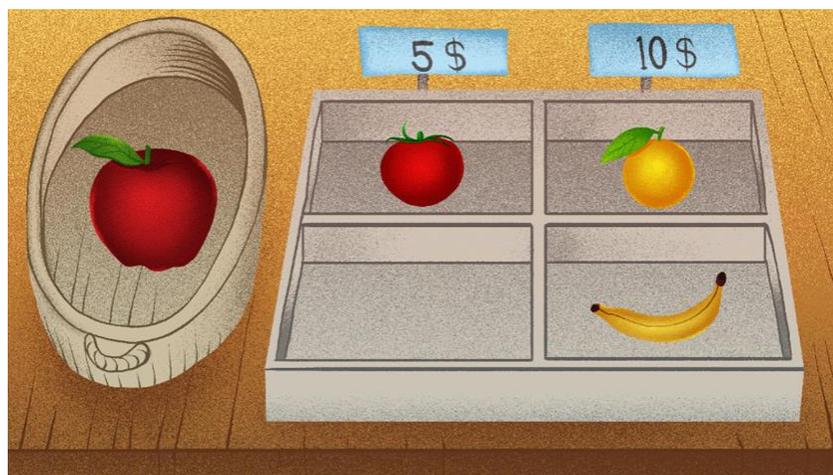


Figura 5.2 Escenario de frutas con la manzana seleccionada.

5.1.2 Generación de intensificaciones

Para la generación de las intensificaciones se optó por iluminar las casillas de cada uno de los objetos independientemente y no por filas y columnas como se hacía en el deletreador. Esto

fue debido a la cantidad de objetos que se tienen. Es decir, para el deletreador se tenía que elegir uno entre 36 caracteres, es por esto que las intensificaciones se realizaban por filas y columnas en la matriz. De esta manera solo existían 12 intensificaciones diferentes, en vez de una para cada carácter. Para el caso de la interfaz propuesta en este capítulo, solo se tienen cuatro objetos y al realizar intensificaciones se puede identificar cual es el objeto deseado de una manera directa y no realizando una correlación entre las filas y columnas como sucede en el deletreador. Para ejemplificar como se dan las intensificaciones, en la figura 5.3 se muestra la casilla del plátano intensificada.

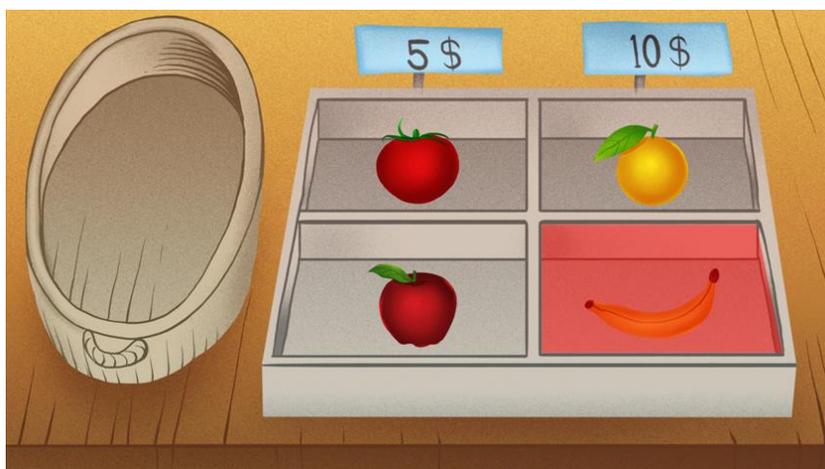


Figura 5.3 Escenario de frutas con uno de los objetos intensificados.

Debido a que la cantidad de objetos es menor en el escenario que en el deletreador, se redujo la frecuencia de intensificación, ya que se volvería más complicado observar y contar cada una de las intensificaciones si la frecuencia de intensificación es la misma que en el deletreador. Esta reducción de frecuencias también se realizó porque la idea es utilizar este escenario en personas con capacidades diferentes, las cuales en ocasiones no tienen el mismo nivel de atención que una persona sana [45]. En consecuencia, cada uno de los objetos dura intensificado 100 ms, pero en este caso el tiempo que se tarda sin intensificación es de 120 ms, lo cual es suficiente para alcanzar a distinguir el cambio entre intensificaciones. El tiempo que dura sin intensificación se definió preguntándoles a los sujetos si no perdían el conteo de las intensificaciones mientras se realizaban las sesiones.

Se eligieron diferentes colores para cada una de las casillas, con la finalidad de hacer más notable cual es el objeto que se está observando. Los colores fueron elegidos tomando en cuenta que no fueran del mismo color de la fruta, y diferente al color de la caja donde se encuentran. Para el tomate se eligió el color amarillo, para la naranja el verde, para la manzana el azul y para el plátano el rojo.

La manera en la que se realizan las sesiones de selección de objetos es la misma que la descrita en la sección 2.4 para la selección de caracteres. A los sujetos se les pide que fijen su vista en el objeto que desean seleccionar y posteriormente las casillas de cada objeto comienzan a intensificarse aleatoriamente. Los sujetos tienen que contar el número de veces que la casilla del objeto deseado fue intensificada, con la finalidad de asegurar que esté poniendo atención en dicho objeto. Se generan 60 intensificaciones, ya que al igual que el deletreador, cada uno de los objetos es intensificado un total de 15 veces, lo cual se hace con la finalidad de tener una mejor resolución de la señal P300 mediante el promediado.

Para el entrenamiento y validación del método CNN300 aplicado a selección de objetos, se generó una base de datos similar a la del deletreador de Donchin modificado. Esta base de datos contiene las señales electroencefalográficas de diez sujetos de pruebas en los canales O1, Oz y O2. Para el entrenamiento se tomaron un total de 28 épocas de selección de objetos y para la validación 20 épocas. Estas cantidades de épocas se definieron de tal manera que fuera una cantidad máxima de sesiones, pero sin ocasionar agotamiento en los sujetos de pruebas. Cada época de selección de objetos consiste en generar 4 diferentes tipos de intensificaciones, una por cada objeto (fruta). Por lo tanto, por cada sujeto de pruebas se tiene un total de 112 muestras para el entrenamiento y 80 para la validación. De las 112 de entrenamiento para cada sujeto, solo 28 deben contener la señal P300 y 83 no deben contenerla, ya que, de las cuatro intensificaciones por época de selección de objeto, solo una debe contener la señal P300. Para las 80 de validación, 20 muestras son las que contienen la señal P300 y 60 no.

Cabe mencionar que los sujetos de pruebas que participaron en el experimento descrito en este capítulo no son los mismos que los del capítulo 3 y 4, debido a que los experimentos se realizaron tiempo después de los experimentos con el deletreador y no fue posible reunir a los

sujetos anteriores. De igual manera los sujetos que conforman la base de datos tienen edad entre 20 y 25 años, con visión normal o corregida.

5.2 Detección de la señal P300 mediante CNN300

Al igual que en el capítulo 4, el método CNN300 para la terapia virtual consta de cuatro etapas: preprocesamiento, agrupamiento por intensificaciones, red neuronal convolucional y detección de objeto.

La etapa de preprocesamiento es similar a la utilizada para el deletreador por CNN300, donde se realiza el proceso de filtrado y normalización, obteniendo la señal $s_{ff}(n)$ de la ecuación 4.1. Sin embargo, para la terapia, n toma valores de 1 hasta 4495, ya que se realiza una cantidad menor de intensificaciones. Los valores de $1 \leq n \leq 1000$ corresponden a los 4 s para que el sujeto detecte el objeto a selección y ayudan a la estabilización del filtro. Como cada objeto dura intensificado durante 25 muestras (100 ms) y para el cambio de intensificación se toman 30 muestras (120 ms), se puede decir que las intensificaciones ocurren cada 55 muestras. Cada época de selección de objetos tiene 60 intensificaciones (15 para cada objeto), por lo tanto, se requieren de 3300 muestras para representarlos, entonces, los valores $1001 \leq n \leq 4300$ contienen las señales producidas por las intensificaciones. Por último, los valores $4301 \leq n \leq 4495$ son utilizados para completar un segundo después de la última intensificación.

El agrupamiento por intensificaciones también se realiza de forma similar que en CNN300, donde la señal $\varphi_{ff}(n)$ se descompone en segmentos $\chi_{ij}(m)$, para ser promediadas y obtener $\overline{\chi_{kj}}(m)$ al igual que en la ecuación 4.2. Posteriormente, las señales $\overline{\chi_{kj}}(m)$ se acondicionan para formar la matriz de entrada $I_k(j,m)$, como el proceso mostrado en la figura 4.2, pero en este caso $k = \{1,2,3,4\}$. En el caso del deletreador k podía tomar 12 valores, debido a que se intensificaban seis filas y seis columnas. No obstante, en el caso de la terapia solo se intensifican cuatro objetos, por eso k obtiene este rango. La CNN utilizada en la terapia tiene la misma arquitectura y parámetros que la descrita en el capítulo 4. El entrenamiento se realiza mediante el proceso descrito en la sección 4.4.5, que corresponde al algoritmo de gradiente estocástico descendiente con momento [14]. Por último, los patrones $I_k(j,m)$ son propagados a través de la CNN para

generar el vector $P(c_b/T_k, \theta)$ de probabilidades para cada uno de los objetos k de pertenecer a la clase $b=1$ (ver ecuación 4.9). Para la predicción del objeto se toma en cuenta la siguiente ecuación:

$$D = \operatorname{argmax} \left(P(c_1 | T_k, \theta) \right) \quad 1 \leq k \leq 4 \quad 5.1$$

donde, D representa el índice del objeto que tiene la mayor probabilidad de contener la señal P300. Posteriormente, el objeto predicho debe sobrepasar de cierto umbral $u=0.8$ de probabilidad de contener la señal P300, como lo descrito en la sección 3.3.2, se tiene la siguiente ecuación:

$$A = \begin{cases} N(D) & \text{si } u < P(c_1 | T_D, \theta) \\ \text{Indefinido} & \text{de otra forma} \end{cases} \quad 5.2$$

donde, A es el objeto predicho y N es el vector que contiene los objetos y se puede definir como:

$$N = \{ \text{Tomate, Naranja, Manzana, Plátano} \} \quad 5.3$$

5.3 Resultados y discusión

En esta sección se muestran los resultados obtenidos del método CNN300 en la detección de la señal P300 y los objetos con la terapia virtual “Frutas”. A diferencia de los capítulos anteriores, para la terapia presentada en este capítulo se presentan resultados con diferentes cantidades de intensificaciones por época de selección de objetos. Es decir, por cada época se analizaron 15, 10 y 5 intensificaciones para la detección de la señal P300 y en la detección de carácter. La finalidad de esta prueba es elegir cual es el número de intensificaciones más bajo para que los sujetos de pruebas seleccionen un objeto lo más rápido posible sin sacrificar el desempeño del método. Esta reducción de intensificaciones por época impacta en reducir el tiempo que dura cada época de selección de objetos, haciendo la terapia más simple. Cuando se realizan 15 intensificaciones, cada época de selección de objeto dura un tiempo de 13.2 s, para 10 se tiene un tiempo de 8.8 s y para 5 se requieren de 4.4 s. Estas pruebas se realizaron para

esta terapia y no para el deletreador debido a que la terapia tiene una cantidad de clases menor. Lo que conlleva a que el desempeño de CNN300 no se vea tan afectado.

Los resultados en la detección de la señal P300 con la terapia “Frutas” se muestran en la sección 5.3.1. Los resultados en la detección de los objetos con la terapia “Frutas” se muestran en la sección 5.3.2.

5.3.1 Resultados para la detección de la señal P300 en cada objeto

Para la detección de la señal P300, en las tablas 5.1, 5.2 y 5.3 se muestra la cantidad (#) y porcentaje de los verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN) para entrenamiento y validación, en 15, 10 y 5 intensificaciones respectivamente. Estos resultados fueron obtenidos de las matrices de confusión generadas para la detección de la señal P300.

Tabla 5.1 Resultados de las matrices de confusión tomando en cuenta 15 intensificaciones.

Sujeto	Entrenamiento								Validación							
	VP		VN		FP		FN		VP		VN		FP		FN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	28	11.7	84	35.0	0	0.0	0	0.0	19	11.3	57	33.9	3	1.8	1	0.6
2	28	11.7	84	35.0	0	0.0	0	0.0	17	10.1	60	35.7	0	0.0	3	1.8
3	28	11.7	84	35.0	0	0.0	0	0.0	19	11.3	59	35.1	1	0.6	1	0.6
4	28	11.7	84	35.0	0	0.0	0	0.0	16	9.5	59	35.1	1	0.6	4	2.4
5	28	11.7	84	35.0	0	0.0	0	0.0	19	11.3	60	35.7	0	0.0	1	0.6
6	28	11.7	84	35.0	0	0.0	0	0.0	17	10.1	60	35.7	0	0.0	3	1.8
7	27	11.3	82	34.2	2	0.8	1	0.4	13	7.7	58	34.5	2	1.2	7	4.2
8	28	11.7	84	35.0	0	0.0	0	0.0	13	7.7	58	34.5	2	1.2	7	4.2
9	28	11.7	84	35.0	0	0.0	0	0.0	18	10.7	59	35.1	1	0.6	2	1.2
10	28	11.7	84	35.0	0	0.0	0	0.0	19	11.3	59	35.1	1	0.6	1	0.6
Ideal	40	16.7	200	83.3	0	0.0	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0

Tabla 5.2 Resultados de las matrices de confusión tomando en cuenta 10 intensificaciones.

Sujeto	Entrenamiento								Validación							
	VP		VN		FP		FN		VP		VN		FP		FN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	28	11.7	84	35.0	0	0.0	0	0.0	17	10.1	50	29.8	10	6.0	3	1.8
2	28	11.7	84	35.0	0	0.0	0	0.0	13	7.7	58	34.5	2	1.2	7	4.2
3	28	11.7	84	35.0	0	0.0	0	0.0	17	10.1	57	33.9	3	1.8	3	1.8
4	28	11.7	84	35.0	0	0.0	0	0.0	13	7.7	55	32.7	5	3.0	7	4.2
5	28	11.7	84	35.0	0	0.0	0	0.0	19	11.3	59	35.1	1	0.6	1	0.6
6	28	11.7	84	35.0	0	0.0	0	0.0	18	10.7	56	33.3	4	2.4	2	1.2
7	25	10.4	82	34.2	2	0.8	3	1.3	10	6.0	59	35.1	1	0.6	10	6.0
8	28	11.7	84	35.0	0	0.0	0	0.0	12	7.1	49	29.2	11	6.5	8	4.8
9	27	11.3	83	34.6	1	0.4	1	0.4	15	8.9	58	34.5	2	1.2	5	3.0
10	28	11.7	84	35.0	0	0.0	0	0.0	18	10.7	59	35.1	1	0.6	2	1.2
Ideal	40	16.7	200	83.3	0	0.0	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0

Tabla 5.3 Resultados de las matrices de confusión tomando en cuenta 5 intensificaciones.

Sujeto	Entrenamiento								Validación							
	VP		VN		FP		FN		VP		VN		FP		FN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	28	11.7	84	35.0	0	0.0	0	0.0	16	9.5	53	31.5	7	4.2	4	2.4
2	28	11.7	84	35.0	0	0.0	0	0.0	11	6.5	54	32.1	6	3.6	7	4.2
3	28	11.7	84	35.0	0	0.0	0	0.0	15	8.9	55	32.7	5	3.0	5	3.0
4	28	11.7	84	35.0	0	0.0	0	0.0	10	6.0	49	29.2	11	6.5	10	6.0
5	28	11.7	84	35.0	0	0.0	0	0.0	14	8.3	58	34.5	2	1.2	6	3.6
6	28	11.7	84	35.0	0	0.0	0	0.0	10	6.0	54	32.1	6	3.6	10	6.0
7	28	11.7	82	34.2	2	0.8	0	0.0	9	5.4	53	31.5	7	4.2	11	6.5
8	28	11.7	84	35.0	0	0.0	0	0.0	15	8.9	49	29.2	11	6.5	5	3.0
9	28	11.7	84	35.0	0	0.0	0	0.0	11	6.5	54	32.1	6	3.6	9	5.4
10	28	11.7	84	35.0	0	0.0	0	0.0	14	8.3	57	33.9	3	1.8	6	3.6
Ideal	40	16.7	200	83.3	0	0.0	0	0.0	28	16.7	140	83.3	0	0.0	0	0.0

En las tablas 5.1, 5.2 y 5.3 se puede observar que los resultados obtenidos en el entrenamiento son casi ideales, ya que para 15, 10 y 5 intensificaciones casi todos los sujetos

obtienen resultados ideales. Sin embargo, también se puede observar en las mismas tablas que en la validación los VP y VN decaen conforme disminuye el número de intensificaciones que se analizan. Esto se debe a que el promedio de las señales baja la relación señal ruido conforme cae la cantidad de intensificaciones. No obstante, se puede observar en las tablas 5.1, 5.2 y 5.3 que el desempeño es similar para los sujetos de pruebas.

Al igual que en los resultados para el capítulo 3 y 4, se obtuvieron las mediciones de reconocimiento, precisión, sensibilidad y *F-measure*. Estas mediciones se muestran en las tablas 5.4, 5.5 y 5.6 y corresponden a las pruebas con 15, 10 y 5 intensificaciones respectivamente. La finalidad de mostrar estas métricas es observar el comportamiento del método detalladamente con las diferentes cantidades de intensificaciones utilizadas. Es de esperarse que las métricas muestren desempeños más bajos conforme se reduce el número de intensificaciones, ya que es más difícil diferenciar cuando se encuentra o no las señales P300, debido a que la relación señal ruido también va disminuyendo.

Tabla 5.4 Resultados con base en las métricas de evaluación para 15 intensificaciones.

Sujeto	Entrenamiento				Validación			
	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>
1	1.00	1.00	1.00	1.00	0.95	0.86	0.95	0.90
2	1.00	1.00	1.00	1.00	0.96	1.00	0.85	0.92
3	1.00	1.00	1.00	1.00	0.98	0.95	0.95	0.95
4	1.00	1.00	1.00	1.00	0.94	0.94	0.80	0.86
5	1.00	1.00	1.00	1.00	0.99	1.00	0.95	0.97
6	1.00	1.00	1.00	1.00	0.96	1.00	0.85	0.92
7	0.97	0.93	0.96	0.95	0.89	0.87	0.65	0.74
8	1.00	1.00	1.00	1.00	0.89	0.87	0.65	0.74
9	1.00	1.00	1.00	1.00	0.96	0.95	0.90	0.92
10	1.00	1.00	1.00	1.00	0.98	0.95	0.95	0.95
Promedio	0.99	0.99	0.99	0.99	0.94	0.93	0.85	0.88

Tabla 5.5 Resultados con base en las métricas de evaluación para 10 intensificaciones.

Sujeto	Entrenamiento				Validación			
	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>
1	1.00	1.00	1.00	1.00	0.84	0.63	0.85	0.72
2	1.00	1.00	1.00	1.00	0.89	0.87	0.65	0.74
3	1.00	1.00	1.00	1.00	0.93	0.85	0.85	0.85
4	1.00	1.00	1.00	1.00	0.85	0.72	0.65	0.68
5	1.00	1.00	1.00	1.00	0.98	0.95	0.95	0.95
6	1.00	1.00	1.00	1.00	0.93	0.82	0.90	0.86
7	0.96	0.93	0.89	0.91	0.86	0.91	0.50	0.65
8	1.00	1.00	1.00	1.00	0.76	0.52	0.60	0.56
9	0.98	0.96	0.96	0.96	0.91	0.88	0.75	0.81
10	1.00	1.00	1.00	1.00	0.96	0.95	0.90	0.92
Promedio	0.99	0.99	0.99	0.99	0.89	0.81	0.76	0.77

Tabla 5.6 Resultados con base en las métricas de evaluación para 5 intensificaciones.

Sujeto	Entrenamiento				Validación			
	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>	Rec.	Precisión	Sensibilidad	<i>F-Measure</i>
1	1.00	1.00	1.00	1.00	0.86	0.70	0.80	0.74
2	1.00	1.00	1.00	1.00	0.83	0.65	0.61	0.63
3	1.00	1.00	1.00	1.00	0.88	0.75	0.75	0.75
4	1.00	1.00	1.00	1.00	0.74	0.48	0.50	0.49
5	1.00	1.00	1.00	1.00	0.90	0.88	0.70	0.78
6	1.00	1.00	1.00	1.00	0.80	0.63	0.50	0.56
7	0.98	0.93	1.00	0.97	0.78	0.56	0.45	0.50
8	1.00	1.00	1.00	1.00	0.80	0.58	0.75	0.65
9	1.00	1.00	1.00	1.00	0.81	0.65	0.55	0.59
10	1.00	1.00	1.00	1.00	0.89	0.82	0.70	0.76
Promedio	0.99	0.99	1.00	0.99	0.82	0.66	0.63	0.64

En lo que concierne al entrenamiento, se puede observar que la tasa de reconocimiento promedio es de 0.99 para los tres casos mostrados en las tablas 5.4, 5.5 y 5.6. Debido a que el desempeño en el entrenamiento es prácticamente ideal, las demás métricas también deben

mostrar resultados similares, lo cual se puede observar en dichas tablas. Estos resultados en el entrenamiento hacen dudar de que algoritmo esté realizando una buena generalización de los datos, o simplemente está memorizando las muestras de entrenamiento. Para esto, se tiene que realizar un análisis de los desempeños en la validación, el cual se muestra a continuación.

De acuerdo con los resultados de las tablas 5.4, 5.5 y 5.6, para la validación se tiene una tasa promedio de reconocimiento de 0.94 cuando se utilizan 15 intensificaciones, 0.89 con 10 intensificaciones y 0.82 cuando solo se utilizan 5 intensificaciones. Es decir, existe una generalización adecuada en las pruebas de 15 intensificaciones, ya que la tasa de reconocimiento en validación es cercana a la de entrenamiento. No obstante, para los casos de 10 y 5 intensificaciones, la tasa de reconocimiento en validación cae considerablemente. Por lo tanto, se asume para estos casos que en los resultados de 10 y 5 intensificaciones el método CNN300 está memorizando las muestras de entrenamiento. Una razón por la que puede estar sucediendo esta memorización es que, al bajar la cantidad de intensificaciones, se reduce la relación señal ruido en el promedio $\overline{\chi_{k_j}(m)}$, por lo tanto, es necesario generalizar más los datos ya que las características extraídas para los casos con P300 y No P300 son más similares.

Esto no significa que la CNN no funcione adecuadamente con menos intensificaciones. Lo que sucede es que la arquitectura de la CNN en CNN300 se diseñó con 15 intensificaciones. Es decir, podría rediseñar la arquitectura de la CNN con menos intensificaciones para tratar de mejorar los desempeños en la validación.

De acuerdo con las tablas 5.4, 5.5 y 5.6, los resultados promedio en validación de la métrica *F-measure* son 0.88, 0.77 y 0.64 para los casos de 15, 10, 5 intensificaciones respectivamente. Dados estos resultados se puede observar que la tasa promedio de *F-measure* también decae a bajar el número de intensificaciones. Sin embargo, para algunos sujetos de pruebas no es tan severo el decaimiento en *F-measure*, como el sujeto 5 que tiene tasas de 0.97, 0.95 y 0.78 en 15, 10 y 5 intensificaciones respectivamente. Esto puede ser causado debido a los diferentes factores que actúan sobre la generación de las señales P300.

Al igual que en el deletreador, los resultados en la terapia virtual para la detección del objeto dependen de los sujetos de prueba. Específicamente, en que tan buena es la amplitud de su señal P300. No se puede hacer un método que sea invariante a las personas, ya que depende de muchos factores como son generadas sus señales P300.

5.3.2 Resultados para la detección del objeto deseado

En esta sección se muestran los resultados obtenidos en la predicción del objeto deseado. En la tabla 5.7 se muestra el porcentaje de objetos reconocidos correctamente para cada uno de los sujetos de pruebas, tanto para entrenamiento (E) como para validación (V) en 15, 10 y 5 intensificaciones.

Tabla 5.7 Porcentaje de épocas de selección de objetos reconocidas correctamente.

Sujeto	Intensificaciones					
	15		10		5	
	E	V	E	V	E	V
1	100%	95%	100%	75%	100%	85%
2	100%	95%	100%	80%	100%	60%
3	100%	100%	100%	85%	100%	80%
4	96.42%	95%	96.42%	80%	96.42%	55%
5	100%	100%	100%	90%	100%	85%
6	100%	100%	100%	90%	100%	65%
7	96.42%	90%	96.42%	75%	100%	45%
8	100%	95%	100%	75%	100%	85%
9	100%	95%	100%	85%	100%	75%
10	100%	100%	100%	95%	100%	75%
Promedio	99.28%	96.5%	99.28%	83%	99.64%	71%

Con base en los resultados de la tabla 5.7, se puede observar que el porcentaje de reconocimiento del objeto deseado en el entrenamiento es cercano al 100%, ya que se obtuvieron resultados de 99.28% para 15 y 10 intensificaciones y 99.64% para 5. Sin embargo, los resultados en validación varían considerablemente de acuerdo con el número de intensificaciones: 96.5% con 15 intensificaciones, 83% con 10 y 71% con 5.

Al igual que en el deletreador, el porcentaje de reconocimiento de la señal P300, puede ser mayor o menor al porcentaje de objetos reconocidos correctamente. Esto es debido a que en una época de selección de objetos se pueden predecir dos o más objetos que contienen la señal P300. Sin embargo, la decisión se toma por aquel objeto que tiene la mayor probabilidad de contener la señal, sobrepasando el umbral $u=0.8$.

Aunado a esto, para el caso de 15 intensificaciones, la tasa de reconocimiento de señales P300 fue de 0.94 en la validación, equivalente a un 94%, y los objetos reconocidos correctamente fueron el 96.5%. Para caso particular fue porque en la detección de la señal P300, los verdaderos positivos tuvieron mayor probabilidad de contener la señal P300 que los falsos positivos. Cuando se trata de 10 intensificaciones, la tasa de reconocimiento para las señales P300 fue de 0.89, equivalente a un 89%, y para los objetos se reconoció el 83% correctamente. Para este caso, los verdaderos positivos tuvieron menor probabilidad de contener la señal P300 que los falsos positivos, por eso el porcentaje cae. Por último, para el caso de 5 intensificaciones sucede lo mismo que para 10, donde la tasa de reconocimiento de señales P300 es de 0.82, equivalente a 82% y los objetos reconocidos correctamente fueron el 71%.

Por lo tanto, al reducir el número de intensificaciones la cantidad de falsos positivos aumenta. Debido a que los promedios de las intensificaciones que no contienen P300, es decir, que solo contienen ruido aleatorio, no tienden a ser cero, disminuyendo la relación señal ruido. En conclusión, el número de intensificaciones recomendadas para utilizar puede variar para cada uno de los sujetos de pruebas, en función de cómo se genera su señal P300, lo cual puede ser analizado visualmente observando el promedio de las señales con P300 contra aquellas que no tienen P300.

CAPÍTULO 6. COMPARACIÓN DE LOS RESULTADOS

En este capítulo se comparan los resultados obtenidos por los métodos FFp300 y CNN300 contra los algoritmos más populares en la literatura, diseñados con deletreadores para detectar la señal P300. Para la comparación se propone un criterio de evaluación que considera el porcentaje de reconocimiento, número de sujetos utilizados, cantidad de canales y el tipo de dispositivo utilizado.

Para realizar la comparación de los métodos, este capítulo se estructura con las siguientes secciones: en la sección 6.1 se muestra el estado del arte de los deletreadores basados en P300 y en la sección 6.2 se analizan y comparan los resultados obtenidos por cada uno de los métodos.

6.1 Estado del arte de los deletreadores basados en P300

En la literatura existen diversos métodos para la detección de las señales P300 generadas por diversos estímulos. El estándar dentro de estos métodos es el uso del deletreador de Donchin como medio para evocar la señal P300. A continuación, se muestran los métodos más populares para deletrear caracteres mediante la detección de la señal P300. Todos los métodos mostrados son evaluados en dos sujetos de pruebas cuyas señales fueron obtenidas por medio del dispositivo g.tec (versión 2003).

6.1.1 Métodos basados en máquinas de vectores de soporte

A. Rakotomamonjy y Vicent Guigue proponen en [16] cinco métodos basados en máquinas de vectores de soporte (SVM por sus siglas en inglés). La diferencia entre los métodos propuestos son la cantidad de canales y clasificadores SVM utilizados. A continuación, se detallan los cinco métodos propuestos:

- **ESVM_1**: este método consiste en 17 clasificadores SVM que modelan el ruido existente y seleccionan los canales basándose en la aportación que provee cada uno en la información a procesar. De esta manera, para cada sujeto de prueba se pueden utilizar diferentes canales. ESVM_1 obtiene un porcentaje de reconocimiento promedio de

96.5% en el reconocimiento de caracteres. Para un sujeto se utilizan 40 canales, y para el otro 16.

- ESVM_2: este método es muy similar a ESVM_1, pero no utiliza el algoritmo de selección de canales. Para este caso se utilizan 64 canales (los que se especifican en la figura 2.2). El porcentaje promedio de reconocimiento de caracteres correctos es de 95.5%.
- ESVM _3: el funcionamiento de este método es igual al ESVM_1 y ESVM_2, pero utiliza ocho canales: Fz, Cz, Pz, C3, C4, P3, P4 y Oz. El porcentaje de reconocimiento es de 80%.
- SSVM_1: a diferencia de los ESVM, este método utiliza solo un clasificador SVM y los 64 canales. Obtiene un porcentaje de reconocimiento promedio de 96.5%.
- SSVM_2: al igual que SSVM_1 este método utiliza solo un clasificador SVM, pero utiliza ocho canales Fz, Cz, Pz, C3, C4, P3, P4 y Oz. De esta manera, dicho método tiene un porcentaje de reconocimiento de caracteres de 70%.

En [46] Vibhore Bhatnagar et al., proponen una modificación al trabajo de A. Rakotomamonjy y Vicent Guigue [16]. La modificación consiste en que los datos de entrada a ESVM se realizan mediante promediados por filas y columnas y no por intensificación. De la misma manera, el método ESVM_1 utiliza 64 canales. Para este método se tiene un porcentaje promedio de reconocimiento de 92.5%. Este porcentaje no logra superar al método propuesto en [16], pero logra reducir el tiempo de procesamiento.

6.1.2 Análisis de componentes independientes para la detección de P300

Yupeng Wang et al., en [47] trabajan con un método de clasificación de la señal P300 para la detección del carácter correcto en el deletreador de Donchin. La primera etapa consiste en extraer características en el dominio del tiempo con análisis de componentes independientes (ICA). La segunda etapa extrae características en el dominio de la frecuencia con transformada Wavelet. Finalmente, se aplica una etapa de clasificación por SVM. Obtienen un porcentaje de reconocimiento promedio de carácter de 89.5%, donde, en un sujeto de prueba utilizan 13 canales, y para el otro sujeto 12 canales.

6.1.3 Redes neuronales convolucionales para detección de P300

En [17] Hubert Cecotti y Axel Gräser proponen siete modelos de clasificación basados en redes neuronales convolucionales. Las diferencias entre estos son la forma en que alimentan los datos a las redes, los canales utilizados y la cantidad de redes. Los modelos que se nombran CNN son aquellos que solo utilizan una red neuronal convolucional, y los modelos que llevan el nombre MCNN, utilizan múltiples redes neuronales convolucionales. A continuación, se define como es el funcionamiento de cada uno de los modelos propuestos:

- CNN-1: la red neuronal convolucional que utiliza este clasificador es alimentada por 64 canales y tiene un porcentaje de reconocimiento promedio de 94.5%.
- CNN-2a: este clasificador es similar al CNN-1, pero la red se alimenta con ocho canales: Fz, Cz, Pz, P3, P4, PO7, PO8 y Oz. El porcentaje de reconocimiento de caracteres promedio es de 87.5%.
- CNN-2b: este clasificador funciona de la misma manera que CNN-2a, pero los ocho canales utilizados son aquellos que tienen mayor valor en los pesos de la primera capa oculta de la CNN-1. Este método tiene un porcentaje de reconocimiento de 87%.
- CNN-3: es igual que CNN-1 pero no combina la información entre los canales. Este clasificador analiza 64 canales por separado. Obtiene un desempeño de reconocimiento promedio de 88.5%.
- MCNN-1: este método está compuesto por cinco redes neuronales convolucionales. Cuando se trabaja con el deletreador, por cada muestra con la señal P300, existen cinco que no contienen P300. Para tener la misma cantidad de muestras con y sin P300 para cada red, son entrenadas con todas las muestras que contienen la señal P300 y un quinto de las muestras que no contienen P300. Este método es el que obtuvo los mejores resultados en este trabajo, por medio de 64 canales utilizados obtienen un porcentaje de reconocimiento promedio de 95.5% en la predicción del carácter.
- MCNN-2: en este método se propone realizar un entrenamiento de cinco redes neuronales convolucionales, cada una con un quinto de la información que se tiene para el entrenamiento. De esta manera es posible modelar el comportamiento de la señal P300

independientemente del tiempo cuando se tomaron las muestras. El porcentaje de reconocimiento promedio de este método para la detección del carácter es de 90.5%, utilizando 64 canales.

- MCNN-3: este método está compuesto por tres redes neuronales convolucionales con la arquitectura de CNN-1. Como los pesos son inicializados aleatoriamente, se espera que se obtengan diferentes soluciones para la CNN-1. El objetivo de MCNN-3 es mejorar la confiabilidad de CNN-1. En caso de que no mejore, el clasificador mostrará que la inicialización aleatoria conlleva a tener el mismo resultado siempre para este caso particular. El reconocimiento promedio de MCNN-3 es el mismo que CNN-1, de 94.5% mediante 64 canales. Este resultado indica que, aunque la inicialización de los pesos sea aleatoria, el ajuste de los pesos en el entrenamiento siempre es el mismo.

Mingfei Liu et al., en [48] proponen un método basado en una red neuronal convolucional, para la detección de la señal P300 en el deletreador de Donchin. Este método es llamado BN³, donde utilizan normalización por lotes y descartan aleatoriamente cierta información de las características extraídas por la CNN para mejorar la generalización de la red. También, utilizan una rectificación lineal en la salida de las capas convolucionales de la red, para acelerar el entrenamiento. Una vez que es detectada la señal P300, identifican cual es el carácter que los sujetos desean deletrear. BN³, utilizando 64 canales del sistema internacional 10-20, logran obtener un desempeño promedio en el reconocimiento del carácter de 96.5%.

6.1.4 Técnicas de promediado para el análisis de ERP

En [49] Nanying Liang y Laurent Bougrain proponen un método para la detección de la señal P300 evocada por medio del deletreador de Donchin con cuantización de vectores de aprendizaje (LVQ por sus siglas en inglés). Este método se denomina múltiple LVQ (mLVQ) ya que utiliza 64 canales. Las pruebas se realizan con dos sujetos y obtienen un porcentaje de reconocimiento promedio para los dos sujetos de 91.5% en la predicción del carácter.

6.2 Análisis y comparación de los métodos

Los resultados de los métodos FFp300 y CNN300 que se van a utilizar en la comparación, son los que se obtuvieron con el deletreador de Donchin modificado. Para ambos casos, se utilizó el dispositivo OpenBCI® y la cantidad de sujetos fue de diez.

Los métodos mostrados en la sección 6.1 fueron seleccionados para este estudio comparativo ya que son métodos populares en la literatura, y todos utilizan el deletreador de Donchin. Las señales utilizadas por todos los métodos usan una base de datos diseñada para una competencia en el año de 2003, la cual se generó con el dispositivo g.tec versión 2003 y recurre a dos sujetos de pruebas. Todos los métodos mostrados evalúan sus resultados de formas distintas ya que en algunos de los métodos muestran resultados para la detección de la señal P300 después de cada intensificación, otros detectan la señal P300 por fila y columna de cada época de deletreo, y otros más, solo muestran el porcentaje de reconocimiento de caracteres. La forma de analizar los desempeños depende de la manera de utilizar la detección de la señal P300. Sin embargo, todos los métodos tienen en común la evaluación por medio del porcentaje promedio de reconocimiento de caracteres.

Para este estudio comparativo, se realizó primero un análisis de todos los métodos que se pueden ver sintetizados en la tabla 6.1, la cual muestra las características principales de cada uno de los métodos propuestos. Estas características son: número de sujetos para las pruebas, cantidad de canales, dispositivo a utilizar y porcentaje de reconocimiento. Adicional a ello, se muestra también un ranking basado únicamente en el porcentaje de reconocimiento promedio sin tomar en cuenta las demás características, debido a que en la literatura es la única métrica que se muestra en común para todos los métodos. Por esto, dicho ranking no puede ser tomado en cuenta para conocer el desempeño real de los métodos, ya que en un sistema BCI, no solo se tiene que contemplar la capacidad de clasificar las señales, sino también se tiene que comprobar que tan práctico es y que pueda reproducir los resultados en varios sujetos de pruebas.

Tabla 6.1 Desempeño y características de los métodos propuestos en la literatura y este trabajo de tesis. El ranking solo toma en cuenta el porcentaje de reconocimiento.

Método	# de sujetos	# de canales	Dispositivo	% de reconocimiento promedio	Ranking por % de reconocimiento
FFp300	10	3	OpenBCI	79.28%	13
CNN300	10	3	OpenBCI	92.85%	4
ESVM_1	2	40 y 16	g.tec	96.5%	1
ESVM_2	2	64	g.tec	95.5%	2
ESVM_3	2	8	g.tec	80%	12
SSVM_1	2	64	g.tec	96.5%	1
SSVM_2	2	8	g.tec	70%	14
Modificación ESVM	2	64	g.tec	92.5%	5
mLVQ	2	64	g.tec	91.5%	6
CNN-1	2	64	g.tec	94.5%	3
CNN-2a	2	8	g.tec	87.5%	10
CNN-2b	2	8	g.tec	87%	11
CNN-3	2	64	g.tec	88.5%	9
MCNN-1	2	64	g.tec	95.5%	2
MCNN-2	2	64	g.tec	90.5%	7
MCNN-3	2	64	g.tec	94.5%	3
BN ³	2	64	g.tec	96.5%	1
ICA	2	13 y 12	g.tec	89.5%	8

En la tabla 6.1 se puede observar que los métodos FFp300 y CNN300 tienen un ranking de 13 y 4 respectivamente. FFp300 tiene un desempeño de 79.28% en la detección de los caracteres correctos, el cual se encuentra en el penúltimo lugar del ranking. Sin embargo, utiliza solo tres canales, a diferencia de los métodos que se encuentra en los primeros lugares que utilizan 64 canales. Para el caso de CNN300 su desempeño de 92.85%, lo cual lo coloca en el cuarto lugar del ranking y cercano a los primeros lugares, utilizando solo tres canales, y su reproducibilidad está comprobada con diez sujetos de pruebas. Un dato importante mostrado en la tabla 6.1, es que los algoritmos reportados en la literatura con desempeño superior al 90%, utilizan una gran cantidad de electrodos.

Un aspecto importante que considerar, es que los métodos revisados en la sección anterior se enfocan a reportar porcentaje de reconocimiento, porque solo se propone el algoritmo. Sin embargo, la contribución de esta tesis es el diseño de un sistema BCI completo que procese señales P300 mediante dos métodos: FFp300 y CNN300. Por lo tanto, enfocarse únicamente al grado de reconocimiento podría no describir en concreto la aportación total del sistema desarrollado. En consecuencia, se propone realizar un criterio de evaluación que considere aspectos relacionados tanto al diseño del sistema y el algoritmo, como se realiza en [50] para SSVEPs. Aparte del porcentaje de reconocimiento, se consideró que el número de sujetos donde se probó el algoritmo, los canales utilizados para la detección de la señal P300 y el dispositivo utilizado son los adecuados para evaluar el funcionamiento del sistema BCI. Estos criterios son elegidos de acuerdo con lo siguiente:

- Porcentaje de reconocimiento promedio (ζr): este criterio es elegido debido a que es el utilizado para evaluar el desempeño del algoritmo.
- Número de sujetos donde se probó el algoritmo (ζs): se sabe que la señal P300 es muy subjetiva en cada una de las personas. Por esta razón, el método propuesto debe probarse con la mayor cantidad de sujetos posible, ya que de esta manera se puede saber la capacidad de generalización de cada método.
- Número de canales utilizados (ζc): se busca la menor cantidad de electrodos posible, ya que, a mayor cantidad de electrodos, aumenta la complejidad computacional de los algoritmos, y es más complicado colocar los electrodos en los sujetos de pruebas. Esto puede repercutir en complicaciones para el uso del sistema BCI en aplicaciones reales.
- Tipo de dispositivo EEG (ζd): la importancia del dispositivo radica principalmente el costo y en la portabilidad. Es decir, un dispositivo caro y complicado de utilizar, es poco factible para una aplicación que pueda ser utilizada por personas que en verdad lo necesitan.

En el caso de los métodos de la sección 6.1, se toman como datos del sistema de adquisición los mostrados en la tabla 6.1.

6.2.1 Pesos de cada criterio de evaluación

Una vez que se definieron los criterios de evaluación, se propone repartir un total de 100 puntos entre los cuatro criterios y a cada criterio se le asignará cierta puntuación máxima. La puntuación para el criterio de porcentaje de reconocimiento es representada por ζ_r , para el número de sujetos es ζ_s , ζ_c representa el número de canales y ζ_d el tipo de dispositivo. De esta manera, el ranking de los métodos será de acuerdo con la puntuación obtenida.

Para establecer una ponderación de los criterios de evaluación y asignar la puntuación máxima en cada uno, se realizó un análisis de los resultados, características y condiciones de los métodos reportados en la sección 6.1.

Para el criterio ζ_s se consideró lo siguiente, todos los métodos reportados en la sección 6.1 trabajan con dos sujetos de pruebas. La variación de ζ_r entre los dos sujetos para todos los métodos en promedio es cercano a un 5%, sin embargo, no se tiene conocimiento de cómo sería dicha variación al agregar más sujetos de prueba. Dado esto, se decidió colocar la puntuación máxima del criterio ζ_s en 7 puntos. Así, la reproducibilidad del algoritmo tendrá cierta contribución en los resultados de esta evaluación.

En el caso del criterio ζ_d se decidió colocar una puntuación máxima menor a la de ζ_s , debido a que se consideró más importante la reproducibilidad del algoritmo que el dispositivo utilizado. Sin embargo, como se mencionó anteriormente es importante que el dispositivo utilizado sea de fácil adquisición para las personas, además de que sea ligero y portable. Es por esto que la puntuación máxima de ζ_d se colocó en 2 puntos.

Para los criterios ζ_r y ζ_c se encontró la siguiente relación. En la tabla 6.1 se muestran tres pares de métodos que funcionan igual, pero se diferencian en la cantidad de electrodos que utilizan. Estos métodos son ESVM_2 con ESVM_3, SSVM_1 con SSVM_2 y CNN-1 con CNN-2a. Uno de los métodos de cada par utiliza 64 electrodos y son los que dan el porcentaje de reconocimiento más alto, los métodos con el desempeño más bajo utilizan ocho electrodos. Al bajar la cantidad de electrodos el porcentaje de reconocimiento decae un 15.5% para los ESVM, un 26.5% para los SSVM y un 7% para los CNN. En promedio para los tres casos el

porcentaje decae 16.33%, lo que es muy cercano a un sexto del 100%. Como la mayoría de los algoritmos utilizan 64 u ocho canales, se propone que la puntuación máxima del criterio de número de canales ξ_c sea un sexto de la puntuación del criterio porcentaje de reconocimiento ξ_r . De esta manera, los algoritmos que son iguales y solo varían la cantidad de canales puedan tener un desempeño similar para la evaluación propuesta, ya que son el mismo algoritmo. No se puede conseguir que se tenga el mismo desempeño, debido a que no en todos los algoritmos cae el porcentaje de la misma manera.

Debido a que la puntuación máxima de ξ_s y ξ_d ya se encuentra definida, es posible encontrar las puntuaciones máximas de ξ_r y ξ_c , las cuales fueron colocados en 78 y 13 respectivamente, debido a que la puntuación máxima de ξ_c tiene que ser un sexto de ξ_r . Con base en estas consideraciones, las ponderaciones para cada uno de los criterios quedan repartidas como se muestra en la tabla 6.2.

Tabla 6.2 Puntuación máxima de los criterios para la evaluación.

Criterio de evaluación	Puntuación máxima
Porcentaje de reconocimiento (ξ_r)	78
Número de sujetos (ξ_s)	7
Número de canales (ξ_c)	13
Tipo de dispositivo (ξ_d)	2
Total (ξ_t)	100

Finalmente, para medir el desempeño de un método tomando en consideración su algoritmo, generalización, modelo de adquisición y la disponibilidad del sistema, se define lo siguiente:

$$\xi_t = \xi_r + \xi_s + \xi_c + \xi_d \quad 6.1$$

donde ξ_t es la puntuación total obtenida para cada método.

6.2.2 Cálculo de puntuación para cada criterio de evaluación

Una vez que se tienen los valores de puntuación máxima de cada uno de los criterios, es necesario definir como se calculará el valor de los criterios para cada método. A continuación, se define como son calculados:

- Cálculo de ξ_r : para el cálculo de este criterio de evaluación se propuso que la puntuación máxima de 78 puntos equivale al 100%. Para aquellos métodos que no tienen un 100% de reconocimiento se tomará la parte proporcional de los puntos. De esta manera el cálculo del criterio de porcentaje de reconocimiento se calcula de la siguiente manera:

$$\xi_r = \frac{78 \cdot \zeta_r}{100} \quad 6.2$$

- Cálculo de ξ_s : se propone que el utilizar diez sujetos corresponda a la puntuación máxima para este criterio y el utilizar un sujeto corresponda a una puntuación de uno. Esto se debe a que el número máximo de sujetos utilizados por los métodos mostrados en la tabla 6.1 es de diez y el mínimo que se podría utilizar es de uno. Por lo tanto, para el cálculo de este criterio se utiliza la siguiente ecuación:

$$\xi_s = \frac{6(\zeta_s - 1)}{9} + 1 \quad 6.3$$

- Cálculo de ξ_c : para este criterio, a menor cantidad de canales utilizados, más alta será la puntuación obtenida. Debido a que existen algunos dispositivos que pueden utilizar hasta 128 canales, y el mínimo número de canales que se podría utilizar es uno. Se propone que cuando se utilizan 128 canales, la puntuación sea uno, y cuando se utiliza un canal, la puntuación sea máxima, es decir 13. Por lo tanto, el cálculo se realiza de la siguiente manera:

$$\xi_c = 14 - \left(\frac{12(\zeta_c - 1)}{127} + 1 \right) \quad 6.4$$

Existen dos métodos donde el número de canales utilizados por sujeto es diferente, definidos mediante algoritmos de optimización. El uso de estos algoritmos ayuda al desempeño, con base en los porcentajes de reconocimiento. Sin embargo, para un sistema BCI esto significa hacer un análisis para cada sujeto de pruebas, con la finalidad de obtener los canales óptimos para cada uno. Esto refleja que los métodos no son capaces de generalizar las señales con distintos sujetos. Por lo tanto, a estos dos métodos se decidió castigarlos como si utilizaran 64 electrodos, ya que para hacer el análisis es necesario colocar todos los electrodos.

- Cálculo de ζ_d : en el caso de este análisis solo se tienen dos dispositivos. Uno de ellos es OpenBCI® y el otro es g.tec. OpenBCI® es aproximadamente 30 veces más barato que g.tec. Además, OpenBCI® es un dispositivo más pequeño que g.tec. Unas de las desventajas de OpenBCI ante g.tec es la cantidad de canales que se pueden utilizar. Sin embargo, este trabajo de tesis busca realizar un producto que sea de fácil acceso para la mayoría de la población y que pueda funcionar con una cantidad mínima de canales, capaz de funcionar fuera de los laboratorios. De esta manera, se propone dar a OpenBCI la puntuación máxima de 2 y a g.tec de 1. La puntuación de g.tec fue definida en función de que aún existen algunos dispositivos más costosos y menos portables que éste.

Una vez que se definió como se calcula la puntuación para cada uno de los criterios de evaluación de los métodos, se procede a realizar la evaluación. Aplicando las ecuaciones 6.1-6.4 y la definición de puntuación de ζ_d , se obtienen los valores mostrados en la tabla 6.3. Además, con base en la puntuación total ζ_t obtenida, se ordenaron los métodos, de tal manera que el que obtuvo la puntuación ζ_t más alta se encuentra en la primera posición del ranking y así sucesivamente con los demás métodos.

Como se puede observar en la tabla 6.3, el método CNN300 tiene la mejor puntuación en el ranking bajo el criterio de evaluación propuesto en este capítulo. Esto se debe a que su porcentaje de reconocimiento es competitivo, utiliza solo tres canales del sistema internacional 10-20, fue probado en diez sujetos y se realiza por medio de un dispositivo relativamente económico y portable.

En segundo lugar, en el ranking, están los métodos ESVM_1, SSVM_1 y BN³. Estos tres métodos obtuvieron el mismo desempeño para el reconocimiento de caracteres. ESVM_1 utiliza una cantidad menor de canales, pero fue castigado debido a que no es capaz de generalizar la cantidad de canales utilizados por sujetos. Además, para el correcto funcionamiento de ESVM_1 se tienen que colocar 64 canales para poder definir cuáles son los que se van a utilizar por cada sujeto.

Tabla 6.3 Puntuación para cada uno de los criterios de evaluación por método.

Método	ζ_r	ζ_s	ζ_c	ζ_d	ζ_t	Ranking
CNN300	66.85	7.00	11.83	2.00	87.68	1
ESVM_1	69.48	1.67	6.54	1.00	78.69	2
SSVM_1	69.48	1.67	6.54	1.00	78.69	2
BN3	69.48	1.67	6.54	1.00	78.69	2
ESVM_2	68.76	1.67	6.54	1.00	77.97	3
MCNN-1	68.76	1.67	6.54	1.00	77.97	3
FFp300	57.08	7.00	11.83	2.00	77.91	4
CNN-1	68.04	1.67	6.54	1.00	77.25	5
MCNN-3	68.04	1.67	6.54	1.00	77.25	5
CNN-2a	63.00	1.67	11.39	1.00	77.06	6
CNN-2b	62.64	1.67	11.39	1.00	76.70	7
Modificación ESVM	66.60	1.67	6.54	1.00	75.81	8
mLVQ	65.88	1.67	6.54	1.00	75.09	9
MCNN-2	65.16	1.67	6.54	1.00	74.37	10
ICA	64.44	1.67	6.54	1.00	73.65	11
CNN-3	63.72	1.67	6.54	1.00	72.93	12
ESVM_3	57.60	1.67	11.39	1.00	71.66	13
SSVM_2	50.40	1.67	11.39	1.00	64.46	14

SSVM_2 es el método con el puntaje menor en el ranking. Este algoritmo utiliza ocho canales, pero tiene el porcentaje más bajo de reconocimiento de carácter con 70%.

FFp300 quedó en el cuarto lugar del ranking. Esto se debe a que presenta un porcentaje de reconocimiento bajo. Como se vio en el capítulo 3, el motivo por el cual exhibe un desempeño

bajo es que no es capaz de generalizar a los sujetos de pruebas, debido a que solo obtuvo resultados competentes para aquellos sujetos que su señal P300 se encontraba cerca de los 300 ms.

En conclusión, el método CNN300 propuesto en este trabajo de tesis hace factible el desarrollo de sistemas BCI basados en P300 con un buen desempeño, y baja cantidad de electrodos en comparación con el resto de los métodos propuestos en la literatura. Otra ventaja de este método es que puede generalizar las señales P300 de diferentes sujetos ya que está comprobada su reproducibilidad mediante la aplicación en diez sujetos de pruebas.

CAPÍTULO 7. RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se propuso un sistema BCI que analiza la señal P300 en un deletreador y terapia virtual. Las señales EEG fueron adquiridas con electrodos de copa y utilizando los canales O1, Oz y O2, por medio del dispositivo OpenBCI®. Para procesar las señales P300 se propusieron dos métodos de extracción de características y clasificación. Estos métodos se denominaron FFp300 y CNN300. FFp300 extrae características basándose en la amplitud de la señal P300 y clasifica mediante una red neuronal artificial de tipo *feedforward*. CNN300 extrae las características y clasifica por medio de una red neuronal convolucional.

Para generar las señales P300 de entrenamiento y prueba, se utilizó un deletreador de Donchin modificado. La razón por la cual se modificó el paradigma original del deletreador de Donchin, se debe a que la señal P300 se presentó con mayor amplitud cuando las intensificaciones que la generaban se daban en las columnas. De esta manera, la modificación consiste en solo realizar intensificaciones por columnas, y para poder detectar donde se encuentra el carácter correcto, la matriz se transpone a la mitad de la época de deletreo.

Los resultados obtenidos por medio de FFp300 mostraron que es un método con desempeño aceptable cuando las señales P300 se localizan cerca de los 300 ms. Sin embargo, si la señal p300 se adelanta o se atrasa, FFp300 puede tener problemas en la detección. Por lo tanto, este método tiene una menor generalización de las señales P300 que los métodos utilizados en la evaluación.

En el caso de CNN300, los resultados fueron mejores que FFp300 ya que puede detectar la señal P300 sin importar el tiempo en que sucede la señal P300. Esto se debe a que fue capaz de generalizar las señales de todos los sujetos de pruebas utilizados. Además, obtuvo un resultado competitivo contra los métodos reportados en la literatura.

Debido al buen desempeño de CNN300, este método fue aplicado en la terapia virtual “Frutas”. Esta terapia utiliza la detección de la señal P300 para poder seleccionar objetos dentro de un escenario que simula una situación real. Además, para esta terapia se evaluó el desempeño de CNN300 reduciendo la cantidad de intensificaciones. Este análisis demostró que para algunos

sujetos de pruebas el reducir la cantidad de intensificaciones por objeto es factible, y hace que la selección de objetos se realice de una manera más rápida.

Respecto a la comparación de FFp300 y CNN300 contra algunos de los métodos más populares en la literatura se puede comentar lo siguiente. FFp300 no pudo generalizar los resultados en todos los sujetos de prueba, por lo cual obtuvo el lugar 13 en el ranking por porcentaje de reconocimiento. Sin embargo, FFp300 utiliza una cantidad menor de electrodos, y fue probado con más sujetos que el resto de los métodos propuestos. Para CNN300 se tiene un desempeño competitivo, colocándose en el cuarto lugar del ranking de trabajos realizados. Adicional a esto, utiliza una cantidad mucho menor de electrodos, cerca de 21 veces menos electrodos que los algoritmos que obtuvieron ligeramente mayor desempeño. Además, es capaz de generalizar las señales P300 de más sujetos de pruebas.

Debido a que, en este trabajo de tesis, se presenta un sistema de interfaz cerebro computadora para la detección de la señal P300, y no solo los algoritmos de extracción de características y clasificación, se propuso un criterio de evaluación para los sistemas BCI, donde, se toman en cuenta diferentes factores. Este criterio se basó en el porcentaje de reconocimiento de caracteres correctos, el número de canales utilizados, la cantidad de sujetos donde se probó y el dispositivo utilizado para adquirir las señales EEG. Bajo el criterio evaluación propuesto, CNN300 fue el que obtuvo el mejor desempeño debido a que es un método con las siguientes características:

- Un porcentaje de reconocimiento competitivo.
- Se utilizan solo tres canales del sistema internacional 10-20.
- Se comprobó su reproducibilidad con diez sujetos de pruebas.
- Las señales fueron adquiridas con un dispositivo portable y de bajo costo.

El método CNN300 fue diseñado para funcionar en un equipo de cómputo genérico como estaciones de trabajo, computadora de escritorio, laptop, o tableta con Windows®. Sin embargo, uno de los enfoques de la tecnología hoy en día es realizar cada vez dispositivos más pequeños, lo cual es posible ya que la capacidad de procesamiento es cada vez mayor. Por lo tanto, como trabajo futuro se propone la implementación de este sistema BCI en una plataforma embebida.

Para la colocación de los electrodos se utilizó una banda elástica. Ahora que está definido cuales electrodos se utilizan, es posible diseñar algún soporte, como una diadema que sea útil para la colocación de los electrodos utilizados. Además, se puede continuar la investigación, de tal manera que se optimice el método CNN300, con la finalidad de reducir al mínimo la cantidad de canales, pudiendo de esta manera detectar la señal P300 con un solo canal.

REFERENCIAS

- [1] S. N. Abdulkader, A. Atia, y M.-S. M. Mostafa, “Brain computer interfacing: Applications and challenges”, *Egypt. Informatics J.*, vol. 16, núm. 2, pp. 213–230, 2015.
- [2] J. Malmivuo y R. Plonsey, *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*, Primera. New York: Oxford University Press, 1995.
- [3] M. Teplan, “Fundamentals of EEG measurement”, *Measurement Science Review*, vol. 2, pp. 1–11, 2002.
- [4] L. F. Nicolas-Alonso y J. Gomez-Gil, “Brain computer interfaces, a review”, *Sensors*, vol. 12, núm. 2, pp. 1211–1279, 2012.
- [5] D. Marshall, D. Coyle, S. Wilson, y M. Callaghan, “Games, gameplay, and BCI: The state of the art”, *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, núm. 2, pp. 82–99, 2013.
- [6] E. A. Larsen y A. I. Wang, “Classification of EEG Signals in a Brain - Computer Interface System”, M.C Tesis, Norwegian University of Science and Technology, 2011.
- [7] “EMOTIV - Brainwear Wireless EEG Technology”, 2016. [En línea]. Disponible en: <https://www.emotiv.com/>. [Consultado: 12-jun-2016].
- [8] “EEG - ECG - Biosensors”, 2016. [En línea]. Disponible en: <http://neurosky.com/>. [Consultado: 12-jun-2016].
- [9] “g.tec - Guger Technologies”. [En línea]. Disponible en: <http://www.gtec.at/>. [Consultado: 12-jun-2016].
- [10] “OpenBCI - Open Source Biosensing Tools (EEG, EMG, EKG, and more)”, 2016. [En línea]. Disponible en: <http://openbci.com/>. [Consultado: 12-jun-2016].
- [11] A. Momenzhad, M. Shamsi, y H. Ebrahimzhad, “Classification of EEG-P300 Signals Extracted from Brain Activities in BCI Systems Using ν -SVM and BLDA Algorithms”, vol. 34, núm. 2, pp. 23–35, 2014.
- [12] S. Theodoridis y K. Koutroumbas, *Pattern Recognition*, Segunda ed. San Diego: Elsevier, 2003.
- [13] A. A. Khodaskar y S. A. Ladhake, “Pattern Recognition: Advanced Development , Techniques and Application for Image Retrieval”, en *2014 International Conference on*

- Communication and Network Technologies (ICCNT)*, 2014, pp. 74–78.
- [14] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 53, núm. 9. 2006.
- [15] S. Sur y V. K. Sinha, “Event-related potential: An overview”, *Ind. Psychiatry J.*, vol. 18, núm. 1, pp. 70–73, 2009.
- [16] A. Rakotomamonjy y V. Guigue, “BCI competition III: Dataset II- ensemble of SVMs for BCI P300 speller”, *IEEE Trans. Biomed. Eng.*, vol. 55, núm. 3, pp. 1147–1154, 2008.
- [17] H. Cecotti y A. Gr, “Convolutional neural networks for P300 Detection with Application to Brain-Computer Interfaces”, vol. 33, núm. 3, pp. 433–445, 2011.
- [18] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, y J. R. Wolpaw, “Toward enhanced P300 speller performance”, *J. Neurosci. Methods*, vol. 167, núm. 1, pp. 15–21, 2008.
- [19] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, y J. R. Wolpaw, “BCI2000: A general-purpose brain-computer interface (BCI) system”, *IEEE Trans. Biomed. Eng.*, vol. 51, núm. 6, pp. 1034–1043, 2004.
- [20] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, y F. Yang, “BCI competition 2003 - Data set IIb: Enhancing P300 wave detection using ICA-based subspace projections for BCI applications”, *IEEE Trans. Biomed. Eng.*, vol. 51, núm. 6, pp. 1067–1072, 2004.
- [21] E. Donchin y L. A. Farwell, “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials”, *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, núm. 6, pp. 510–523, 1988.
- [22] A. Arasteh, M. H. Moradi, y A. Janghorbani, “A Novel Method Based on Empirical Mode Decomposition for P300-Based Detection of Deception”, vol. 11, núm. 11, pp. 2584–2593, 2016.
- [23] A. F. Perez Vidal, M. A. Oliver Salazar, y G. Salas Lopez, “Development of a Brain-Computer Interface Based on Visual Stimuli for the Movement of a Robot Joints”, *IEEE Lat. Am. Trans.*, vol. 14, núm. 2, pp. 477–484, 2016.
- [24] H. Wang, Y. Zhang, N. R. Waytowich, D. J. Krusienski, G. Zhou, J. Jin, X. Wang, y A. Cichocki, “Discriminative Feature Extraction via Multivariate Linear Regression for SSVEP-based BCI”, *Neural Syst. Rehabil. Eng. IEEE Trans.*, vol. 24, núm. 99, pp. 532–541, 2016.

- [25] B. Blankertz, “Documentation Wadsworth BCI Dataset (P300 Evoked Potentials) - BCI Competition III Challenge 2004”, en *Proceedings of the BCI Classification Contest*, 2004.
- [26] L. Madrid, J. A. Ramírez, y M. I. Chacón, “Extracción de características y clasificación de la onda P300 mediante un canal en el deletreador de donchin”, en *ELECTRO 2017*, 2017, vol. 39, pp. 83–88.
- [27] C. Lindig y O. Yáñez-Suárez, “Optimized Detection of the Infrequent Response in P300-Based Brain-Computer Interfaces”, *Rev. Mex. Ing. Biomédica*, vol. 34, núm. 1, pp. 53–69, 2013.
- [28] U. Hoffmann, J. M. Vesin, T. Ebrahimi, y K. Diserens, “An efficient {P300}-based brain-computerinterface for disabled subjects”, *J. Neurosci. Methods*, vol. 167, núm. 1, pp. 115–125, 2008.
- [29] A. Riccio, L. Simione, F. Schettini, A. Pizzimenti, M. Inghilleri, M. O. Belardinelli, D. Mattia, y F. Cincotti, “Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis”, *Front. Hum. Neurosci.*, vol. 7, núm. 732, pp. 1–9, 2013.
- [30] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, y G. Edlinger, “How many people are able to control a P300-based brain-computer interface (BCI)?”, *Neurosci. Lett.*, vol. 462, núm. 1, pp. 94–98, 2009.
- [31] “RFDuino”, 2017. [En línea]. Disponible en: <http://www.rfduino.com/>.
- [32] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. D. R. Millán, M. Schröder, y N. Birbaumer, “The BCI competition III: Validating alternative approaches to actual BCI problems”, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, núm. 2, pp. 153–159, 2006.
- [33] C. C. Postelnicu y D. Talaba, “P300-based brain-neuronal computer interaction for spelling applications”, *IEEE Trans. Biomed. Eng.*, vol. 60, núm. 2, pp. 534–543, 2013.
- [34] W. Pritchard, “The psychophysiology of P300”, *Psychol. Bull.*, vol. 89, pp. 506–540, 1981.
- [35] J. G. Proakis y D. G. Manolakis, *Tratamiento digital de señales*, 4ta Ed. Boston: PRENTICE HALL, 2007.
- [36] C. Parker Anthony y G. A. Thibodeau, *Anatomía y Fisiología*. D.F, México, 1983.

- [37] L. H. Tsoukalas y R. E. Uhrig, *Fuzzy and neural approaches in engineering*. Wiley, 1997.
- [38] M. F. Møller, “A scaled conjugate gradient algorithm for fast supervised learning”, *Neural Networks*, vol. 6, núm. 4, pp. 525–533, 1993.
- [39] R. Fletcher y C. M. Reeves, “Function minimization by conjugate gradients”, *The Computer Journal*, vol. 7, núm. 2. pp. 149–154, 1964.
- [40] R. Swarnkar, A. G. Keskar, P. M. S. Prasad, y N. C. Shivprakash, “A new approach to detect P300 in a single trial based on PCA and SVM classifier”, en *2016 IEEE Region 10 Symposium (TENSYP)*, 2016, pp. 355–360.
- [41] J. Lemley, S. Bazrafkan, y P. Corcoran, “Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision.”, *IEEE Consum. Electron. Mag.*, vol. 6, núm. 2, pp. 48–56, 2017.
- [42] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, y J. Schmidhuber, “Flexible, high performance convolutional neural networks for image classification”, en *IJCAI International Joint Conference on Artificial Intelligence*, 2011, pp. 1237–1242.
- [43] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Massachusetts: The MIT Press, 2012.
- [44] K. Shi, N. Gao, Q. Li, y O. Bai, “A P300 brain-computer interface design for virtual remote control system”, en *2017 3rd IEEE International Conference on Control Science and Systems Engineering (ICCSSE)*, 2017, pp. 326–329.
- [45] L. M. McCane, S. M. Heckman, D. J. McFarland, G. Townsend, J. N. Mak, E. W. Sellers, D. Zeitlin, L. M. Tenteromano, J. R. Wolpaw, y T. M. Vaughan, “P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs. age-matched controls”, *Clin. Neurophysiol.*, vol. 126, núm. 11, pp. 2124–2131, 2015.
- [46] V. Bhatnagar, N. Yede, R. S. Keram, y R. K. Chaurasiya, “A modified approach to ensemble of SVM for P300 based brain computer interface”, en *2016 International Conference on Advances in Human Machine Interaction, HMI 2016*, 2016.
- [47] Y. Wang, J. Shen, J. Liang, y Y. Ji, “Research of P300 feature extraction algorithm based on ICA and wavelet transform”, en *Proceedings - 2014 6th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2014*, 2014, pp. 41–45.

- [48] M. Liu, W. Wu, Z. Gu, Z. Yu, F. Qi, y Y. Li, “Deep learning based on Batch Normalization for P300 signal detection”, *Neurocomputing*, p. to be published, 2017.
- [49] N. Liang y L. Bougrain, “Averaging techniques for single-trial analysis of oddball event-related potentials”, en *4th International Brain-Computer Interface workshop*, 2008, pp. 1–6.
- [50] F. B. Vialatte, M. Maurice, J. Dauwels, y A. Cichocki, “Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives”, *Prog. Neurobiol.*, vol. 90, núm. 4, pp. 418–438, 2010.