

Discriminador binario de imaginación visual a partir de señales EEG basado en redes neuronales convolucionales

Fabio Ricardo Llorella^{a, *}, Eduardo Iáñez^b, José María Azorín^b, Gustavo Patow^c

^aViRVIG, Universitat de Girona, Girona, Cataluña, España

^bBrain-Machine Interface Systems Lab, Universidad Miguel Hernández, Elche, Valencia, España

^cViRVIG, Universitat de Girona, Girona, Cataluña, España

To cite this article: Llorella, F. R., Iáñez, E., Azorín, J. M., Patow, G. 2022. Binary visual imagery discriminator from EEG signals based on convolutional neural networks. Revista Iberoamericana de Automática e Informática Industrial 19, 108-116. <https://doi.org/10.4995/riai.2021.14987>

Resumen

Las interfaces cerebro-máquina (Brain-Computer Interface, BCI, en inglés) son una tecnología que permite la comunicación directa entre el cerebro y el mundo exterior sin necesidad de utilizar el sistema nervioso periférico. La mayoría de sistemas BCI se centran en la utilización de la imaginación motora, los potenciales evocados o los ritmos corticales lentos. En este trabajo se ha estudiado la posibilidad de utilizar la imaginación visual para construir un discriminador binario (brain-switch, en inglés). Concretamente, a partir del registro de señales EEG de siete personas mientras imaginaban siete figuras geométricas, se ha desarrollado un BCI basado en redes neuronales convolucionales y en la densidad de potencia espectral en la banda α (8-12 Hz), que ha conseguido distinguir entre la imaginación de una figura geométrica cualquiera y el relax, con un acierto promedio del 91 %, con un valor Kappa de Cohen de 0.77 y un porcentaje de falsos positivos del 9 %.

Palabras clave: Discriminador binario, interfaz cerebro-máquina, red neuronal convolucional, densidad potencia espectral, EEG.

Binary visual imagery discriminator from EEG signals based on convolutional neural networks

Abstract

A Brain-Computer Interface (BCI) is a technology that allows direct communication between the brain and the outside world without the need to use the peripheral nervous system. Most BCI systems focus on the use of motor imagination, evoked potentials, or slow cortical rhythms. In this work, the possibility of using visual imagination to construct a binary discriminator has been studied. EEG signals from seven people have been recorded while imagining seven geometric figures. Using convolutional neural networks it has been possible to distinguish between the imagination of a geometric figure and relaxation with an average success rate of 91 % with a Cohen kappa value of 0.77 and a percentage of false positives of 9 %.

Keywords: Brain-switch, visual imagery, convolutional neuronal network, power spectral density, EEG.

1. Introducción

Las interfaces cerebro-máquina (Brain-Computer Interface, BCI, en inglés) permiten crear un canal de comunicación directo entre el cerebro y el mundo exterior, sin necesidad de utilizar el sistema nervioso periférico (Fernando et al., 2012), posibi-

litando de esta forma que personas con problemas de movilidad puedan interactuar con el entorno que les rodea. La base de esta tecnología es el registro y procesamiento de la actividad cerebral. La electrocorticografía (ECoG) y la electroencefalografía (EEG) son los métodos más comúnmente utilizados para el registro, siendo el primero invasivo y el segundo no invasi-

*Autor para correspondencia: fricardocorreo@gmail.com

vo (Reza et al., 2018). Sin embargo, en las tecnologías invasivas se debe realizar una operación quirúrgica para poder introducir un conjunto de electrodos en la corteza cerebral, lo que conlleva un alto riesgo para la salud. En cambio, en las tecnologías no invasivas no se debe realizar ninguna intervención, reduciendo drásticamente las complicaciones que se puedan derivar de su uso.

Las interfaces cerebro-máquina tienen diversas aplicaciones, como neurorehabilitación (Hortal et al., 2015), comunicación (Birbaumer et al., 1999), o videojuegos (Ahn et al., 2014). Todas estas aplicaciones demuestran que las interfaces cerebro-máquina son una realidad y que la comunidad científica está investigando activamente para extender sus posibilidades.

El objetivo básico de cualquier interfaz cerebral basada en señales EEG es la clasificación de dichas señales a partir de unos paradigmas de entrenamiento, es decir, asociar las señales EEG a una actividad concreta, como puede ser imaginar el movimiento de una mano, de los pies (Jiang et al., 2015), o imaginar una palabra (Abhiram et al., 2019). Los paradigmas utilizados son diversos, pero los principales son: la imaginación motora, donde la persona debe imaginar mover las manos, los pies u otra parte del cuerpo; los potenciales evocados (Gong et al., 2020), donde estímulos externos hacen disparar una respuesta del cerebro y por último los ritmos sensorimotrices lentos, utilizados ampliamente para el control (Zhang et al., 2018).

Sin embargo, el ser humano tiene la capacidad de realizar muchas otras tareas cognitivas, más allá de las que hemos enumerado. Entre estas capacidades encontramos la de imaginar objetos, en la que hemos centrado este trabajo. La imaginación visual se puede definir como la capacidad de usar información visual que no procede de la percepción sino de la memoria, tal y como indica el trabajo de (Knauff et al., 2000). La imaginación visual no ha sido muy estudiada en el uso de las interfaces cerebrales, como se explica en el trabajo de (Kosmyna et al., 2018). En el trabajo de (Bobrov et al., 2011) se utilizó la imaginación de caras y de casas, juntamente con el relax, obteniendo que era posible la distinción entre estos tres estados mentales. En el caso del trabajo de (Esfahani and Sundararajan, 2012) se usaron cinco figuras geométricas imaginadas, obteniendo también resultados positivos a la hora de clasificar las distintas figuras. En cambio, en el trabajo de (Kosmyna et al., 2018) no se pudo distinguir entre la imaginación visual de una flor y un martillo, pero sí se pudo distinguir más allá del azar entre una figura imaginada y el relax. Todo ello lleva a pensar que se debe ahondar más en el estudio del uso de la imaginación visual aplicada a las interfaces cerebrales. También se ha estudiado la diferencia entre la imaginación visual y la percepción visual (Xie et al., 2020), mostrando que existen similitudes entre ambas categorías en la banda α .

En el presente trabajo se ha estudiado la posibilidad de utilizar la imaginación visual para crear un discriminador binario basado en señales EEG y de esta forma poder ampliar las aplicaciones en las que se puedan utilizar las interfaces cerebrales, siendo este trabajo uno de los primeros en abordar el estudio del uso de la imaginación visual para ser utilizada en un discriminador binario.

El interés por desarrollar un codificador binario basado en un clasificador multiclase radica en su simplicidad, tanto de uso como de construcción, ya que cuantas más clases se deban clasificar, más difícil se hace para el clasificador poder acertar correctamente y proporcionar al usuario una buena experiencia de uso. Pero también se debe destacar que, mediante una clasificación binaria entre imaginación visual y el estado de relax, podemos estudiar con más facilidad qué frecuencias son predominantes con respecto a las del relax.

2. Materiales y métodos

2.1. Participantes

Para realizar el análisis de la imaginación visual, se ha registrado la actividad cerebral mediante señales EEG de 7 personas (2 mujeres y 5 hombres), con una media de edad de 31 años. Todas las personas han participado de forma voluntaria. Ninguna de las personas que han participado en el experimento tenía problemas de visión y no había usado nunca una interfaz cerebral. Los participantes fueron elegidos al azar entre un grupo de participantes de entre 18 y 40 años, todos personal relacionado en el ámbito universitario.

El número de usuarios utilizado en experimentos BCI no es un tema resuelto, si bien se considera que lo normal es entre 10 y 20 sujetos (Melnik et al., 2017), también es cierto que no siempre se puede alcanzar dicha cantidad de sujetos, y existen trabajos con una cantidad de sujetos menores (Planelles et al., 2014; Anderson and Sijercic, 1996). En nuestro trabajo hemos podido registrar señales EEG a siete personas, que es una cantidad que se encuentra dentro los parámetros aceptados dentro la investigación de sistemas BCI.

2.2. Dispositivo de registro

El dispositivo utilizado para registrar las señales EEG ha sido el dispositivo *g.NAUTILUS* de la empresa *g.TEC*, que consta de ocho electrodos húmedos. La frecuencia de muestreo ha sido de 250 Hz con una resolución de 16 bits, y se ha aplicado un filtro Notch centrado en 50 Hz para eliminar la influencia de la línea eléctrica.

Los electrodos utilizados (siguiendo el sistema internacional 10-10) (Jurcak et al., 2007) son Oz, Pz, P3, P4, PO3, PO4, PO7 y PO8, y como electrodo de referencia se ha utilizado el electrodo AFz (Figura 1). Estos electrodos son los más aconsejables para discriminar la imaginación visual de objetos según se desprende del trabajo de (Bobrov et al., 2011). A la hora de realizar los registros de las señales EEG, la impedancia de los electrodos era inferior a los 10K Ω .

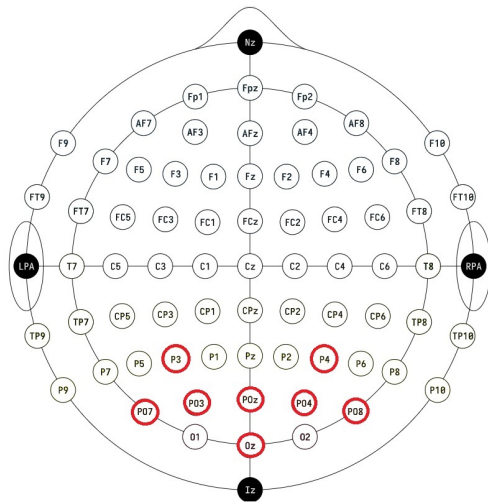


Figura 1: Electrodo del sistema internacional 10-10. Los círculos en rojo representan los electrodos utilizados en este trabajo. Recuperado del trabajo de (Seeck et al., 2017).

2.3. Protocolo experimental

Las siete personas que han participado en este experimento se sentaron en una silla delante de un monitor TFT plano de 17 pulgadas, siendo la distancia entre la persona y el monitor de 85 cm (Figura 2). La sala donde se realizaron los registros estaba iluminada con una luz tenue para no deslumbrar a los participantes. En el momento del registro de las señales EEG las personas se encontraban solas en la sala, para evitar que fueran molestadas. Además, se indicó a las personas que realizaban el experimento que intentaran no moverse durante el proceso de imaginación visual, para disminuir de esta forma los artefactos musculares que se pudiesen producir, contaminando de esta forma las señales EEG.

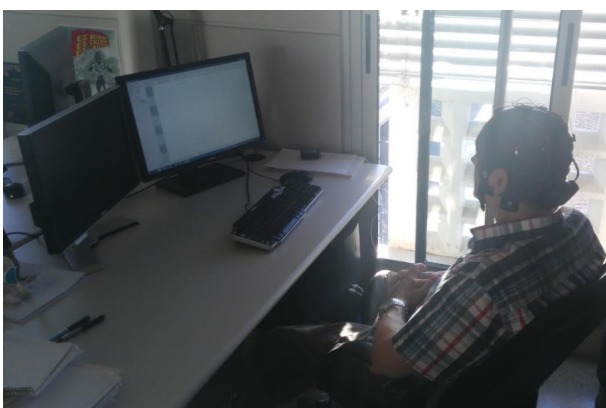


Figura 2: Entorno experimental.

Cada persona ha registrado en un mismo día un total de 20 sesiones, donde cada sesión consta de 14 pruebas (dos pruebas para cada figura geométrica), teniendo por tanto un total de 40 pruebas por figura geométrica. Se han utilizado siete figuras geométricas (Figura 3), triángulo, círculo, cuadrado, hexágono, línea recta, paralelogramo y pentágono.

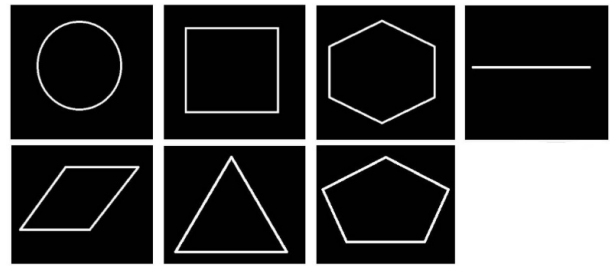


Figura 3: Imágenes de las figuras utilizadas en el experimento.

Cada prueba empieza mostrando una cruz blanca sobre fondo negro, durante un segundo. A continuación aparece, de forma aleatoria, una de las figuras geométricas indicada, que estará en pantalla durante dos segundos. Al terminar los dos segundos, la figura desaparece de la pantalla y es cuando la persona debe imaginar la figura que se le ha indicado, durante cinco segundos. En la Figura 4 se puede observar de forma esquemática el protocolo utilizado. El protocolo que hemos diseñado ha tomado como referencia los trabajos de (Kosmyna et al., 2018) y (Bobrov et al., 2011).

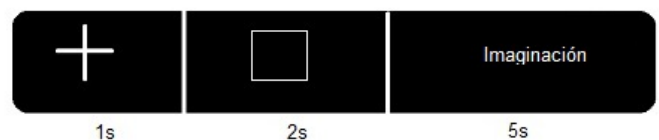


Figura 4: Protocolo de registro.

3. Procesamiento de los datos

3.1. Pre-procesamiento de las señales EEG

Las pruebas han sido segmentadas en fragmentos de un segundo, de la misma manera que en el trabajo de (Bobrov et al., 2011), separando los segmentos correspondientes a la línea base, es decir cuando el usuario no está imaginando nada, y los de imaginación visual. Luego hemos procedido a eliminar aquellos canales que presentaran una desviación estándar superior a dos veces el valor promedio del resto de canales (Pedroni et al. (2019)), para eliminar los que puedan tener artefactos que distorsionan la señal (Bigdely-Shamlo et al., 2015). Por otra parte, se han filtrado las señales EEG en el rango 1 a 30 Hz, ya que mantiene el rango α .

3.2. Extracción de características

Las señales EEG contienen distintas frecuencias, y es importante poder observar como se distribuyen dichas frecuencias y cuál es su potencia para poder clasificar distintos estados cognitivos de las señales EEG, ya que es difícil ver características importantes en el dominio del tiempo. En cambio, observar la distribución frecuencial puede facilitarnos dicha tarea. La densidad de potencia espectral ha sido ampliamente utilizada en diversos estudios de interfaces cerebrales, por lo que, para caracterizar las señales EEG, se ha optado por utilizar la densidad de potencia espectral mediante el método de Welch (Proakis et al., 1996), que ha obtenido muy buenos resultados en el pasado (Kim et al., 2018; Aggarwal and Chugh, 2019). Una de las

ventajas de esta técnica es que es de fácil integración al disponer de librerías que la implementan. La librería utilizada en este trabajo ha sido *SciPy* versión 1.5.3 para *Python* 3.5. El método de Welch divide la señal en bloques no solapados sucesivos, y de cada bloque se extrae el periodograma y se promedian los resultados (1):

$$x_m(n) = w(n)x(n + mR), n = 0, 1, \dots, M - 1, m = 0, 1, \dots, K - 1 \quad (1)$$

donde R indica el tamaño de la ventana temporal utilizado y K en número de ventanas. El periodograma para cada ventana se calcula como indica la (2).

$$P_{x,M}(w_k) = \frac{1}{M} |FFT_{N,k}(x_m)|^2 \quad (2)$$

Donde FFT es la transformada rápida de Fourier. Por último, la densidad de potencia espectral se obtiene a partir de (3):

$$S_x^W(w_k) = \frac{1}{K} \sum_{m=0}^{K-1} P_{x_m,M}(w_k) \quad (3)$$

Una vez hemos obtenido la densidad de potencia espectral, se puede obtener la potencia de la banda alfa mediante (4):

$$PSD_\alpha = S_x^W([8, 12]) \quad (4)$$

Se ha optado por utilizar la banda α ya que diversos estudios han mostrado que esta banda es una banda importante a la hora de procesar la información visual (Kosmyrna et al., 2018; Xie et al., 2020).

4. Clasificación

Una vez tenemos las características de las señales, se pasa a la etapa de clasificación. Se han utilizado distintos clasificadores, de los llamados *clásicos*, como Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) y K Nearest Neighbors (KNN). También se ha testado el uso de redes neuronales convolucionales (CNN), juntamente con un algoritmo genético para encontrar la estructura más adecuada para dichas redes. Para el uso de los clasificadores se ha utilizado la librería de python *Sklearn*. Dicha librería implementa los tres clasificadores clásicos arriba mencionados. Para la implementación de las redes CNN se ha utilizado la librería *Keras*, versión 2.2.4. Los parámetros utilizados para el clasificador KNN han sido de $k = 3$, y el kernel utilizado para el clasificador SVM ha sido las Radial Basis Functions, con el resto de los parámetros por defecto. Para saber más sobre los clasificadores clásicos, consultar el trabajo de (Lotte et al., 2007).

4.1. Redes neuronales convolucionales

Las redes neuronales convolucionales (Convolutional Neural Network CNN, en inglés), son un tipo de redes neuronales que incluyen un filtraje convolucional que permite extraer información espacial sobre los datos de entrada. Este tipo de redes ha tenido un gran éxito a la hora de trabajar con imágenes (Gavali and Banu, 2019), y actualmente existe una gran variedad de estudios que aplican redes CNN a la clasificación de señales EEG (Craik et al., 2019).

Una de las partes más importantes en las redes CNN es la operación de convolución. Esta operación transforma dos funciones en una tercera función que representa la superposición de las dos funciones de entrada:

$$f[m] * g[m] = \sum_{-\infty}^{+\infty} f(n)g(t - n)dn \quad (5)$$

Esta operación permite crear un mapa de características que será utilizado por una red neuronal para clasificar dichas características.

Las redes CNN constan de muchas otras capas, como las capas de entrada, capas de reducción de dimensionalidad, que se suelen aplicar después de las capas convolucionales, y por último las capas de salida (Figura 5). Las capas convolucionales tienen el problema que el mapa de salida de características es sensible a la posición de los píxeles de la imagen, es decir, un pequeño cambio en la posición produce un mapa de características distinto. Una de las posibles soluciones es utilizar capas de reducción, que progresivamente reducen el tamaño espacial de los mapas de características producidas por las capas convolucionales.

La entrada a nuestra red CNN es una matriz que representa la potencia espectral de la señal EEG. Esta matriz será pasada por sucesivas capas que realizarán una operación de convolución sobre la entrada; y por capas de reducción, que disminuirán la dimensión de la entrada, y por último, se procederá a la clasificación mediante una red neuronal (figura 5).

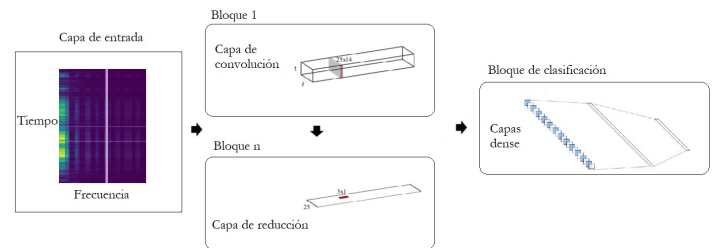


Figura 5: Esquema de la estructura de una red CNN para el procesamiento de las matrices de entrada, que representan la potencia espectral de las señales EEG.

Uno de los problemas que debe afrontar el investigador cuando usa las redes CNN es saber qué estructura debe tener la red CNN, es decir, cuántas capas convolucionales, el tamaño de los filtros, el número de neuronas por capa, qué funciones de activación debemos utilizar, etc. Muchas veces el investigador utiliza su conocimiento previo del problema para ir creando las redes CNN e ir testeando qué redes son las mejores, pero esta forma de proceder requiere mucho tiempo y no garantiza obtener el mejor resultado. Es por ello que se ha optado por utilizar un algoritmo genético para buscar qué estructura de red CNN es la mejor para cada sujeto.

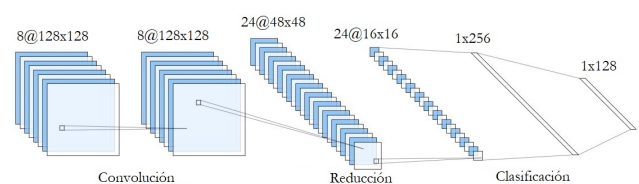


Figura 6: Ejemplo de estructura de una red CNN.

4.2. Algoritmos genéticos

Los algoritmos genéticos son un tipo de algoritmos heurísticos inspirados en la teoría de la evolución de Charles Darwin (McCall, 2005), que permiten encontrar soluciones óptimas en espacios de búsqueda n -dimensionales. Los algoritmos genéticos se suelen utilizar cuando no conocemos que procedimiento es el mejor para encontrar la mejor respuesta a un problema de búsqueda, pero si conocemos el mecanismo para saber si una respuesta es mejor que otra (Algoritmo 1).

Los individuos de los algoritmos genéticos, representan las posibles soluciones, es decir, las posibles estructuras de redes CNN. Estos individuos son evaluados todos con los mismos datos, ya que los datos de las señales EEG se han dividido en dos grupos, uno de entrenamiento y otro de validación.

Dada una función de calidad que deberá ser optimizada, se crean aleatoriamente un conjunto de n posibles individuos que representan una solución, y aplicaremos la función a optimizar para saber qué tan bueno es. Una vez tenemos los valores de la función para cada solución, pasaremos a elegir los m mejores individuos y aplicaremos un conjunto de operaciones para generar nuevos individuos a partir de los individuos seleccionados. De esta forma se estará transmitiendo información entre soluciones con el objetivo de llegar a una solución óptima que satisfaga los requisitos de nuestro problema. En primer lugar debemos definir la función a optimizar y la representación de nuestras soluciones. Para la representación de los individuos existen distintas técnicas, si bien se utilizan cadenas binarias, o de números naturales o reales.

Sobre las posibles soluciones se suelen aplicar dos operadores: la recombinación y la mutación. El primer operador tiene como objetivo generar nuevas soluciones a partir de dos soluciones dadas, recombinando la información de las soluciones de entrada. Esto hace que se transmita la información que mejor se adapte al problema. El segundo operador, dada una solución, varía de forma aleatoria una de sus variables. Este operador tiene como finalidad realizar saltos en el espacio de búsqueda y evitar que el algoritmo genético caiga en soluciones locales. A continuación se muestra el pseudocódigo del algoritmo genético utilizado.

Algorithm 1: Algoritmo genético

Result: Estructura CNN óptima

begin

 INICIAR población de forma aleatoria

while no condición de fin **do**

 EVALUAR soluciones de la población

 SELECCIONAR mejores soluciones

 RECOMBINAR mejores soluciones

 MUTAR hijos

 SELECCIONAR supervivientes

 resultado = mejor solución encontrada

return resultado

4.3. Representación de redes CNN

El objetivo de utilizar los algoritmos genéticos en este trabajo es poder encontrar una estructura de red CNN adecuada para poder clasificar correctamente las señales de imaginación visual de una figura geométrica y el relax. Para ello, primero debemos representar las redes CNN de una forma que pueda ser utilizado

por el algoritmo genético. La representación que hemos realizado es la siguiente: las redes CNN están compuestas por dos partes bien diferenciadas, por una parte tenemos las capas de convolución y reducción que crean un mapa de características, y por la otra las capas neuronales dedicadas a aprender sobre el mapa de características. Por lo tanto, hemos codificado las soluciones en un objeto que contiene dos vectores de números enteros, que llamaremos brazos. Cada brazo contiene una cantidad n de genes, que son de dos tipos: tipo convolucional o tipo neuronal. El tipo convolucional codifica una capa convolucional que contiene un vector con la función de activación, la cantidad de filtros, el tamaño de los filtros y por último un valor que indica si hay o no una capa de reducción, y si la hay, qué tamaño tiene. Los genes neuronales codifican la función de activación de la capa, la cantidad de neuronas y el porcentaje de dropout, ver Figura 7.

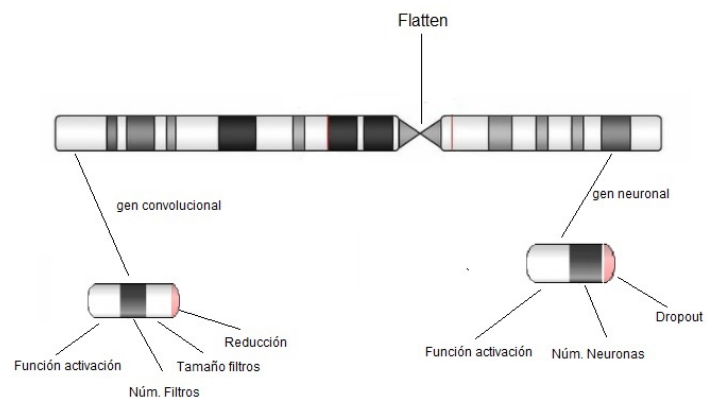


Figura 7: Representación de una red CNN.

En la Figura 6 podemos observar de forma esquemática la representación de las redes CNN. En la izquierda codificamos las capas convolucionales y en la derecha la red neuronal para clasificar. En el medio tenemos una capa de tipo Flatten, que realiza un redimensionamiento de los mapas de características y lo transforma en un vector unidimensional.

4.4. Función de adaptación

La función de adaptación (fitness function, en inglés) (6) es la función que utilizamos para evaluar las redes CNN. En nuestro caso estamos tratando un problema de clasificación binario, por lo tanto, la función de adaptación deberá reflejar la calidad de la red CNN a la hora de clasificar las señales EEG. Para ello se ha utilizado el valor Kappa de Cohen (Altman, 2006). Éste es un valor estadístico que indica si los resultados a la hora de clasificar son causados por el azar o no. El valor de Kappa se encuentra entre el rango $[-1,1]$. Valores cercanos a 0 significa que los resultados se deben al azar, valores cercanos a 1 implican que la clasificación es muy buena, y valores cercanos a -1 implican que la clasificación falla mucho. Este valor se calcula como:

$$k = \frac{p_0 - p_e}{1 - p_e} \quad (6)$$

donde p_0 es el acierto observado y p_e es el acierto teórico, que en nuestro problema es del 50 %.

Para la interpretación de los valores de Kappa mayores a cero se ha utilizado la siguiente escala (Swets, 2014):

- < 0.20 concordancia pobre
- 0.21 - 0.40 concordancia débil
- 0.41 - 0.60 concordancia moderada
- 0.61 - 0.80 concordancia buena
- 0.81 - 1.00 concordancia muy buena

A la hora de calcular la función de adaptación, esto se ha realizado mediante una validación cruzada con un valor de pliegues igual a cinco.

4.5. Recombinación

El operador de recombinación tiene como objetivo crear nuevas soluciones a partir de otras soluciones previas. El operador de recombinación que hemos realizado es el de cruce de los brazos de cada padre, es decir, el primer hijo se formará a partir del brazo convolucional del primer padre y el brazo neuronal del segundo padre, y el segundo hijo se formará a partir del brazo convolucional del segundo padre y del brazo neuronal del primer padre. Ver Figura 8.

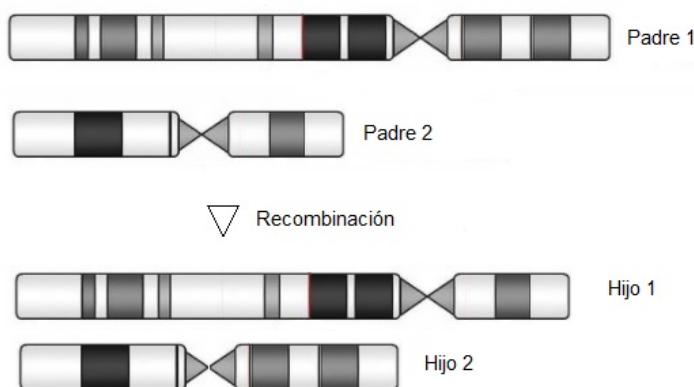


Figura 8: Ejemplo de recombinación.

4.6. Mutación

La mutación es un operador aleatorio, que en nuestro trabajo hemos realizado como una mutación en uno de los brazos. Nuestra mutación se compone de un cuarteto $[m, b, p, v]$, donde el valor m está dentro el rango $[0, 1]$, y si es mayor a la tasa de mutación, entonces se producirá una mutación en el cromosoma. El valor b indica en qué brazo se realizará la mutación, el valor p indica el gen que se mutará y v es el valor que tendrá el gen p .

5. Resultados

El código para procesar las señales EEG se ha escrito en Python 3.5 y se han utilizado las librerías *Scikit-learn*, *Keras* y *TensorFlow*, para construir y evaluar los modelos.

Existen diversos trabajos donde se han implementado sistemas discriminadores binarios, pero prácticamente todos los sistemas se basan en la detección de la imaginación motora (Han

et al., 2020). Para poder comprobar si el clasificador binario que proponemos en este trabajo aporta alguna ventaja, se ha comparado los resultados obtenidos por otros discriminadores. La métrica utilizada para evaluar los distintos modelos ha sido el ratio de positivos verdaderos (true positive rate, TPR, en inglés) (7):

$$TPR = TP / (TP + FP) \quad (7)$$

donde TP indica los positivos verdaderos y FP los falsos positivos. Como falsos positivos se considera cuando el clasificador indica que la señal EEG es una señal asociada a la imaginación de una figura geométrica, pero en realidad se trata de una señal no asociada a la imaginación.

A la hora de reportar los resultados también se han recogido otros indicadores que dan información importante de como se está comportando el discriminador binario, entre ellos los aciertos (accuracy, en inglés) (8):

$$ACC = \frac{TP + TN}{P + N} \quad (8)$$

dónde P son los positivos reales y N los negativos reales. También se calculó el ratio de falsos positivos (false positive rate, en inglés) (9):

$$FPR = \frac{FP}{FP + TN} \quad (9)$$

donde FP son los falsos positivos y TN los verdaderos negativos.

Este último valor es importante a la hora de evaluar un discriminador binario, ya que indica el porcentaje de veces que el discriminador detecta la imaginación de una figura geométrica cuando en realidad no lo es, lo que representa un problema porque, si nuestro sistema se utiliza para activar algún dispositivo, este podría activarse involuntariamente. Es aconsejable mantener el FPR lo más bajo posible.

Para testear los clasificadores LDA, SVM y KNN, se ha utilizado la validación cruzada (k-fold cross fold validation, en inglés) con $k = 5$. Los datos de entrada se dividen en cinco grupos del mismo tamaño, se usan cuatro grupos para crear el modelo y uno para testear, y se repite el proceso cinco veces, cambiando los grupos de modelo y testeo, para luego realizar el promedio para obtener el resultado final. Para la validación de las redes CNN, juntamente con los algoritmos genéticos, se han dividido los datos en dos grupos: un 70 % de los datos de entrada han sido para encontrar la mejor estructura de la red CNN y luego el 30 % restante para validar la estructura encontrada mediante validación cruzada con $k = 4$. Esto se ha repetido n veces con una $n = 5$ y se ha promediado los n resultados obtenidos.

En la Tabla 1 se muestra el TPR de otros trabajos a la hora de clasificar una tarea cognitiva es un discriminador cerebral, viendo que el promedio de TPR es 74 % .

En la primera fase hemos utilizado los clasificadores LDA, SVM y KNN, usando la PSD en la banda α (8-12) Hz para cada canal. En la evaluación de los modelos se ha utilizado 5-fold cross validation.

El intervalo temporal (3-4)s lo hemos desechado ya que el primer segundo puede verse afectado por la retención de la imagen que se muestra a la hora de indicar al usuario que figura debe imaginar. También hemos estudiado que intervalo temporal

Tabla 1: Comparación entre discriminadores binarios (Han et al., 2020)

Tarea	Características	TPR
Finger MI	ERD/ERS	≥ 70.00 %
Foot MI	ERD/ERS	61.40 %
Foot MI	MRCP	71.00 %
Hand or Foot MI	ERD/ERS	64.32 %
Hand MI	RG	91.30 %
Hand MI	CSP	84.46 %
Foot MI	ERD/ERS	62.50 %
Foot MI	MRCP	71.00 %
Foot MI	Sample Entropy	89.50 %
Avg		74.00 % ± 11.60

Tabla 2: Resultados promedio de todas las figuras. Se indica el TPR (%).

Ventana	LDA	SVM	KNN
(4-5)s	63.73	56.60	60.96
(5-6)s	63.33	59.87	61.91
(6-7)s	62.61	58.95	65.04
(7-8)s	62.74	58.57	62.80

de un segundo es el que ofrece mejor resultado, desplazando la ventana un segundo y sin solapamiento. Los resultados de TPR según la ventana observada pueden verse en la Tabla 2.

El mejor resultado se obtiene mediante el clasificador LDA y en la ventana temporal entre 4 y 5 segundos, pero los resultados no son mejores que el promedio obtenido por los discriminadores que se presentan en la Tabla 1.

En la Tabla 3, mostramos los resultados obtenidos al clasificar cada figura contra el relax, utilizando el clasificador LDA en la ventana temporal (4-5) segundos. En la tabla también se reportan distintos indicadores: Acc es el acierto total, TPR es el ratio de positivos verdaderos y FPR el ratio de falsos positivos. Como podemos observar en la Tabla 3, el pentágono es la figura con mejores resultados, obteniéndose un acierto del 68 % y un valor TPR y FPR de 63 % y 27 % respectivamente, mostrando resultados inferiores a los discriminadores binarios observados en el trabajo Han et al. (2020). Es por ello que hemos estudiado si mediante las redes CNN podemos obtener resultados mejores.

Los resultados obtenidos mediante las redes CNN conjuntamente con los algoritmos genéticos para encontrar las mejores estructuras, se muestran en la Tabla 4. Podemos ver que, mediante la ventana de 4 a 5 segundos, obtenemos los mejores aciertos con un TPR de 90 % y un FPR del 9 %. Estos resultados sobresalen comparándolos con las técnicas clásicas y con los discriminadores citados en el trabajo de (Han et al., 2020).

En la tabla 5 mostramos los resultados obtenidos por cada figura utilizando las redes CNN conjuntamente con los algoritmos genéticos en la ventana temporal de entre 4 y 5 segundos. Observando la tabla, se puede ver que el pentágono, juntamente con el cuadrado, son las figuras geométricas con mejores resultados, y que los resultados obtenidos mediante redes CNN son muy superiores a los resultados obtenidos con las técnicas clásicas.

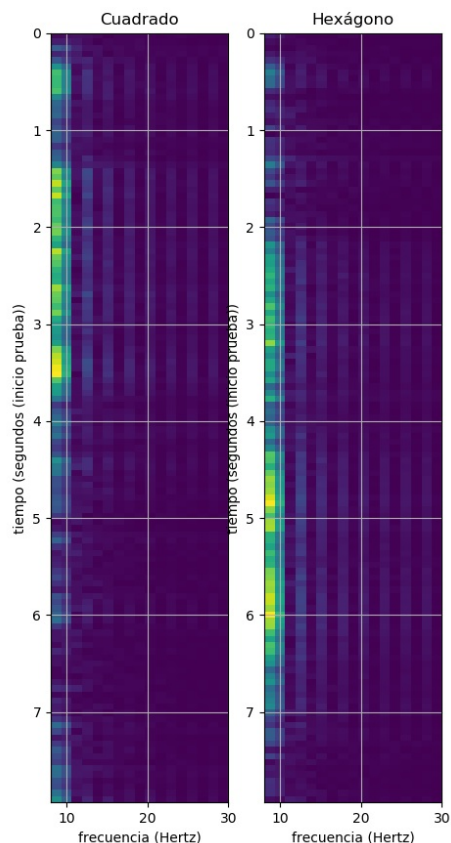


Figura 9: Espectrograma en el canal Oz de las figuras cuadrado y hexágono.

En la figura 9 se puede observar el espectrograma promediando las señales EEG de las figuras cuadrado y hexágono en el electrodo Oz, ya que son las que tiene mayor y peor acierto respectivamente. En el caso del cuadrado se puede observar una disminución de las frecuencias en entre 8 y 10 Hz a partir del segundo 4, que es donde empezaría propiamente la tarea de imaginar la figura geométrica indicada, mientras que en el hexágono dicha disminución no parece producirse con la mis-

Tabla 3: Resultados obtenidos con el clasificador LDA. Se indica el accuracy (Acc), Kappa, la cantidad de aciertos positivos (TPR) y los falsos positivos (FPR)

Figura	Acc(%)	Kappa	TPR(%)	FPR(%)
Triángulo	62±10	0.27	60	35
Círculo	66±9	0.34	65	32
Cuadrado	66±10	0.31	61	30
Pentágono	68±9	0.36	63	27
Línea	68±11	0.37	71	35
Hexágono	68±10	0.38	64	28
Paralelogramo	65±10	0.30	59	28

Tabla 4: Resultados promedio de todas las figuras. Se indica el accuracy (Acc), Kappa, la cantidad de aciertos positivos (TPR) y los falsos positivos (FPR).

Ventana	Acc(%)	Kappa	TPR(%)	FPR(%)
(4-5)s	91±12	0.77	90	9
(5-6)s	90±15	0.76	91	12
(6-7)s	87±17	0.73	85	10
(7-8)s	90±13	0.79	84	4

Tabla 5: Resultados obtenidos con el clasificador CNN + AG. Se indica el accuracy (Acc), Kappa, la cantidad de aciertos positivos (TPR) y los falsos positivos (FPR).

Figura	Acc(%)	Kappa	TPR(%)	FPR(%)
Triángulo	95±10	0.90	91	0
Círculo	92±16	0.86	100	17
Cuadrado	96±8	0.92	91	0
Pentágono	96±8	0.92	92	0
Línea	88±10	0.75	92	17
Hexágono	85±20	0.53	90	20
Paralelogramo	85±12	0.52	75	8

ma intensidad que en el caso del cuadrado. Esta disminución en la banda α podría explicar que, con una figura geométrica, se obtengan mejores resultados.

6. Discusión

En este trabajo se ha llegado a la conclusión que el uso de redes convolucionales, juntamente con la densidad de potencia espectral en la banda α , permite que todos los sujetos que han participado en este estudio puedan utilizar cualquier figura geométrica de las siete propuestas para ser utilizada como un interruptor binario, con un acierto promedio del 91%. En general, las interfaces cerebro-computador no suelen generalizarse para un amplio grupo de sujetos, sino que se buscan ciertas configuraciones de electrodos o técnicas que son comunes para varios de ellos, pero siempre personalizando cada modelo para cada usuario particular. Poder generalizar un modelo para cualquier sujeto es un campo de estudio abierto en interfaces cerebrales. Este problema se conoce como problema de transferencia de conocimiento y se suele llevar a cabo mediante redes neuronales profundas. En un futuro se debería estudiar si se puede crear un modelo discriminador basado en redes CNN genéricas, es decir, una estructura que pueda ser utilizada por cualquier sujeto.

7. Conclusiones

Se ha demostrado que es posible utilizar la imaginación visual de las figuras geométricas utilizadas en este trabajo para

crear un interfaz cerebro-máquina basado en la detección de la imaginación visual, y que la mejor técnica es utilizando redes neuronales convolucionales ya que ofrecen un rendimiento off-line muy superior a los clasificadores clásicos como LDA, SVM y KNN.

Como trabajo futuro sería interesante continuar investigando el uso de la imaginación visual en sistemas cerebro-máquina, más concretamente, investigar si es posible distinguir entre distintas figuras imaginadas y utilizar nuevos objetos imaginados e implementar el sistema propuesto en este trabajo en online. También sería interesante estudiar la generalización de los modelos para que se pueda utilizar un discriminador binario de forma general para cualquier sujeto.

8. Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto TIN2017-88515-C2-2-R del Ministerio de Economía y Competitividad.

Referencias

- Abhiram, Singh, A., Gumaste, 2019. Decoding imagined speech and computer control using brain waves. arXiv.
- Aggarwal, S., Chugh, N., Jan. 2019. Signal processing techniques for motor imagery brain computer interface: A review. Array 1-2, 100003. URL: <https://doi.org/10.1016/j.array.2019.100003> DOI: 10.1016/j.array.2019.100003
- Ahn, M., Lee, M., Choi, J., Jun, S., Aug. 2014. A review of brain-computer interface games and an opinion survey from researchers, developers and users.

- Sensors 14 (8), 14601–14633.
URL: <https://doi.org/10.3390/s140814601>
DOI: 10.3390/s140814601
- Altman, D. G., 2006. Practical Statistics for Medical Research. Chapman & Hall/CRC.
- Anderson, C. W., Sijercic, Z., 1996. Classification of eeg signals from four subjects during five mental tasks. In: Solving engineering problems with neural networks: proceedings of the conference on engineering applications in neural networks (EANN'96). Turkey, pp. 407–414.
- Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., Robbins, K. A., Jun, 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics* 9.
URL: <https://doi.org/10.3389/fninf.2015.00016>
DOI: 10.3389/fninf.2015.00016
- Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., Flor, H., Mar, 1999. A spelling device for the paralysed. *Nature* 398 (6725), 297–298.
URL: <https://doi.org/10.1038/18581>
DOI: 10.1038/18581
- Bobrov, P., Frolov, A., Cantor, C., Fedulova, I., Bakhnyan, M., Zhavoronkov, A., Jun, 2011. Brain-computer interface based on generation of visual images. *PLoS ONE* 6 (6), e20674.
URL: <https://doi.org/10.1371/journal.pone.0020674>
DOI: 10.1371/journal.pone.0020674
- Craik, A., He, Y., Contreras-Vidal, J. L., Apr. 2019. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of Neural Engineering* 16 (3), 031001.
URL: <https://doi.org/10.1088/1741-2552/ab0ab5>
DOI: 10.1088/1741-2552/ab0ab5
- Esfahani, E. T., Sundararajan, V., Oct. 2012. Classification of primitive shapes using brain-computer interfaces. *Computer-Aided Design* 44 (10), 1011–1019.
URL: <https://doi.org/10.1016/j.cad.2011.04.008>
DOI: 10.1016/j.cad.2011.04.008
- Fernando, L., Nicolas-Alonso, J., Gomez-Gil, 2012. Brain computer interface, a review. *Sensors* 12 (2), 1211–1279.
- Gavali, P., Banu, J. S., 2019. Deep convolutional neural network for image classification on CUDA platform. In: *Deep Learning and Parallel Computing Environment for Bioengineering Systems*. Elsevier, pp. 99–122.
URL: <https://doi.org/10.1016/b978-0-12-816718-2.00013-0>
DOI: 10.1016/b978-0-12-816718-2.00013-0
- Gong, M., Xu, G., Li, M., Lin, F., May 2020. An idle state-detecting method based on transient visual evoked potentials for an asynchronous ERP-based BCI. *Journal of Neuroscience Methods* 337, 108670.
URL: <https://doi.org/10.1016/j.jneumeth.2020.108670>
DOI: 10.1016/j.jneumeth.2020.108670
- Han, C.-H., Müller, K.-R., Hwang, H.-J., Mar. 2020. Brain-switches for asynchronous brain-computer interfaces: A systematic review. *Electronics* 9 (3), 422.
URL: <https://doi.org/10.3390/electronics9030422>
DOI: 10.3390/electronics9030422
- Hortal, E., Planelles, D., Resquin, F., Climent, J. M., Azorín, J. M., Pons, J. L., Oct. 2015. Using a brain-machine interface to control a hybrid upper limb exoskeleton during rehabilitation of patients with neurological conditions. *Journal of NeuroEngineering and Rehabilitation* 12 (1).
URL: <https://doi.org/10.1186/s12984-015-0082-9>
DOI: 10.1186/s12984-015-0082-9
- Jiang, J., Zhou, Z., Yin, E., Yu, Y., Liu, Y., Hu, D., Nov. 2015. A novel morse code-inspired method for multiclass motor imagery brain-computer interface (BCI) design. *Computers in Biology and Medicine* 66, 11–19.
URL: <https://doi.org/10.1016/j.compbiomed.2015.08.011>
DOI: 10.1016/j.compbiomed.2015.08.011
- Jurcak, V., Tsuzuki, D., Dan, I., Feb. 2007. 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *NeuroImage* 34 (4), 1600–1611.
URL: <https://doi.org/10.1016/j.neuroimage.2006.09.024>
DOI: 10.1016/j.neuroimage.2006.09.024
- Kim, C., Sun, J., Liu, D., Wang, Q., Paek, S., Mar. 2018. An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI. *Medical & Biological Engineering & Computing* 56 (9), 1645–1658.
URL: <https://doi.org/10.1007/s11517-017-1761-4>
DOI: 10.1007/s11517-017-1761-4
- Knauff, M., Kassubek, J., Mulack, T., Greenlee, M. W., Dec. 2000. Cortical activation evoked by visual mental imagery as measured by fMRI. *NeuroReport* 11 (18), 3957–3962.
URL: <https://doi.org/10.1097/00001756-200012180-00011>
DOI: 10.1097/00001756-200012180-00011
- Kosmyna, N., Lindgren, J. T., Lécuyer, A., Sep. 2018. Attending to visual stimuli versus performing visual imagery as a control strategy for EEG-based brain-computer interfaces. *Scientific Reports* 8 (1).
URL: <https://doi.org/10.1038/s41598-018-31472-9>
DOI: 10.1038/s41598-018-31472-9
- Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B., Jan. 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering* 4 (2), R1–R13.
URL: <https://doi.org/10.1088/1741-2560/4/2/r01>
DOI: 10.1088/1741-2560/4/2/r01
- McCall, J., Dec. 2005. Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics* 184 (1), 205–222.
URL: <https://doi.org/10.1016/j.cam.2004.07.034>
DOI: 10.1016/j.cam.2004.07.034
- Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., König, P., Mar. 2017. Systems, subjects, sessions: To what extent do these factors influence EEG data? *Frontiers in Human Neuroscience* 11.
URL: <https://doi.org/10.3389/fnhum.2017.00150>
DOI: 10.3389/fnhum.2017.00150
- Pedroni, A., Bahreini, A., Langer, N., Nov. 2019. Automagic: Standardized preprocessing of big EEG data. *bioRxiv*.
URL: <https://doi.org/10.1101/460469>
DOI: 10.1101/460469
- Planelles, D., Hortal, E., Costa, Á., Úbeda, A., Iáez, E., Azorín, J., Sep. 2014. Evaluating classifiers to detect arm movement intention from EEG signals. *Sensors* 14 (10), 18172–18186.
URL: <https://doi.org/10.3390/s141018172>
DOI: 10.3390/s141018172
- Proakis, John, M., Dimitri, 1996. Digital signal processing: principles, algorithms and applications. Prentice-Hall, 910–913.
- Reza, Abiri, S., Borhani, E. W., Sellers, Y., Jiang, X., Zhao, 2018. A comprehensive review of eeg-based brain-computer interface paradigms. *Journal of Neural Engineering* 16.
- Seeck, M., Koessler, L., Bast, T., Leijten, F., Michel, C., Baumgartner, C., He, B., Beniczky, S., Oct. 2017. The standardized EEG electrode array of the IFCN. *Clinical Neurophysiology* 128 (10), 2070–2077.
URL: <https://doi.org/10.1016/j.clinph.2017.06.254>
DOI: 10.1016/j.clinph.2017.06.254
- Swets, J., 2014. Signal detection theory and ROC analysis in psychology and diagnostics : collected papers. Psychology Press, New York, New York Hove, England.
- Xie, S., Kaiser, D., Cichy, R. M., Aug. 2020. Visual imagery and perception share neural representations in the alpha frequency band. *Current Biology* 30 (15), 3062.
URL: <https://doi.org/10.1016/j.cub.2020.07.023>
DOI: 10.1016/j.cub.2020.07.023
- Zhang, W., Tan, C., Sun, F., Wu, H., Zhang, B., Dec. 2018. A review of EEG-based brain-computer interface systems design. *Brain Science Advances* 4 (2), 156–167.
URL: <https://doi.org/10.26599/bsa.2018.9050010>
DOI: 10.26599/bsa.2018.9050010