

Salamanda, Gto., a 27 de Julio del 2020

M. en I. HERIBERTO GUTIÉRREZ MARTIN
JEFE DE LA UNIDAD DE ADMINISTRACIÓN ESCOLAR
PRESENTE.-

Por medio de la presente, se otorga autorización para proceder a los trámites de impresión, empastado de tesis y titulación al alumno Martín Montiel Rodríguez del **Programa de Maestría en Ingeniería Eléctrica** y cuyo número de **NUA** es: 142368 del cual soy director. El título de la tesis es: Diseño e implementación de un sistema de identificación de personas basado en EEG y aprendizaje profundo

Hago constar que he revisado dicho trabajo y he tenido comunicación con los sinodales asignados para la revisión de la tesis, por lo que no hay impedimento alguno para fijar la fecha de examen de titulación.

ATENTAMENTE



NOMBRE Y FIRMA
DIRECTOR DE TESIS
Dr. José Ruiz Pinales



NOMBRE Y FIRMA
PRESIDENTE
Dr. José Amparo Andrade Lucio



NOMBRE Y FIRMA
VOCAL
Dr. Horacio Rostro González



UNIVERSIDAD DE GUANAJUATO

CAMPUS IRAPUATO SALAMANCA

DIVISIÓN DE INGENIERÍAS

“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE
IDENTIFICACIÓN DE PERSONAS BASADO EN EEG Y APRENDIZAJE
PROFUNDO”

TESIS PROFESIONAL PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERÍA ELÉCTRICA

PRESENTA:

ING. MARTÍN MONTIEL RODRÍGUEZ

ASESOR: DR. JOSÉ RUÍZ PINALES

Salamanca, Guanajuato. Marzo 2020

Contenido

Índice de figuras.....	6
Agradecimientos	8
Agradecimientos Institucionales.....	9
Abstract	10
Resumen.....	10
Capítulo 1 Introducción.....	12
1.1 Objetivo General.....	14
1.2 Justificación	15
1.3 Antecedentes.....	16
1.4 Metodología y desarrollo	18
1.5 Organización de la tesis	19
Capítulo 2 Electroencefalograma	20
2.1 Historia de la Electroencefalografía	20
2.2 Definición y Contenido de un Electroencefalograma.....	23
2.3 Los tipos de ondas cerebrales	26
Ondas Delta (1 a 3 Hz).....	27
Ondas Theta (3,5 a 7,5 Hz)	27

Ondas Alfa (8 a 13 Hz)	27
Ondas Beta (12 a 33 Hz)	27
Ondas Gamma (25 a 100 Hz).....	28
2.4 Uso de la diadema EMOTIV EPOC	28
Capítulo 3 Deep Learning	30
3.1 Evolución de las redes neuronales	30
3.2 Redes Neuronales	33
3.3 Tipos de redes neuronales profundas.....	34
3.3.1 Redes Perceptrón multicapa.....	34
3.3.2 Autoencoder	35
3.3.3 Redes neuronales convolucionales.....	37
3.3.4 Redes neuronales recurrentes.....	41
3.4 Dropout	48
3.5 Función de activación	49
Sigmoide	50
Tangente hiperbólica.....	50
Unidad Rectificadora Lineal (ReLU).....	51
Leaky ReLU.....	51
SoftMax.....	52
3.5 Esquemas de variación de tasa de aprendizaje	52

3.5.1 Tasa de aprendizaje cíclica.....	53
Capítulo 4 Desarrollo del sistema.....	56
4.1 Base de datos	56
4.1.1 Adquisición	56
4.1.1 Acerca de la base de datos	59
4.2 Entrenamiento de la Red neuronal.....	60
4.3 Pruebas y Resultados	62
Capítulo 5 Conclusiones.....	65
5.1 Conclusiones.....	65
5.2 Trabajos a futuro y mejoras	66
Bibliografía.....	67

Índice de figuras

Figura 1 Dispositivo Emotive EPOC	14
Figura 2 Uso de señales EEG para control de una computadora	16
Figura 3 etapas del trabajo.....	18
Figura 4. Obtención de un electroencefalograma.....	24
Figura 5. Sistema estándar internacional 10-20 para la colocación de electrodos.	25
Figura 6 Colocación de electrodos en un feto.	26
Figura 7 Red Neuronal profunda.....	33
Figura 8 Arquitectura de una red neuronal profunda del tipo Perceptrón multicapa.	35
Figura 9 Arquitectura de un Autoencoder	36
Figura 10 Autoencoder convolucional.	37
Figura 11 LeNet-5.	38
Figura 12 convolución de una imagen.....	39
Figura 13 Max-pooling de dos señales.....	40
Figura 14 Capa de neuronas de una Red Neuronal Recurrente	42
Figura 15 Tipos de unidades de memoria de redes recurrentes.....	43
Figura 16 Partes de una unidad LSTM.....	44
Figura 17 Celda de memoria de una LSTM.	45
Figura 18 Compuerta de olvido de una LSTM.....	46

Figura 19 Izquierda: obtención de la actualización. Derecha: suma para obtener el nuevo estado de memoria de la celda LSTM.	47
Figura 20 Compuerta de Salida.	48
Figura 21 Aplicación de Dropout a una red neuronal.	49
Figura 22 Curva de la función tangente hiperbólica.	51
Figura 23 función de activación Leaky Relu con $a=0.5$	52
Figura 24 Esquema de tasa de aprendizaje cíclica	54
Figura 25 Ventana de la interfaz usada para el muestreo.	57
Figura 26 Contenido de un archivo.	58
Figura 27 finalización e inicio de un nuevo muestreo.	59
Figura 28 Error de entropía cruzada de la red para la partición 1.	62
Figura 29 Resultados del entrenamiento para la partición 1.	63

Agradecimientos

Quiero agradecer a mi familia por el apoyo otorgado durante la realización de este proyecto llamado maestría por su apoyo incondicional tanto económico como emocional cuando lo necesité.

Un agradecimiento especial al doctor José Ruiz Pinales por su guía y apoyo no solo en la realización de este trabajo de tesis sino el apoyo mostrado desde el inicio de los cursos hasta la finalización de ésta.

A mis amigos que me motivaron y apoyaron en la decisión de iniciar y concluir estos estudios, especialmente a Tomás Serrano y Ana María Barrera.

“De dónde viene el ser humano todos lo sabemos, a donde quiere llegar pocos lo conocen”

Immanuel Kant.

Agradecimientos Institucionales.

Quedo agradecido con el CONACYT, por el apoyo y soporte económico otorgado por esta institución sin el cual hubiera sido muy difícil la obtención de este grado. Al interés colectivo del gobierno de formar profesionales para el crecimiento del país



Al postgrado de la Universidad Guanajuato por el apoyo mostrado con recursos materiales y humanos para la realización de estos estudios.



Abstract

In the present time is necessary an identification system which is reliable and secure, brain-computer interfaces play an important role in the control process, therefore the need to identify a user of this interface.

EEG (Electroencephalography) signals are a good option for this application but these signals are not easy to process with traditional methods of digital signal processing. One solution this problem is a machine learning technique such as deep neuronal networks. The use of LSTM (Long Short Time Memory) networks constitutes a good option for processing this kind of signals.

This work is about the design of a neural network for person identification based on EEG signals. Neural network training will be performed by using parallel computing devices such as GPUs.

Resumen

En el presente es necesario un buen sistema de identificación, seguro e invulnerable, las interfaces cerebro computadora son una parte muy importante en los procesos de control, de ahí la necesidad de identificar al usuario de esta interfaz

Las señales EEG (electroencefalográficas) son una buena opción para esta aplicación, pero estas señales no son fáciles de procesar con los métodos tradicionales de procesamiento digital de señales. La solución a este problema son las técnicas de aprendizaje de máquina tales como las redes neuronales profundas. El uso de redes LSTM (memoria de largo-corto plazo) representa una buena opción para este tipo de señales.

Este trabajo trata sobre el diseño de una red neuronal para identificar personas por medio de señales EEG, el entrenamiento de la red neuronal se hará usando dispositivos de cómputo paralelo tales como GPUs.

Capítulo 1

Introducción

Desde hace mucho tiempo las personas han enviado comandos a las máquinas para comunicarse con ellas por medio del teclado y el mouse, comúnmente conocido como WIMP (Windows Icon Menu Print and Click Devices). El término máquina comprende una amplia gama de dispositivos no refiriéndose a sistemas computacionales, por ejemplo, usamos un tipo de teclado para interactuar con un teléfono móvil, controlar una televisión, el uso de varios botones para controlar las funciones de un automóvil, en un cajero automático para hacer transacciones de dinero, se encuentran infinidad de aplicaciones en las cuales interactuamos con máquinas o dispositivos a base de este sistema[1].

Los dispositivos personales son el ejemplo más evidente de cómo las nuevas formas de interfaz humano-computadora (HCI, Human computer Interface) reducen la brecha ente los modelos mentales humanos y la tecnología. Muchos sensores de bajo costo han impulsado la llegada de teléfonos inteligentes, los gestos están siendo la modalidad de mayor interacción para activar las funciones de estos. Al mismo tiempo las tecnologías de reconocimiento de voz y el incremento de la potencia de las CPU les permite a los usuarios proporcionar entradas efectivamente cuando no pueden desarrollar gestos[2].

Muchos usuarios de video juegos han estado exigiendo a los fabricantes nuevas formas de controlar a los personajes, los desarrolladores de consolas de video juegos han propuesto controladores para liberar a los jugadores de las restricciones de usar un

control con teclado para comunicarse con las consolas. Los sensores, tales como el Microsoft Kinect representan un paso adicional hacia la implementación de interfaces completamente naturales en donde el cuerpo humano se transforme en el controlador. El dispositivo permite a los usuarios proporcionar comandos a la máquina por medio de gestos y posturas corporales tal como el hardware incluido desarrolla un procesamiento en tiempo real de los datos de base desde una cámara que obtiene un esquema del esqueleto humano conformado por un conjunto de huesos y juntas. Al reconocer la posición y orientación de los huesos le permite al hardware identificar las posturas y gestos, los cuales puede ser mapeados con comandos para controlar la máquina[3].

Se han propuesto sensores que pueden realizar el seguimiento de las manos de un usuario. Por ejemplo, el Leap Motion puede realizar un seguimiento interactivo de las manos de un usuario al identificar la posición de los dedos y del centro de la palma, y luego computar las juntas de los dedos utilizando un resolvente cinemático inverso[4]. Algunos fabricantes de automóviles están proponiendo actualmente una alternativa de seguimiento de manos como modalidad de interacción en reemplazo de las pantallas táctiles tradicionales dirigidas a la gestión de las funciones de información y entretenimiento. En forma similar algunos televisores inteligentes les permiten el control a los usuarios por medio de gestos, con lo cual se puede reemplazar el control remoto tradicional.

Los escenarios mencionados sólo podían encontrarse en las películas de ciencia ficción, hace sólo unos años atrás, pero ahora están presentes en la realidad de la HCI. Por otra parte, nuevos e intrigantes escenarios aparecen como inminentes. Tales como interfaces cerebrales que parecen adecuadas para invertir la relación entre los humanos y las máquinas. El éxito de este nuevo paradigma de interacción descansa en avances tecnológicos futuros, que apuntan a transformar los dispositivos de interfaces en objetos usables y embebidos. Las interfaces basadas en tecnologías de realidad aumentada (AR), son ejemplos claros de esta transformación. Muchas aplicaciones para turismo, entretenimiento, mantenimiento, compras y redes sociales ya están disponibles para los

dispositivos personales pero nuevos sensores usables podrían pronto cambiar nuestros hábitos[5].

1.1 Objetivo General

Diseñar un sistema de identificación de personas basado en señales Electroencefalográficas, EEG, obtenidas mediante un dispositivo de captura EMOTIV EPOC, ver Figura 1. Aplicar técnicas de procesamiento de señales, reconocimiento de patrones basado en redes neuronales convolucionales y redes neuronales LSTM (Long Short Term Memory)[6].



Figura 1 Dispositivo Emotive EPOC [25]

Parte del trabajo consiste en construir una base de datos de señales EEG de varias personas, analizar y normalizar cada una de las señales.

Con esto se plantea una manera más fácil y segura de acceder a una computadora debido a que no se usaría un sistema de claves para acceder sino a la persona misma como llave de este sistema, incluso las señales del cerebro provenientes de la misma persona no siempre se comportan de la misma manera. A pesar de los esfuerzos, la investigación sobre identificación basada en EEG aún está en su etapa inicial, y existen

varios problemas a enfrentar. El principal problema es que las señales EEG pueden variar debido a cambios fisiológicos y psicológicos de la persona tales como angustia, fatiga.

Al final obtendremos un sistema que sea capaz de reconocer a una persona en base a las señales cerebrales capturadas por medio de una diadema EMOTIV EPOC.

1.2 Justificación

Actualmente se está buscando un método eficaz e invulnerable de identificación de personas ya sea para acceder a una cuenta bancaria o a un sistema de cómputo. Se ha detectado que los sistemas lógicos (contraseñas) e incluso los de huellas dactilares son fácilmente violados y con este tipo de señales técnicamente es imposible debido a que estas señales son muy variables.

Hay mucho interés en este tipo de trabajos sobre la identificación de señales cerebrales sobre todo en video juegos donde se desea hacer interfaces de usuario en base a estas señales e identificar sentimientos sin utilizar sensores u otro dispositivo que estorbe al usuario. Lograr una plena identificación del usuario en los sistemas actuales de identificación debido a los robos de identidad. En la figura 2 observamos a una persona usando una interfaz para controlar una computadora.



Figura 2 Uso de señales EEG para control de una computadora

Recientemente se ha incrementado el interés en la investigación sobre las señales obtenidas de un electroencefalograma debido a que estas siguen un patrón único de persona a persona permitiéndonos usar este tipo de señales como un método de biometría singular para identificación de personas.

En este trabajo se afrontarán algunos retos como lo son el obtener una base de datos adecuada para el entrenamiento de la red neuronal a usarse, sobre todo que las señales obtenidas del electroencefalograma son señales muy pequeñas, incluso haciendo necesario un preprocesamiento. Otro reto es el encontrar un método para lograr una correcta identificación de las personas en este caso esto se hará con técnicas de inteligencia artificial como las redes neuronales profundas, el uso de GPU para procesamiento paralelo nos proporciona una gran herramienta la cual nos ayudará a resolver este problema haciendo que el proyecto sea realizable y exitoso.

1.3 Antecedentes

Existen diversas partes de nuestro cuerpo que nos hacen únicos respecto al resto de las personas. Desde el iris hasta la huella dactilar, el reconocimiento facial o incluso la mano. Los científicos trabajan en el "anti-spoofing" en la mayoría de las modalidades de identificación biométrica.

Por una parte, evalúan la robustez de los sistemas biométricos actuales frente a distintos tipos de ataques, por otro lado, crean algoritmos, dispositivos, técnicas colaterales y políticas de uso que eviten o detecten esos intentos de fraude.

Muchos de los intentos de fraude en la identificación biométrica se pueden ver en las películas, aunque algunos pertenecen al terreno de la ciencia ficción. Uno que resulta factible, según los investigadores, es la reproducción de huellas mediante silicona u otros plásticos, mientras que cortar el dedo o la mano a una persona para identificarse en su lugar por la huella o la geometría de la mano sólo es factible en los sistemas de gama baja.

Una buena opción es el uso de las señales cerebrales debido a que estas son particulares a cada individuo y una característica fundamental es que el sujeto a identificar debe estar vivo si no su actividad cerebral sería nula, haciendo este método de identificación muy seguro y muy difícil de corromper.

La electroencefalografía consiste en detectar las variaciones en los potenciales eléctricos entre los puntos del cerebro que son resultado de la actividad neuronal del área observada [7], hay métodos invasivos y no invasivos, debido a que los electrodos pueden implantarse sobre la corteza cerebral, en el cráneo o mantenerse en el cuero cabelludo de forma superficial. Como el potencial eléctrico depende de si hay o no actividad cerebral no siempre se obtienen las mismas señales, sino que dependen de la actividad que realice la persona a examinar [8].

1.4 Metodología y desarrollo

El trabajo se realizará en 4 etapas mostradas en la Figura 3, armar una base de datos, procesamiento de las señales EEG, entrenar la red neuronal propuesta con la base de datos y probar el identificador y documentar los resultados obtenidos.

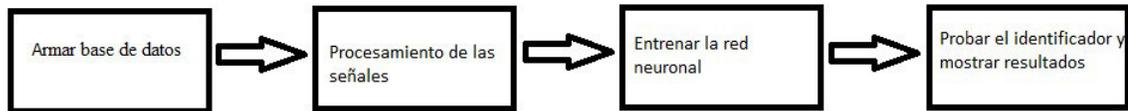


Figura 3 etapas del trabajo

Base de datos. Consiste en obtener un conjunto de señales por medio de la diadema EPOC, las cuales consisten en 14 señales provenientes de cada uno de los electrodos que contiene la diadema el cual se conecta por medio de una interfaz bluetooth a la computadora a 2.4 GHz con la cual se registrarán las señales obtenidas de cada persona individualmente y con la etiqueta correspondiente a la persona a la que pertenecen.

Procesamiento de las Señales. Se procesarán las señales obtenidas en la base de datos para que sean adecuadas para el entrenamiento de la red neuronal.

Entrenamiento de la Red Neuronal. Las arquitecturas propuestas para nuestro clasificador se entrenarán para encontrar la arquitectura idónea y con los mejores resultados para realizar la clasificación de las personas, dividiendo la base de datos en una parte de prueba y otra de validación. Se usará un entrenamiento llamado de entrenamiento profundo (Deep learning)[9] con el uso de un GPU de Nvidia.

Probar el identificador y mostrar resultados. Se realizarán pruebas con diferentes personas y se checará que tan efectivo es el clasificador para identificar a las personas,

se hará una comparativa con los resultados esperados y se comparará la efectividad del sistema con otros sistemas existentes, se documentarán todos los resultados.

1.5 Organización de la tesis

La presente tesis está estructurada y dividida en cinco capítulos los cuales se describen a continuación.

Capítulo 2: Se explicará la forma y características de las señales EEG y la forma de obtenerlas a base de la diadema EPOC de Emotiv. Así como las características de estas señales, su procesamiento y su uso en interfaces humano-máquina.

Capítulo 3: Se analizarán los tipos de redes neuronales usadas en el reconocimiento y clasificación de las señales obtenidas y que se describieron en el capítulo 2 así como el tipo de entrenamiento profundo (Deep Learning).

Capítulo 4: Se presenta y describe la arquitectura de la red neuronal usada, los resultados obtenidos del entrenamiento de esta y se presentan los resultados obtenidos en las pruebas hechas con la base de datos construida. Los resultados serán discutidos en base a las métricas propias de un calificador, comparándose su efectividad con identificadores que usan otros métodos, también se analizarán los resultados con otras bases de datos diferentes a la usada en el entrenamiento de la red.

Capítulo 5: Se discutirán los resultados obtenidos y se darán las conclusiones a las cuales se llegó con la presente investigación, así como los alcances de ésta. Se propondrán trabajos a futuro en base a estos resultados y posibles mejoras al proyecto.

Capítulo 2

Electroencefalograma

El cuerpo humano está emitiendo señales eléctricas que parecen imperceptibles, la medición de estas señales es de suma importancia, dos formas conocidas de medir estas señales son por ejemplo el electrocardiograma (mide las señales o impulsos eléctricos registrados en el corazón) y el electroencefalograma (mide y censa la actividad eléctrica del cerebro).

En este capítulo se hablará sobre el electroencefalograma y su aplicación en interfaces hombre-máquina.

2.1 Historia de la Electroencefalografía

La ciencia se encuentra con dos grandes obstáculos: la pobreza de nuestros sentidos para descubrir los hechos y la incapacidad de nuestro lenguaje para describirlos. La finalidad de un método gráfico es el poder apreciar los detalles finos y describirlos para su mayor entendimiento. Uno de estos métodos gráficos es el electroencefalograma que trata de una forma gráfica mostrarnos el funcionamiento del cerebro humano, a falta de un sensor capaz de descifrar esta actividad el hombre descubrió una forma de poder detectar los impulsos eléctricos derivados de la acción de pensar.

Teniendo en cuenta que las EEG son manifestaciones de los ritmos eléctricos del cerebro podemos remontarnos al año 1770 en el cual Luigi Galvani publicó sus observaciones

sobre la electricidad animal[10]. Galvani demostró la existencia de la electricidad intrínseca de los animales desgraciadamente estos descubrimientos a pesar de que le siguieron múltiples investigaciones tuvieron un impacto muy bajo.

Fue hasta que el científico alemán Emil du Bois Raymond, motivado por el fenómeno de la electricidad intrínseca de los nervios y músculos escribió dos libros: “Investigaciones sobre la electricidad animal” en 1848 y 1849. De los cuales el segundo libro fue muy importante ya que estableció las bases de la electroencefalografía clínica debido a que hablaba sobre los potenciales eléctricos que hay de los músculos sobre la piel [11].

Las primeras investigaciones sobre la actividad eléctrica del cerebro las hizo Richard Caton, profesor en La Escuela Real de Medicina de la Universidad de Liverpool. Caton fue influenciado por los trabajos de Edouard Hitzig y Gustav Theodor Fritsch quienes demostraron la respuesta motora en perros debido a la estimulación eléctrica en perros incluso llegando a producir convulsiones en ellos a partir de aplicar intensos estímulos eléctricos en el cráneo de los perros [11]. Caton presentó sus conclusiones en el British Medical Journal en 1875, fue el primero en observar la actividad eléctrica cerebral espontánea y continúa refiriéndose a éstas como “corrientes eléctricas en la sustancia gris”.

Aldof Beck junto con su mentor el profesor Cybulsky en la universidad de Cracovia, Polonia. Beck en su tesis doctoral describe la observación de potenciales evocados visuales. Describió también la supresión de la actividad continua de fondo de las ondas cerebrales al aplicar diferentes estímulos sensoriales. La última publicación conjunta de Beck y Cybulsky fue realizada en 1892 [11].

En 1912 los rusos Pavel Kaufman y Pradvich Neminski establecieron que los potenciales eléctricos cerebrales se pueden recoger a través del cráneo intacto, forma no invasiva. Habiendo realizado previamente estudios en perros a los que se les aplicaba una craneotomía y se les colocaban electrodos sobre la corteza cerebral, registrando la

actividad electroencefalográfica epileptiforme logrando además fotografiarla consiguiendo por primera vez una fotografía de la actividad cerebral [11].

Sin embargo, a quien se le considera el verdadero padre de la electroencefalografía es Hans Berger. El trabajo de Berger es un ejemplo de cómo partiendo de hipótesis erróneas se puede llegar a conclusiones revolucionarias, al principio en su trabajo en la Universidad de Jena en Alemania se interesó fuertemente en las bases físicas de la psicología, investigando el tema de la “la energía psíquica”, después de tratar de demostrar sin éxito la relación de la temperatura cerebral con la actividad psíquica, se convirtió en un neurofisiólogo autodidacta. Haciendo todos sus trabajos en extremo secreto trabajando cuando el laboratorio estaba cerrado ocasionando que sus compañeros de trabajo llegaran a dudar de su salud mental, otra de las grandes limitaciones con las que tuvo que lidiar fue la gran limitación de la tecnología con la que hacía sus experimentos. A pesar de las limitantes comenzó a registrar la actividad cerebral espontánea de perros y gatos. El 6 de julio de 1924 en un chico de 17 años que tenía una deformación craneal logró por primera vez detectar la actividad eléctrica del cerebro [8]. Sin embargo, sus descubrimientos no fueron recibidos muy bien por la comunidad científica de su tiempo y fue ridiculizado en numerosas ocasiones, pero gracias a su perseverancia continuó con sus trabajos publicando una quincena de artículos en los diez años siguientes. Los aportes de Berger fueron muy significativos, pero no se generalizarían hasta los años 50 a 60 del siglo XX gracias a los avances tecnológicos que pudieron comprobar sus descubrimientos. Al principio registró la actividad eléctrica en cerebros expuestos en personas que habían sufrido heridas en la primera Guerra Mundial, avanzando en sus técnicas hasta hacerlo en pacientes con el cráneo intacto en los que se incluyen setenta y tres en su hijo y cincuenta y cuatro a sí mismo[12]. Los principales avances y que constituyen el nacimiento de la electroencefalografía son los siguientes:

- Describió las ondas alfa y beta, generadas por la corteza cerebral.
- Comprobó cómo las ondas alfa desaparecían con la apertura palpebral dando paso a las ondas beta y cómo este fenómeno obedecía a otros estímulos

sensoriales. Concluyendo que la actividad alfa es fundamental de la corteza cerebral.

- Describió las ondas electroencefalográficas en niños y ancianos.
- Describió el electroencefalograma isoelectrico cuando toda actividad cerebral finalmente ha cesado.
- Comprobó los cambios observados en el electroencefalograma de pacientes con hemorragias cerebrales, tumores y contusiones.
- Registró la actividad cerebral durante una crisis de ausencia (punta – onda a tres ciclos por segundo).

En 1934, Fisher y Lowenback demostraron por primera vez los picos epileptiformes en las EEG. Gibbs, Davis y Lennox en 1935 describen ondas pico interictal y los 3 ciclos/s patrón de las crisis de ausencia clínicos, que comenzó con el campo de la electroencefalografía clínica.

En la década de 1950, William Grey Walter desarrolló un complemento de la EEG denominada topografía EEG, lo que permitió el mapeo de la actividad eléctrica a través de la superficie del cerebro. Esto disfrutó de un breve período de popularidad en la década de 1980 y parecía especialmente prometedor para la psiquiatría. Nunca fue aceptada por neurólogos y sigue siendo principalmente una herramienta de investigación.

2.2 Definición y Contenido de un Electroencefalograma

El electroencefalograma es una imagen sobre el censado de la actividad eléctrica del cerebro en situación basal, reposo y con métodos de activación, estimulación visual, auditiva, etc., observe la Figura 4.

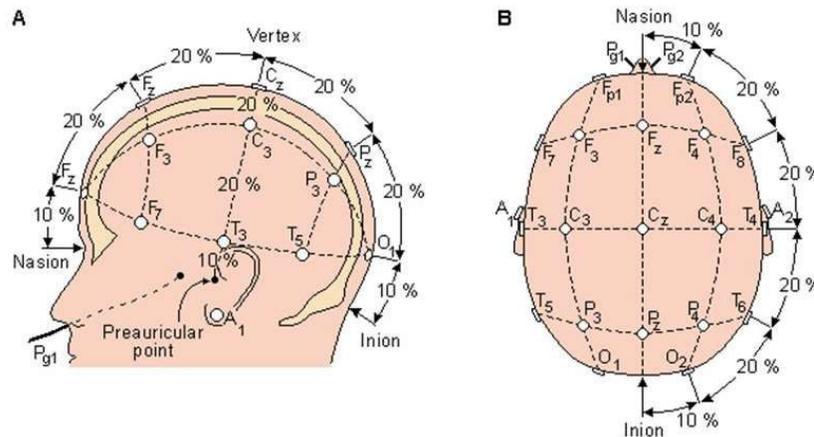


Figura 4. Obtención de un electroencefalograma.

El electroencefalograma permite medir los cambios de potencial generados en el cerebro debido a la actividad de este en tiempo real. El origen de las señales EEG es en las células piramidales de la corteza cerebral. Cada célula consiste en un diminuto dipolo eléctrico cuya polaridad depende de si el impulso a la célula es inhibitorio o excitatorio. Para poder recoger y registrar la señal resultante de la actividad cerebral se colocan electrodos que captan la diferencia de potencial entre ellos, la medición, estudio y análisis de la actividad cerebral se hace aplicando una diferencia de potencial entre los electrodos receptores de la señal. Los electrodos pueden estar sobre el cuello cabelludo y de forma intracraneal, el mejor resultado se obtiene con los electrodos puestos a cráneo abierto pero este método es invasivo y es necesario cirugía, sobre el cuero cabelludo las señales pierden potencia debido a que son del orden de microvolts pero es más amigable con el sujeto de prueba.

El EEG que se aplica comúnmente es no invasivo de bajo costo y muy fácil de realizar. Se hace mediante la colocación de los electrodos en la superficie del cuero cabelludo, son adheridos con un gel conductor o una solución salina. La forma en que son posicionados los electrodos se hace de acuerdo al sistema estándar internacional 10-20 el cual dice que se mide la circunferencia en forma lateral dejando el 10% de esta en

los extremos y después se pone el primer electrodo y a una distancia del 20% de la semicircunferencia se pone el otro y así sucesivamente, ver figura 5. Cada canal de registro mide la diferencia de potencial entre dos electrodos, uno activo y otro el de referencia. En general se usan de 16 a 24 derivaciones en cada montaje.



**A = Lóbulo de oreja, C = Centro, Pg = Nasofaríngeo, P = Parietal
F = Frontal, Fp = Polo frontal, O = occipital**

Figura 5. Sistema estándar internacional 10-20 para la colocación de electrodos.[26]

La combinación de los pares de electrodos da como resultando varios tipos de montajes, estos pueden ser de dos tipos: bipolar (transversal y longitudinal) y monopolar (de referencia). El bipolar calcula la diferencia de potencial entre dos áreas de actividad cerebral. El monopolar registra el voltaje entre un electrodo de referencia (zona sin actividad cerebral) y un electrodo de una zona activa del cerebro[13].

El uso de electrodos invasivos ayuda principalmente a acceder a señales originadas en zonas difíciles de estudiar por su difícil acceso como lo es la superficie mesial del lóbulo temporal (electrodos nasofaríngeos, electrodos fenoidales, del feromen oval temporales, anteriores, etc.). Esta técnica es muy usada cuando es necesario monitorear un feto desde el vientre de la madre, ver Figura 6.

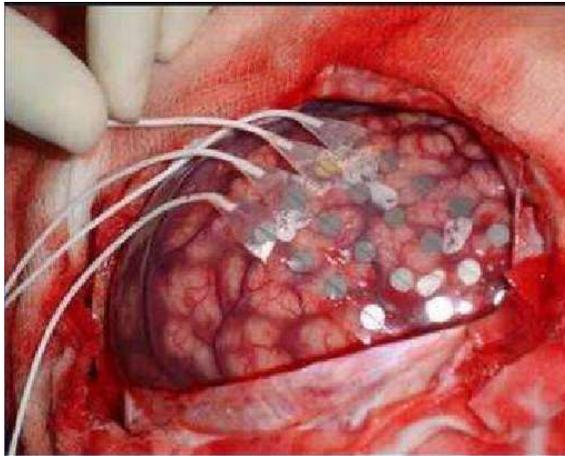


Figura 6 Colocación de electrodos en un feto.

2.3 Los tipos de ondas cerebrales

Tal y como hemos visto, la actividad de las neuronas del cerebro no es absolutamente caótica, sino que sigue una lógica muy compleja en la que puede notarse cómo diferentes neuronas disparan señales eléctricas casi al mismo tiempo de un modo continuado.

Esta frecuencia constituida por la actividad de varias neuronas forma lo que se conoce como ondas cerebrales, patrones de activación que, a diferencia de lo que ocurre con la frecuencia de activación de una sola neurona, son lo suficientemente potentes y claros como para poder ser registrados colocando sensores por fuera del cuero cabelludo (mediante la encefalografía, una de las tecnologías más utilizadas en la investigación sobre el sistema nervioso).

A su vez, las ondas cerebrales pueden ser clasificadas en diferentes tipos según su frecuencia, es decir, el tiempo que pasa entre los momentos en los que muchas neuronas disparan señales eléctricas a la vez.

Estos tipos de ondas cerebrales reciben en nombre de ondas Delta, ondas Theta, ondas Alfa, ondas Beta y ondas Gamma.

Ondas Delta (1 a 3 Hz)

Las ondas Delta son las que tienen mayor amplitud de onda, es decir, que su frecuencia es muy baja. Son características de la fase de sueño profundo, que es la fase en la que raramente soñamos. Sin embargo, que representen los patrones de activación de esta fase de sueño profundo no significa que el cerebro esté relativamente apagado. Si bien se encuentra en estado de descanso, no deja de estar activado, eso sí, ocupado con procesos que no dependen de que se esté en un estado de consciencia.

Ondas Theta (3,5 a 7,5 Hz)

Después de las ondas Delta, las Theta son las que presentan una mayor amplitud de onda. Están asociadas a los estados de calma profunda, relajación e inmersión en los recuerdos y las fantasías, y también con la etapa de sueño REM, que es aquella en la que soñamos. Por consiguiente, cuando aparecen estas ondas se estima que sí hay consciencia o que es muy probable que la haya, si bien es una consciencia desconectada de lo que ocurre a nuestro alrededor y centrada en experiencias imaginarias.

Ondas Alfa (8 a 13 Hz)

Las ondas Alfa son un tipo de onda cerebral que presenta más frecuencia que las ondas theta, aunque sigue estando relacionada con los estados de relajación. Por ejemplo, pueden aparecer durante los paseos por un parque, al tumbarse en una playa o mirando la televisión. Así pues, no son propias del estado de sueño, pero sí de calma profunda, un paso intermedio.

Ondas Beta (12 a 33 Hz)

En las ondas Beta la actividad neuronal es intensa. Están relacionadas con acciones que requieren permanecer en un cierto estado de alerta y de gestión ágil de la atención, como por ejemplo un discurso ante un público amplio, el proceso de responder a una pregunta de examen, etc.

Así pues, este tipo de ondas cerebrales está vinculado a un manejo ágil del foco de la atención, dependiendo de los objetivos, y con la preocupación por lo que ocurre en el presente, normalmente a nuestro alrededor, ya que hay que reaccionar rápidamente ante posibles imprevistos.

Ondas Gamma (25 a 100 Hz)

Estas son el tipo de ondas cerebrales con una mayor frecuencia y menor amplitud, con la ampliación del foco atencional y con la gestión de la memoria.

2.4 Uso de la diadema EMOTIV EPOC

El dispositivo Emotiv EPOC es un casco multicanal, parecido a unos auriculares, que tiene 7 pares de sensores, una pequeña batería incorporada, un transmisor, y al que le ha agregado un giroscopio. Es muy liviano, está construido en plástico resistente y flexible. Es cómodo de usar, y muy fácil de instalar. Este aparato mediante los sensores captura las ondas cerebrales del usuario. Estas ondas, luego de ser interpretadas por el software correspondiente, le dicen al sistema qué es lo que el usuario quiere hacer. En otras palabras, pensamos en “arriba” y el personaje virtual se levanta. Pensamos en movernos a la izquierda, y nuestro protagonista del juego se desplaza hacia esa dirección. Todo ocurre sólo con pensarlo, no hay que mover ni un dedo: ¡controlamos la PC con nuestra mente!

En realidad, la diadema Emotiv EPOC es un aparato de electroencefalografía (EEG), como el que han usado los neurólogos desde hace años para medir las corrientes eléctricas que nuestro cerebro produce, simplificado y concebido para ser llevado al consumidor final. A diferencia de los aparatos médicos, el EPOC es inalámbrico y se comunica con un dongle USB conectado a cualquier PC, por lo que el usuario no utiliza ningún tipo de cables y permite total movilidad mientras se usa.

Además de simplificar el aparato en sí, Emotiv simplificó lo que es la interpretación de las muchas señales que el aparato detecta. Esto significa que el usuario no va a ver nada de esto, solamente ve las acciones asociables a diferentes teclas, por lo que el uso de las señales cerebrales y su interpretación están ocultos al usuario, restando gran porcentaje de la complejidad que por naturaleza poseen los aparatos médicos.

El software que acompaña la diadema o casco tiene varias funciones. Una de ellas es detectar y calibrar todas las acciones asociadas a los pensamientos. Otra es activar y desactivar el giroscopio para poder utilizarlo como mouse, permitiéndonos mover el cursor sólo con los movimientos de nuestra cabeza. Otra utilidad es que grafica los estados de calma, ansiedad, atención o aburrimiento del usuario en un plazo de 30 minutos. También posee una carita de un robot muy simpático, que imita nuestros movimientos faciales, ya sea que pestañeamos, sonreímos, nos enojamos o guiñamos los ojos, el robotito nos imita.

A pesar de que se cuenta con los avances necesarios en el campo de la electroencefalografía y con este tipo dispositivos no se ha logrado una correcta aplicación para desarrollar un sistema antropométrico que sea eficiente y con buenos resultados y para eso se ha tomado la decisión de hacer un sistema basado en inteligencia artificial específicamente redes neuronales de aprendizaje profundo.

Capítulo 3

Deep Learning

Actualmente las redes neuronales profundas (Deep Neural Networks) han tenido un gran auge, debido en gran medida a los avances en recursos computacionales.

Para poder entender el funcionamiento de las redes neuronales hay que hacer un recorrido por la historia y evolución de la inteligencia artificial hasta llegar a los diferentes tipos de redes neuronales y sus aplicaciones en el reconocimiento de personas.

3.1 Evolución de las redes neuronales

Las redes neuronales son parte de la inteligencia artificial (IA), puede tomarse como origen el mismo de éstas.

Antes del siglo XX ya se utilizaban máquinas mecánicas las cuales contaban con una independencia del ser humano para realizar tareas en cálculos con operaciones aritméticas. El inicio de la IA se toma con la famosa prueba de Alan Turing [14] en la cual se trata de, mediante un juego de preguntas y respuestas, identificar si una máquina era capaz de pensar. La prueba consiste en separar a un interrogador, a un hombre y a una mujer mediante las preguntas el interrogador debe identificar quien es quien, pero ambos deben tratar de confundir al interrogador.

Lo interesante de la prueba es cuando se incorpora la IA en la cual una máquina toma el lugar de la persona A y el punto es cómo esta máquina trata de confundir al interrogador de que es la persona B, debido a la sencillez de la prueba se puede realizar en diferentes escenarios y actores, esto permite la certeza que una computadora responda con el conocimiento adquirido en tres ejes principales: entendimiento, razonamiento y aprendizaje. Con este test y gracias al concurso Loebner Prize se puede medir la inteligencia de los bots y sus capacidades de aprendizaje.

Los primeros trabajos acerca de la construcción de una red neuronal con fundamentos computacionales fueron hechos por Warren McCulloch un neurofisiólogo, y Wallter Pitts, un mecánico, quienes lanzaron una teoría de la forma de trabajar de las neuronas, modelaron una neurona por medio de circuitos eléctricos. Durante la década de los 50 muchos matemáticos, biólogos y psicólogos desarrollaron modelos de simulación en computadora de neuronas y redes neuronales dando lugar a la forma más general del trabajo de estos sistemas. Estos esfuerzos pronto tuvieron resultados muy favorables con los trabajos de Frank Rosenblatt quien desarrolló el perceptrón [15], la primera neurona artificial especificada con toda precisión y orientada computacionalmente. Al ser una máquina que podía aprender y demostrar un comportamiento adaptativo complejo. Su procedimiento de aprendizaje fue un avance definitivo sobre la teoría de Hebb. Rosenblatt desechó la idea del cerebro como una computadora lógica y lo consideró como un sistema asociador y clasificador, cuya función era asociar respuestas de clasificación a estímulos específicos.

En 1962 Rosenblatt presentó al perceptrón como modelo para construir Redes Neuronales Artificiales (RNAs) [16]. Los perceptrones se aplicaron rápidamente a resolver problemas como la predicción climática, la interpretación de electrocardiogramas y otros. Sin embargo, pronto se comprobó que las redes de una capa de perceptrones eran incapaces de resolver problemas como la simulación de una compuerta lógica del tipo XOR debido a que éste es un problema no lineal. Tras estas afirmaciones las redes neuronales quedaron en el olvido ocasionando el primer invierno en el estudio de éstas.

Parecía que las redes neuronales estaban muertas hasta que en 1986 surgió el entrenamiento de redes neuronales multicapa con propagación hacia atrás [17]. Este método consiste en partir de una red con pesos aleatorios y presentar patrones a la entrada de la red para obtener una respuesta que es comparada con la solución deseada obteniéndose un error. El error se utiliza para ajustar los pesos de la capa de salida y así sucesivamente siempre propagando el error a la capa anterior ajustándose los pesos en función del error obtenido hasta que se obtiene un error suficientemente bajo que satisfaga una solución al problema planteado.

Con el diseño de Hinton se da por terminado el largo invierno de las redes neuronales impulsando fuertemente el uso de las redes neuronales resolviendo grandes problemas. El principal problema de este tipo de redes es el caer en mínimos locales. Una respuesta a este tipo de problema fueron las Support Vector Machines (SVM) diseñadas por Corina Cortes y Vladimir Vapnik [18]. Básicamente una SVM representa los puntos de una muestra en el espacio, separando las clases a 2 espacios lo más amplios posibles por medio de un hiperplano o conjunto de hiperplanos en un espacio de dimensionalidad muy alta incluso infinita, este hiperplano es definido como el vector entre los 2 puntos, de las 2 clases, más cercanos al que se llama vector de soporte.

Las SVM por su fácil entrenamiento y su capacidad de resolver el problema de dimensionalidad de las bases de datos para el entrenamiento y su facilidad para encontrar un mínimo global hacen de esta una poderosa herramienta de clasificación sin embargo sólo es capaz de clasificar en 2 clases y cuando es necesario clasificar en varias clases se hace en forma de que se clasifica la clase deseada con respecto a las demás clases restante y así sucesivamente hasta que se abarcan todas las clases deseadas.

El aumento de la potencia de cálculo junto con la gran abundancia de datos ha lanzado de nuevo el uso de IA para resolver problemas para que finalmente en el 2006 Geoffrey Hinton acuñara el término “Deep Learning” (aprendizaje profundo) para explicar nuevas arquitecturas de redes neuronales las cuales cuentan con una gran cantidad de capas ocultas. Algunos trabajos a destacar son la reducción de la

dimensionalidad de la base de datos para el entrenamiento [19] y cómo entrenar una red con bastantes capas ocultas [20].

En el 2012 Hinton lideró el equipo ganador del concurso de visión por computadora Imagenet utilizando una red Neuronal de Aprendizaje profundo, venciendo por un amplio margen de diferencia dando nacimiento a la actual explosión de Machine Learning basado en aprendizaje profundo.

3.2 Redes Neuronales

Para el análisis en la presente tesis se tomarán en cuenta las redes de aprendizaje profundo, no se tiene claro y definido el tamaño para considerar una red como de aprendizaje profundo o de aprendizaje simple (Shallow), pero en general se consideran redes con más de cinco capas como de aprendizaje profundo (Deep Learning).

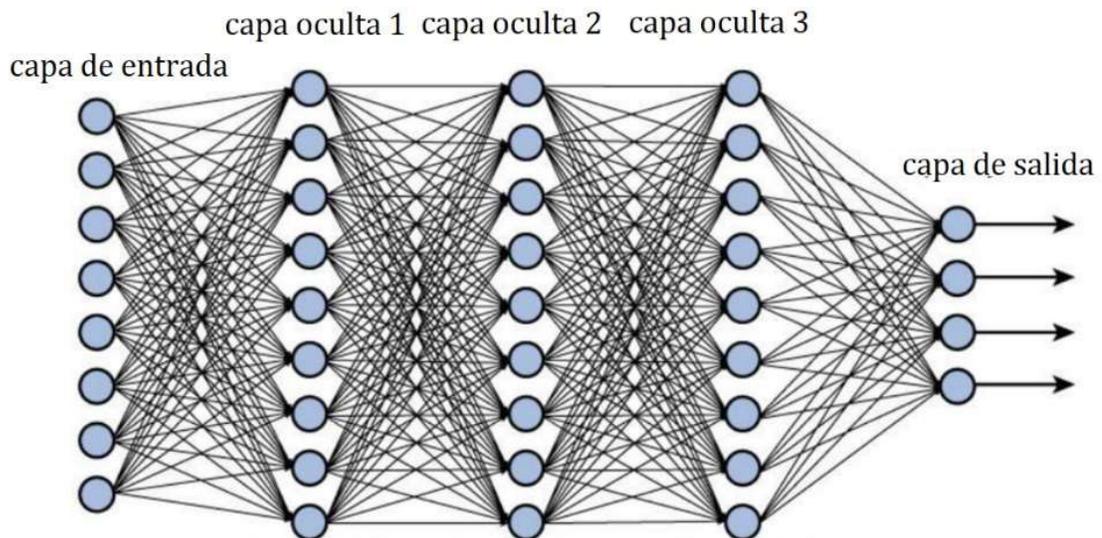


Figura 7 Red Neuronal profunda.

Con el gran avance en los equipos de cómputo y en una era en que la minería de datos se ha convertido en parte del vivir diario se ha llegado a un avance muy importante

en lo que es el uso de redes neuronales y se pasó de aquellas redes de tres capas a super redes de más de 10 capas.

Actualmente se han desarrollado varios tipos de redes que usan el aprendizaje profundo, cada tipo de red tiene sus características propias y en base a estas características se aplican para resolver problemas específicos.

3.3 Tipos de redes neuronales profundas

Las redes neuronales profundas se dividen en general en cuatro grandes grupos, Perceptrón Multicapa, Autoencoder, Redes Neuronales Convolucionales, Redes Neuronales Recurrentes.

3.3.1 Redes Perceptrón multicapa

Estas redes evolucionaron de las redes Perceptrón simple surgidas en 1986 cuyo propósito principal era el de propagar los errores cometidos en la salida hacia atrás (back propagation) ya que en un sistema de este tipo el error cometido en una capa intermedia depende de los errores de salida a las que esta capa está conectada. Al agregar capas de neuronas ocultas se consigue representar funciones lineales, debido a estas capas ocultas se convierte en una red de aprendizaje profundo.

El perceptrón multicapa está compuesto por una capa de entrada, una capa de salida y n capas ocultas intermedias. Las salidas de esta arquitectura son individuales, pero están relacionadas entre si debido a que la salida de las neuronas de una capa son las entradas de la siguiente capa y así sucesivamente Figura 8.

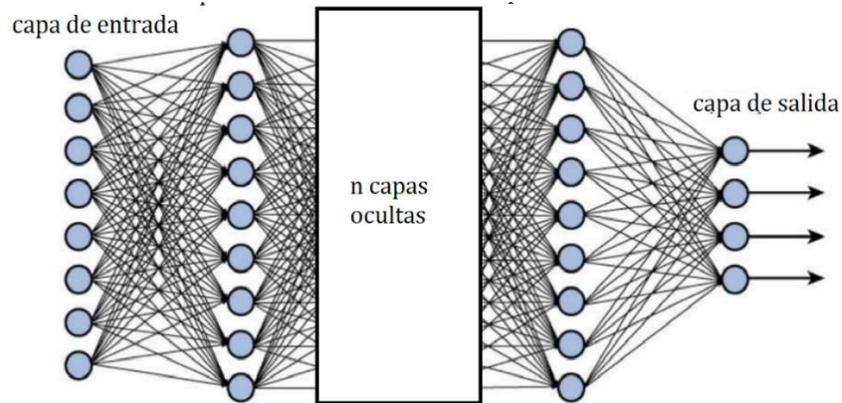


Figura 8 Arquitectura de una red neuronal profunda del tipo Perceptrón multicapa.

En la entrada se tiene un nodo por cada una de las variables de entrada y cada uno se conecta a las neuronas de la capa siguiente y así consecutivamente. Cada neurona tiene aparte de la función de activación un conjunto de pesos de entrada. En la capa de salida se tiene una neurona por cada salida deseada.

Sus principales aplicaciones son el de resolver problemas de aproximación de funciones, asociación de patrones (clasificación), segmentación de imágenes, compresión de datos entre otras.

3.3.2 Autoencoder

El Autoencoder es un tipo de red el cual va progresivamente comprimiendo los datos de entrada hasta un espacio de dimensión reducida (espacio latente) y luego va progresivamente reconstruyendo la entrada [19]. Este tipo de redes básicamente se define en dos partes el encoder y el decoder (Figura 9).

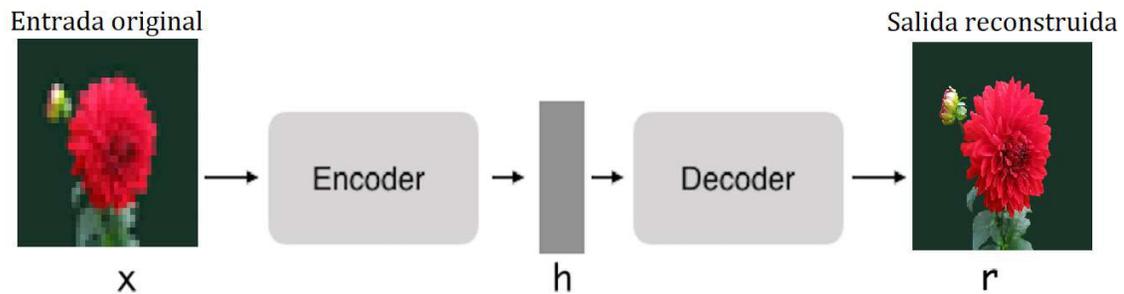


Figura 9 Arquitectura de un Autoencoder

Lo que esperamos al comprimir la entrada y luego reconstruirla en la salida es que en el espacio de \mathbf{h} podamos obtener características importantes y útiles para nosotros. Esto se puede hacer imponiendo límites de codificación y forzando que la dimensión de \mathbf{h} sea menor que la dimensión de la entrada \mathbf{x} , al hacer esto se obliga al encoder a guardar los datos más relevantes de la entrada, a este autoencoder se le llama undercomplete, esto puede suceder incluso cuando el subespacio latente tiene la misma dimensión que el espacio de entrada. Otro caso es cuando el subespacio latente es mayor que el espacio de entrada, esto significa que podemos entrenar una red y si no se hace de manera correcta podemos obtener una salida exactamente igual a la entrada sin aprender nada de los datos, a este autoencoder se le llama overcomplete.

Las principales funciones del autoencoder son la reducción del ruido y la reducción de dimensionalidad para la visualización, estas redes neuronales están capacitadas para conservar la mayor cantidad de información posible cuando se insertan en el encoder y luego en el decoder, pero de igual manera las nuevas representaciones adquieren diferentes tipos de propiedades.

Los cuatro tipos de autoencoder son:

1. Autoencoder Vanilla.
2. Autoencoder Multicapa.
3. Autoencoder Convolutivo (figura 3.4).
4. Autoencoder Regularizado.

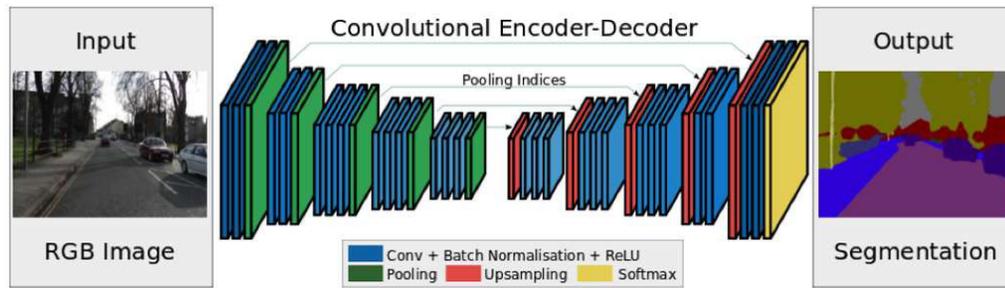


Figura 10 Autoencoder convolucional.

3.3.3 Redes neuronales convolucionales

En una Red Neuronal Convolucional (RNC) las neuronas presentan campos receptivos de una manera muy similar las neuronas de la corteza visual primaria (V1) de un cerebro biológico.

Estas redes tienen una arquitectura jerárquica de extracción de características en forma creciente empezando con características simples creciendo en complejidad y abstracción.

Los fundamentos de las redes neuronales convolucionales se basan en el Neocognitrón, introducido por Kuniyuki Fukushima en 1980. Este modelo fue más tarde mejorado por Yann LeCun en 1998 al introducir un método de aprendizaje basado en Backpropagation para poder entrenar el sistema correctamente[21].

3.3.3.1 Arquitectura de una red convolucional

La idea básica de una RNC es el lograr cierto grado de tolerancia a desplazamientos, variaciones de escala y distorsiones: esto se logra con campos receptivos locales, pesos compartidos y submuestreo espacial o temporal.

Las RNC se componen básicamente de tres tipos de capas: capas convolucionales, capas de pooling y capas full connection (totalmente conectadas). Una

red representativa es la denominada LeNet-5[21], ver Figura 11. La cual se utilizó para el reconocimiento de dígitos manuscritos.

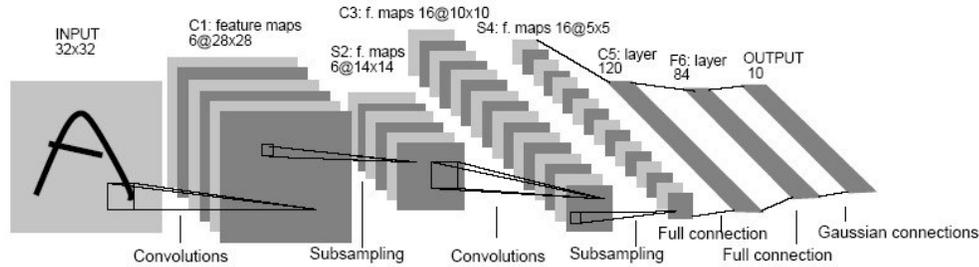


Figura 11 LeNet-5.[21]

El plano de entrada recibe imágenes de tamaño normalizado y centradas. Cada unidad en una capa recibe información de unidades localizadas en una pequeña vecindad de la capa anterior. La idea de conectar unidades a campos receptivos locales en la entrada viene del Perceptrón en la década de los 60's y fue casi simultáneamente con los descubrimientos de Hubel y Wiesel de las neuronas simple[22]. Las neuronas simples responden selectivamente a rectas/barras, bordes con una orientación y longitud dadas.

3.3.3.1.1 Capa convolucional

Para la extracción de características las neuronas de una capa convolucional realizan una operación de convolución sobre la imagen o señal de entrada, su salida se calcula como:

$$Y_j = g \left(\mathbf{b}_j + \sum_i \mathbf{K}_{ij} \otimes Y_i \right)$$

donde Y_j es la salida de la capa j , siendo ésta una matriz que se calcula por medio de la convolución de las salidas de la capa anterior Y_i con el núcleo (o kernel) de convolución K_{ij} , más un umbral b_j y luego pasando el resultado por una función de activación $g(\cdot)$ no lineal.

El operador convolucional tiene el efecto de filtrar la imagen de entrada con un núcleo previamente entrenado, ver Figura 12. Esto transforma los datos de tal manera que ciertas características (determinadas por la forma del núcleo) se vuelven más dominantes en la salida al tener éstas un nivel más alto asignado a los píxeles o valor de la señal que las representan. Los núcleos tienen habilidades específicas por ejemplo para la detección de bordes se pueden realizar con núcleos que resaltan el gradiente en una dirección particular. Sin embargo, los núcleos que son entrenados por una red neuronal convolucional generalmente son más complejos para poder estos extraer otras características más abstractas y no triviales.

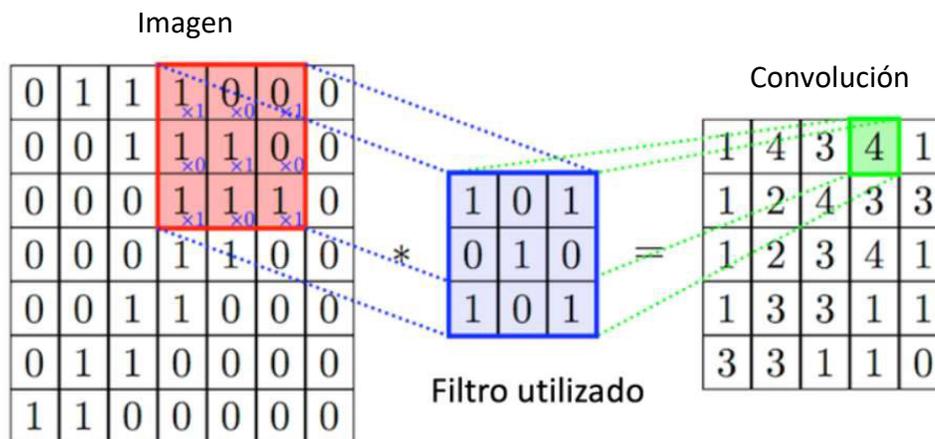


Figura 12 convolución de una imagen.[27]

3.3.3.1.2 Capa de Reducción de Escala (Pooling)

A veces también se le llama capa de submuestreo. Las redes neuronales convolucionales cuentan con cierta tolerancia a pequeñas variaciones en los datos de entrada. Por ejemplo, si dos imágenes casi idénticas (diferenciadas únicamente por cierta traslación de algunos píxeles lateralmente) se analizan con una red neuronal, el resultado debería de ser esencialmente el mismo. Esto se obtiene, en parte, debido a la reducción de escala que ocurre dentro de una red neuronal convolucional. Al reducir la resolución, las mismas características corresponderán a un mayor campo de activación en la señal de entrada.

Originalmente, las redes neuronales convolucionales utilizaban un proceso de promediado para llevar a cabo esta operación. Sin embargo, estudios recientes han demostrado que otras operaciones, como por ejemplo max-pooling, son mucho más eficaces en resumir características sobre una región. Además, existe evidencia que este tipo de operación es similar a como la corteza visual puede resumir información internamente.

La operación de max-pooling encuentra el valor máximo en vez del valor promedio del campo receptivo y pasa este valor como el resumen de características sobre esa área, reduciendo el tamaño de los datos de entrada por un factor igual al tamaño de la ventana de muestra sobre la cual opera. En la Figura 13 vemos cómo el tamaño del vector ventana o kernel es de cuatro y al aplicarse obtenemos una reducción de cuatro a uno logrando así la reducción de la señal de entrada, como se usa el max-pooling se obtiene el valor máximo de la ventana.

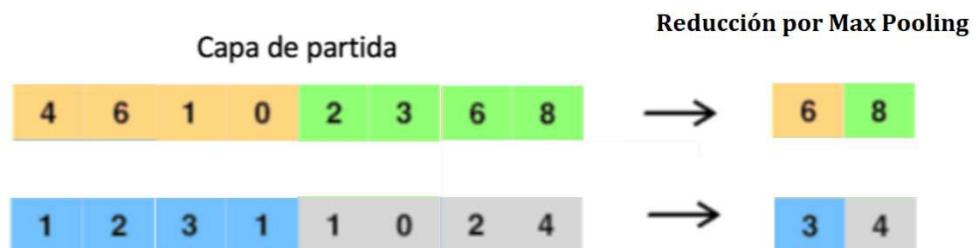


Figura 13 Max-pooling de dos señales.

3.3.3.3 Conexión total (Full Connection)

Al final de varias capas de extracción de características, los datos llegan a la última fase que es la de clasificación. Al pasar de capa en capa los datos han sido depurados hasta llegar a obtener una serie de características únicas para la señal de entrada y hay que clasificarla en una etiqueta u otra en base a las características encontradas según el objetivo del entrenamiento.

Las neuronas en esta capa funcionan igual que una capa de un Perceptrón multicapa descrito en páginas anteriores. La salida de cada neurona se calcula de la siguiente forma:

$$y_j = g \left(b_j + \sum_i w_{ij} \cdot y_i \right)$$

donde la salida y_j de la neurona j es un valor que se calcula por medio de la combinación lineal de las salidas y_i de las neuronas en la capa anterior cada una de ellas multiplicadas con un peso w_{ij} correspondiente a esa conexión. Esta cantidad es sumada a un umbral b_j y luego se pasa por una función de activación $g(\cdot)$ no-lineal.

3.3.4 Redes neuronales recurrentes

Las Redes Neuronales Recurrentes (RNR) son un tipo de redes neuronales especializadas para procesar datos secuenciales, estas pueden ser señales. Operan en el tiempo, es decir en cada paso del tiempo (épocas).

La principal diferencia de estas redes es que en las redes comunes la activación se propaga hacia la salida, no se guarda información temporal. En el caso de las RNR, éstas incluyen conexiones de retroalimentación entre las neuronas dentro de las capas, incluso asemejando una especie de memoria.

Una capa recurrente puede ser representada de tal forma que recibe dos entradas. Una es la salida de una capa anterior y la otra entrada es la salida de la capa en un instante anterior (ver Figura 14).

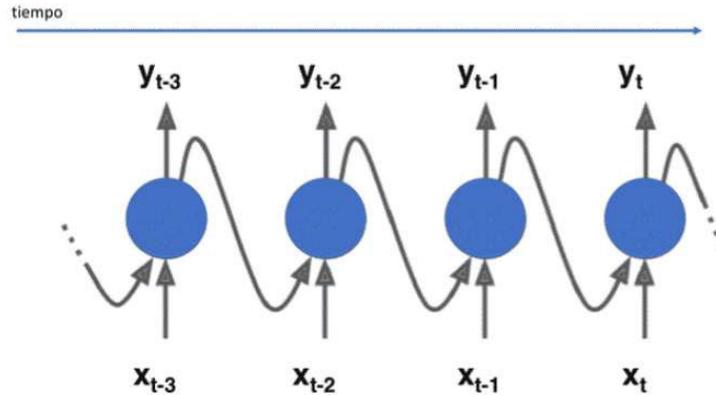


Figura 14 Capa de neuronas de una Red Neuronal Recurrente[28]

Se ha llegado a pensar que en este tipo de redes las neuronas no son como tal una neurona si no un módulo debido a que en cada módulo ocurren varios procesos.

El principal problema de este tipo de redes es el entrenamiento de ésta ya que este entrenamiento debe propagarse desde el instante inicial y en cada paso temporal de esta, ocasionando un alto costo en tiempo y memoria RAM. Esto se simplifica dividiendo la red en tantas capas como pasos temporales o de datos dispone en la secuencia temporal de entrenamiento como si fuese una red no recurrente. Cada capa tiene los mismos pesos para acelerar el proceso. En cuanto más larga se la secuencia temporal a analizar, mayor será la cantidad de capas que debe desenrollar provocando la aparición el problema de desvanecimiento de gradiente (vanishing gradient). Esto se soluciona incorporando capas de tipo LSTM (Long Short -Term Memory) [20] o GRU (Gated Recurrent Units), ver Figura 15.

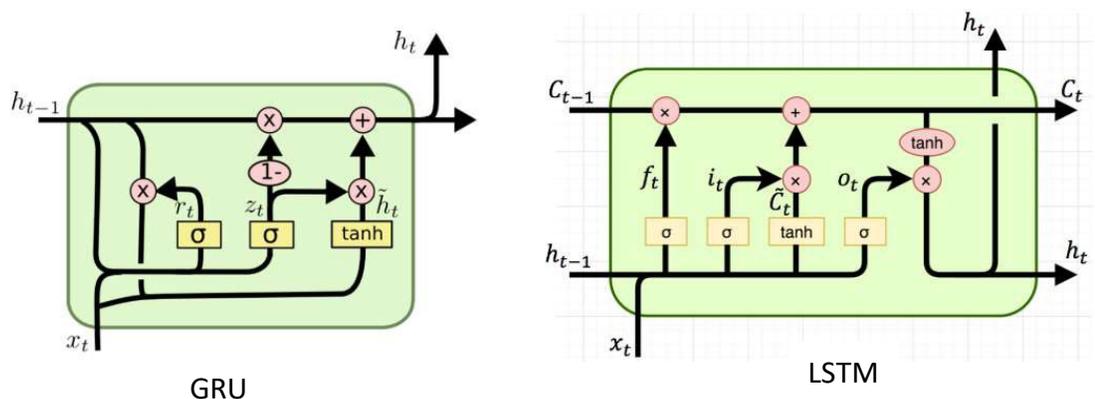


Figura 15 Tipos de unidades de memoria de redes recurrentes.[29]

La unidad de memoria GRU incluye dos compuertas, una de actualización y una de reajuste, la compuerta de actualización (Z_t) funciona como memoria debido a que es la que decide cuánto del contenido anterior se guarda o mantiene. La compuerta de reajuste incorpora el valor de entrada nuevo con el de la celda anterior (h_t) este tipo de red puede representar una red recurrente simple si el valor de la compuerta de actualización se pone en cero y la puerta de reajuste a uno.

La LSTM a diferencia de la GRU puede retener el valor durante un periodo de tiempo corto o largo en función de sus entradas, lo que le permite recordar lo que es importante e incluso usarlo en capas más profundas y no sólo usar el último valor calculado. La celda de memoria contiene tres compuertas, la compuerta de entrada controla cuando la información nueva ($h_{t-1} y X_t$) puede entrar en la memoria, la compuerta de olvido controla cuando se olvida una parte de la información para así obtener nueva información finalmente la compuerta de salida controla cuando se utiliza en el resultado de la celda la información contenida en la celda. La salida h_t es la combinación de la información almacenada y la nueva, C_t es la parte de la señal guardada en memoria puede ser una señal nueva combinación de las guardadas y la nueva información o simplemente la información conservada por celdas anteriores.

Las GRU son más fáciles y rápidas de entrenar por su simplicidad, pero las LSTM son muy usadas por su capacidad de conservar características importantes en varias capas anteriores. Actualmente este tipo de redes se usa en combinación con redes convolucionales.

3.3.4.1 Redes LSTM

Las redes LSTM (Long Short-Term Memory) [23] como se analizó anteriormente son redes neuronales recurrentes las cuales cuentan con tres compuertas de memoria la cual es su principal característica. Las tres compuertas son divididas en una de entrada, una de salida y una de olvido, de éstas hay que mencionar la celda de estado, que funciona como memoria en la cual se le puede quitar y/o añadir datos que (no) se desean guardar.

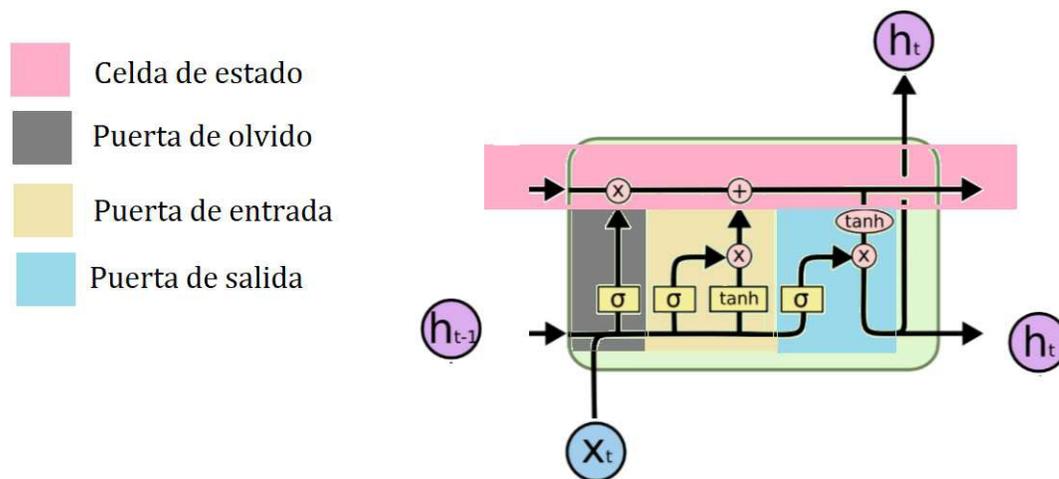


Figura 16 Partes de una unidad LSTM.

Celda de estado

Funciona como la memoria guardada en cada secuencia que entra, la entrada en la celda es la información guardada de una celda anterior C_{t-1} la cual a lo largo de toda la celda es primero multiplicada por la compuerta de olvido y después es sumada por la

compuerta de entrada al hacer este proceso es cuando se decide qué parte de la información guardada conviene mantener para la siguiente celda y cuál parte debe ser olvidada por parte de la compuerta de olvido, al sumarse con la compuerta de entrada se obtiene una salida C_t la cual es transmitida a la siguiente celda con las características guardadas de celdas anteriores y con la suma de una nueva característica obtenida en esta celda.

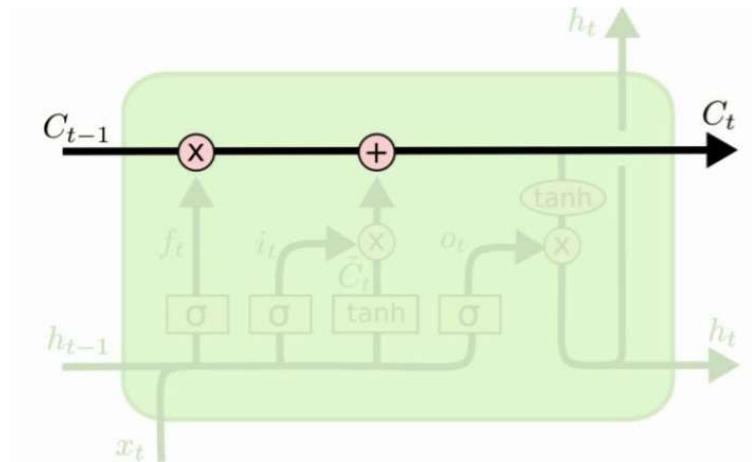


Figura 17 Celda de memoria de una LSTM.[30]

Compuerta de olvido

La compuerta del olvido tiene dos entradas una es la salida de la celda anterior h_{t-1} , también conocida como estado oculto y el valor de entrada de la nueva secuencia de la señal X_t , estas dos entradas se multiplican por el peso W_f y el bias b_f los cuales son valores a modificar durante el entrenamiento de la red, y todo esto es multiplicado por la función de activación σ . El resultado de esta compuerta nos indica la cantidad de información que es relevante y si se guarda, por cuánto tiempo se guarda o si simplemente se olvidan esas características. Al final es filtrada, multiplicada con la salida del estado anterior obteniendo así el estado de memoria de esta celda.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

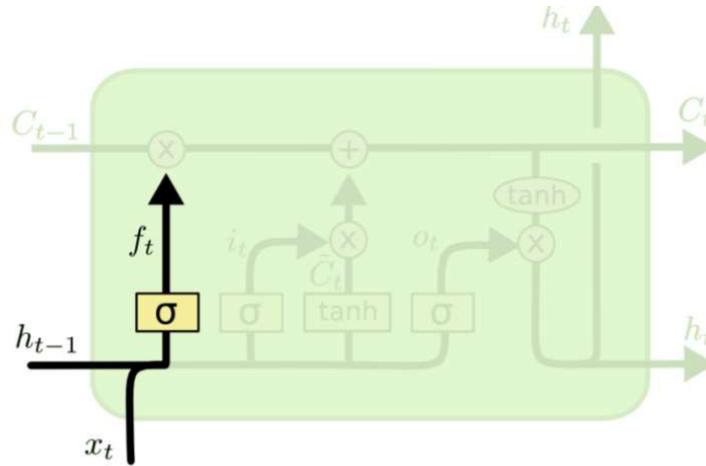


Figura 18 Compuerta de olvido de una LSTM.[30]

Puerta de entrada

Esta puerta nos sirve o su función básica es la de actualización de la barra de estado o de la memoria de nuestra celda. Básicamente se divide en dos pasos simples el primero es el comparar los datos del estado anterior con la nueva información y el segundo es mezclar esta información resultante y actualizar la barra de estado o memoria. Al principio tenemos como entrada la secuencia actual y la entrada de la salida del estado oculto de la celda anterior la cual es multiplicada por el peso W_i y el bias b_i multiplicado por la señal de activación σ obteniendo i_t y esta se mezcla o se filtra con valores candidatos a ser parte del nuevo estado de memoria por la LSTM, \hat{C}_t cuyo valor proviene del producto de la función de activación de tangente hiperbólica del producto de h_{t-1} y de X_t por los pesos W_C más el bias b_C . Una vez mezclado se le suma al estado en la barra de estado y se obtiene el nuevo estado de la memoria de la celda, C_t , ver Figura 19. Este estado de memoria es el valor que se propagará a la siguiente celda LSTM.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \hat{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t$$

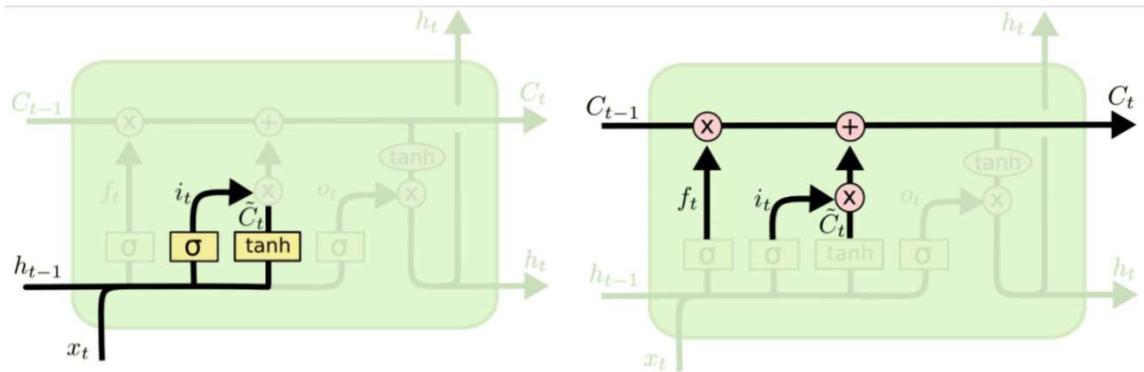


Figura 19 Izquierda: obtención de la actualización. Derecha: suma para obtener el nuevo estado de memoria de la celda LSTM.[30]

Compuerta de salida

Por último, tenemos la compuerta de salida, como su nombre lo indica esta etapa nos proporciona las salidas de la celda LSTM que son h_t que es la salida de estado oculto para la siguiente celda o para la siguiente época, recordemos que la salida de estado de memoria ya se calculó en la compuerta anterior.

El estado oculto se calcula con un filtro o multiplicando el estado de memoria de la celda que se obtuvo en la puerta de entrada con O_t , para garantizar que el valor C_t sea válido se multiplica por la tangente hiperbólica obteniendo valores entre -1 y 1. O_t proviene de la combinación de las señales de entrada h_{t-1} y el valor de la secuencia X_t multiplicada por el peso de la celda W_o y sumándole el bias b_o multiplicado por la señal de activación σ , ver Figura 20.

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o)$$

$$h_t = O_t * \tanh(C_t)$$

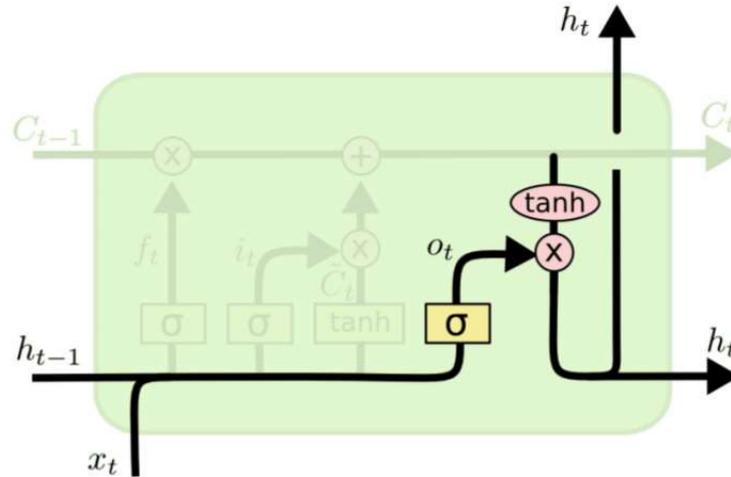


Figura 20 Compuerta de Salida.[30]

3.4 Dropout

El término dropout significa desconectar neuronas de una capa oculta o visible[21], ver Figura 21. Las neuronas cercanas suelen aprender patrones que se relacionan y estas relaciones pueden llegar a formar un patrón específico provocando linealidad en el entrenamiento de las neuronas. El dropout lo que hace es desactivar neuronas en forma aleatoria, estas neuronas desactivadas no se toman en cuenta en el método de propagación a hacia atrás ni hacia adelante obligando a las neuronas cercanas a no depender tanto de las neuronas desactivadas, evitando el sobre entrenamiento.

Al aplicar el dropout las neuronas necesitan y empiezan a trabajar de forma solitaria y así no depender tanto de las relaciones con las neuronas vecinas.

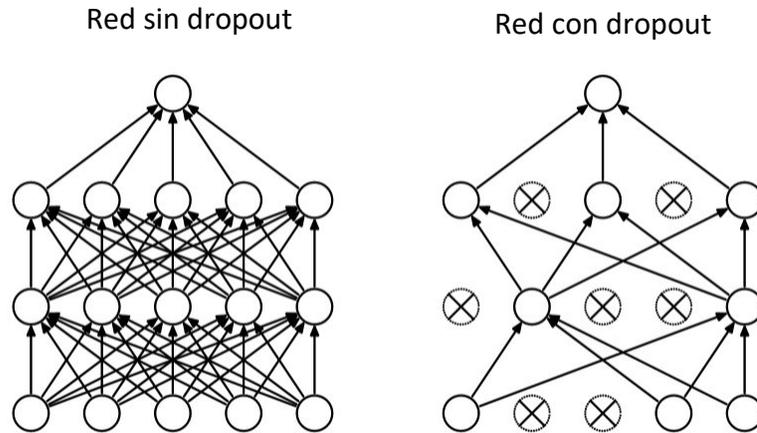


Figura 21 Aplicación de Dropout a una red neuronal.[21]

El dropout tiene un parámetro que indica la probabilidad de que se desactive una neurona este va de los valores de cero a uno, entre más cercano sea el valor a uno este significa que la mayoría de las neuronas serán desactivadas, si el valor es cercano a cero la mayoría de las neuronas permanecerán activas ocasionando que muy pocas se desactiven, generalmente el valor máximo de este parámetro es igual 0.5, lo cual nos indica que la mitad de las neuronas se desactivarán. Se puede establecer un valor de dropout en cada capa, pero generalmente se usa un valor muy alto en las capas de entrada y en las capas ocultas se usa un valor más bajo. En la capa de salida no se usa ya que es necesario tener activas todas las neuronas a la salida.

El dropout sólo se usa en la etapa de entrenamiento, en la etapa de prueba ninguna neurona se debe desactivar por lo que se escala por la probabilidad del dropout para compensar las neuronas desactivadas.

3.5 Función de activación

Básicamente son cuatro funciones de activación, hay más, pero las principales y más usadas son, sigmoide, función escalón, tangente hiperbólica, unidad de rectificación lineal (Rectified Lineal Unit, ReLU), esta se divide en tres que son ReLU, Leaky ReLU y SoftMax.

Sigmoide

Esta función se caracteriza por tener un rango de salida de $[0, 1]$, debido a esta característica se usa cuando nuestro valor de salida lo queremos mantener en esos valores. Los valores muy altos tienden a 1 y los muy bajos a 0 de forma asintótica. Está definida por la ecuación:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Tangente hiperbólica

Transforma los valores de entrada a un rango de $[-1, 1]$. Su función es muy similar a la sigmoide debido a que aquí los valores muy bajos tienden a -1 y los muy altos a +1, está centrada en cero como se muestra en la Figura 22. Es muy utilizada para decidir entre una opción y otra y tiene un buen desempeño en redes recurrentes. Es definida por la ecuación:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

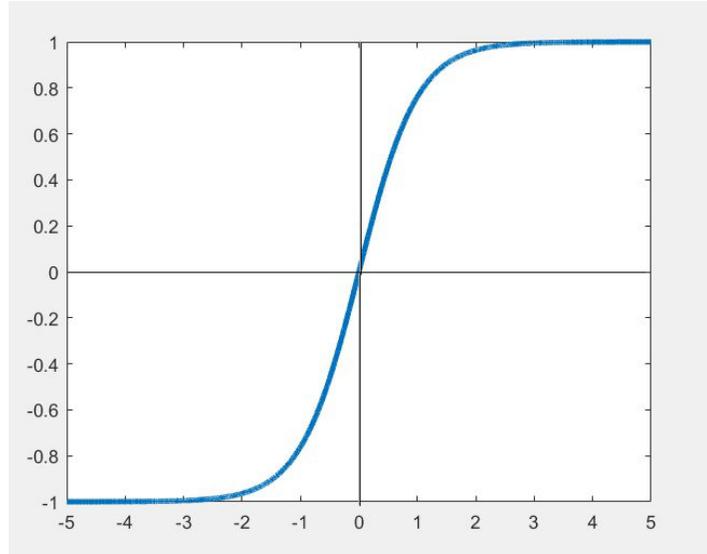


Figura 22 Curva de la función tangente hiperbólica.

Unidad Rectificadora Lineal (ReLU)

Este tipo de función de activación permite o deja pasar los valores mayores a cero y los valores menores los elimina, La principal virtud de esta función es la capacidad de propagación que tiene sobre todo en redes de aprendizaje profundo. Está definida por la ecuación:

$$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ x & \text{para } x \geq 0 \end{cases}$$

Leaky ReLU

Esta función transforma los valores negativos al multiplicarlos por un coeficiente rectificador y deja los positivos como entran, pero en esencia es el mismo funcionamiento que la ReLU, algo importante es que al igual que la ReLU no está acotada. Tiene muy buen desempeño en redes convolucionales y sobre todo en el procesamiento de imágenes. Está definida por la ecuación:

$$f(x) = \begin{cases} -ax, & \text{para } x < 0 \\ x, & \text{para } x \geq 0 \end{cases}$$

En donde el parámetro a es el parámetro con el que se multiplican los valores negativos.

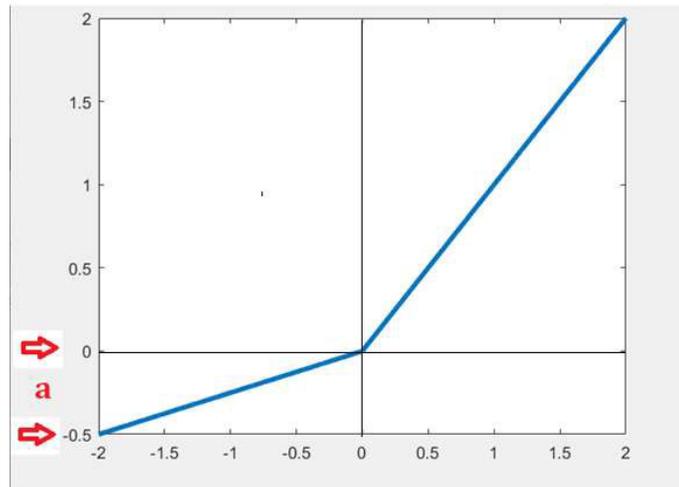


Figura 23 función de activación Leaky Relu con $a=0.5$.

SoftMax

La función SoftMax transforma las salidas en probabilidades de tal forma que la suma de todas las probabilidades de salida es 1 por lo tanto está acotada entre $[0,1]$. Es muy utilizada para normalizar en clasificación multi-clase, y tiene un buen rendimiento en las últimas capas. Su función está dada por la ecuación:

$$f(x)_j = \frac{e^{x_j}}{\sum_{k=1}^K e^{x_k}}$$

3.5 Esquemas de variación de tasa de aprendizaje

Estos esquemas se usan en conjunto con un algoritmo de aprendizaje para ir variando la tasa de aprendizaje conforme se va entrenando la red. La motivación es que para una tasa de aprendizaje pequeña se requieren muchas actualizaciones de pesos para

llegar al mínimo y para una tasa de aprendizaje demasiado grande puede hacer que no haya convergencia. El objetivo de estos esquemas es acelerar el entrenamiento y evitar mínimos locales. Hay varios tipos de esquemas, pero el más novedoso es el de tasa de aprendizaje cíclica (Cyclical Learning Rate, CLR).

3.5.1 Tasa de aprendizaje cíclica

Este tipo de esquema se basa en hacer oscilar los valores de la tasa de aprendizaje de forma cíclica[24]. Se establece un límite inferior y un límite superior para la tasa de aprendizaje, ver Figura 24. El tamaño de paso es el número de iteraciones en medio ciclo, lo recomendable para este valor es un valor entre dos y diez veces el número de iteraciones en cada época, hay que destacar que la figura nos muestra una oscilación triangular, pero esta puede tomar diferentes formas (p. e., parabólica), incluso el límite superior de oscilación puede variar en cada ciclo o bien el tamaño de ciclo puede ir aumentando.

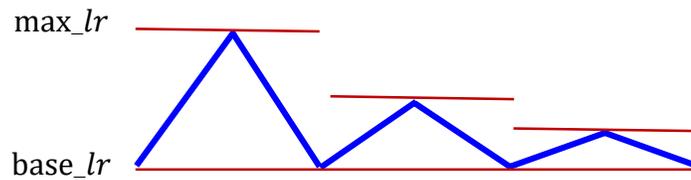
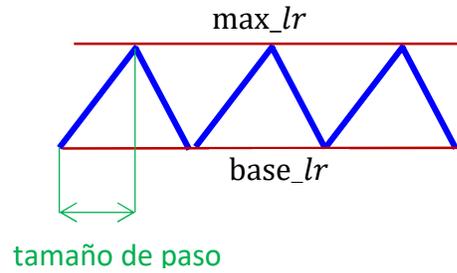


Figura 24 Esquema de tasa de aprendizaje cíclica[31][32]

Algunas ventajas del uso de este método es que el algoritmo de aprendizaje puede, bajo ciertas condiciones, escapar de mínimos locales y llegar más rápido al mínimo local/global.

En este capítulo se analizaron y estudiaron los conceptos básicos en el entrenamiento de redes neuronales de aprendizaje profundo, empezando desde la definición de éstas y su clasificación justificando la razón por las cuales se eligió este tipo de algoritmos para solucionar nuestro problema de identificación, algo muy importante a resaltar es que con el uso de estos sistemas inteligentes no es necesario hacer ningún preprocesamiento a la señal capturada.

Actualmente existen numerosas herramientas para poder entrenar una red usando GPU. Como se explicará en el siguiente capítulo se usaron Keras y Tensor Flow bajo Python 3.6.

Capítulo 4

Desarrollo del sistema

Durante este capítulo se hablará sobre el desarrollo del sistema, se divide básicamente en 3 partes, la primera describe la construcción de la base de datos de señales EEG que se utilizó en el desarrollo del sistema, la segunda parte abarca lo que es el entrenamiento, diseño de la red neuronal que se utilizó, al final se presentarán y analizarán los resultados obtenidos por el sistema.

4.1 Base de datos

La base de datos se construyó capturando las señales EEG de tres personas diferentes con características físicas diferentes y en diversos horarios y diferentes días. Como se mencionó, se utilizó una diadema de la compañía Emotiv, modelo EPOC.

4.1.1 Adquisición

Para la adquisición de muestras se desarrolló una interfaz gráfica de usuario en C# utilizando Visual Studio 2017. En la Figura 25 se observa la ventana usada para esta interfaz que consiste básicamente en dos botones y dos controles de entrada de texto.

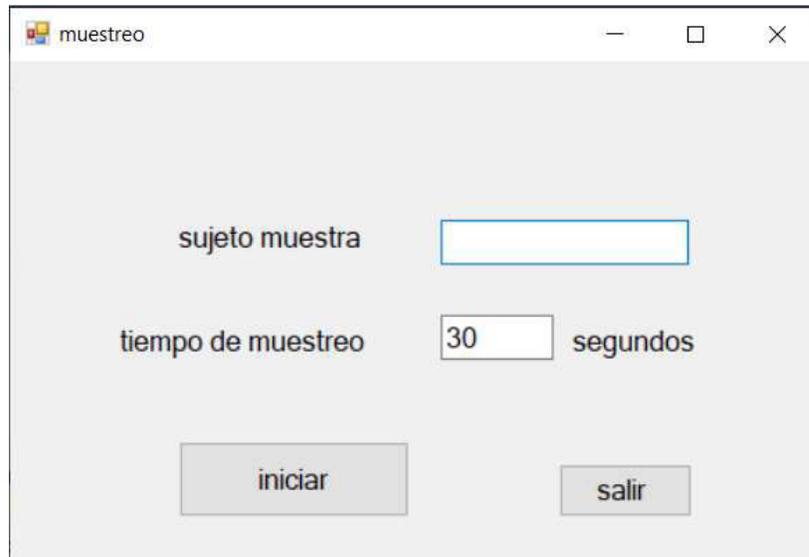


Figura 25 Ventana de la interfaz usada para el muestreo.

El primer cuadro de adquisición contiene la etiqueta del sujeto de prueba, el nombre o carácter escrito en este cuadro es el nombre con el que se etiquetara la muestra capturada, dependiendo del sujeto o persona utilizado.

La segunda etiqueta que aparece es “tiempo de muestreo” y enseguida aparece otro control de adquisición de texto, el valor que aparece ahí es el tiempo de muestreo usado para obtener la señal EEG, por default aparece 30 segundos, pero este valor puede ser modificado manualmente simplemente con poner el número de segundos con el que se quiere muestrear.

Los dos botones con los que se cuenta son “iniciar” y “salir”. El botón de iniciar una vez que se ha agregado la etiqueta y el tiempo de muestreo lo que hace es que llama al programa de adquisición de la muestra este botón bloquea la interfaz mientras realiza el muestreo para evitar conflictos mientras lo hace y también nos pone un indicador en el puntero del mouse para indicar que está en proceso. El botón salir simplemente al ser activado hace cerrar la ventana y se termina la interfaz.

En cuanto es activado el botón iniciar se guarda el nombre y el número de los segundos en variables tipo string. El número es usado para calcular el número de muestras a tomar multiplicándolo por 128 que es la frecuencia de muestreo por segundo. Con la etiqueta obtenida del sujeto de prueba lo que se hace es buscar el directorio de este sujeto, si no se encuentra crea la carpeta y empieza a crear el primer archivo con nombre 1, si encuentra una carpeta existente o directorio lo que hace es que busca el último archivo y nombra al archivo con el siguiente número. El tipo de archivo a crear es del tipo CSV (comma-separated values) de Excel.

El programa primero inicializa el archivo escribiendo un encabezado donde se indica el nombre del canal, y el contador de muestra, ver Figura 26. También se indica la fecha y hora de la muestra, estos parámetros son importantes sobre todo para tener la certeza que hay una secuencia en las muestras y que no se pierde ningún valor a la hora de grabar los datos en el archivo. Enseguida, se establece la conexión con la diadema y se programa el timer para generar eventos. Durante cada uno de los eventos generados se leen señales almacenadas en el buffer y se guardan en archivo.

COUNTER	AF3	F7	F3	FC5	T7	P7	O1	O2	P8
4	4195.89744	4206.15385	4194.87179	4198.97436	4200.51282	4201.53846	4198.46154	4207.69231	4203.07692
5	4213.84615	4182.5641	4209.74359	4215.89744	4210.25641	4186.66667	4197.4359	4178.97436	4190.25641
6	4206.66667	4163.07692	4177.4359	4246.66667	4236.41026	4142.05128	4160.51282	4176.92308	4176.41026

Figura 26 Contenido de un archivo.

Una vez que ha transcurrido el tiempo indicado en la interfaz, se desactiva el time, se hace la desconexión de la diadema, se interrumpe el muestreo y se cierra el archivo.

Al interrumpirse el muestreo se desbloquean los botones de la interfaz si se desea salir de la aplicación se puede usar el botón de salir y si se desea obtener otra muestra sólo es cuestión de ajustar los parámetros y activar de nuevo el programa de muestreo.

4.1.1 Acerca de la base de datos

Se usaron tres sujetos, y se creó una carpeta por cada sujeto. Para cada sujeto se hicieron 150 capturas de señales EEG. Cada captura duró 30 segundos, y la frecuencia del muestreo de la diadema es de 128 muestras por segundo. Por lo tanto, cada archivo de captura contiene 3840 muestras. Como se puede ver en la Figura 26, cada 128 muestras se reinicia el contador de muestra indicando que se terminó un segundo de muestreo e inicia otro segundo.

3771	125	4147.69231	5935.38462	4388.71795	3687.69231	6083.07692
3772	126	4147.17949	5895.38462	4387.69231	3688.20513	6305.64103
3773	127	4145.12821	5883.58974	4389.74359	3689.74359	6360
3774	128	4145.12821	5871.79487	4391.28205	3688.71795	6416.92308
3775	0	4146.15385	5870.76923	4388.71795	3686.15385	6423.58974
3776	1	4144.61538	5883.58974	4387.69231	3686.66667	6335.89744

Figura 27 finalización e inicio de un nuevo muestreo.

La cantidad de muestras se obtiene por la siguiente fórmula:

$$n = 128 * t$$

donde n es el número de muestras obtenidas, 128 es la frecuencia de muestreo y t es el tiempo en segundos de cada captura.

La base de datos consistió en tres carpetas, una por cada sujeto de prueba, a las cuales se les nombró como “b”, para el sujeto uno “g”, sujeto dos “m”, sujeto tres. Se seleccionaron sujetos con diferentes características: edad, género y otras particularidades. El sujeto uno es un adulto masculino de treinta y dos años, el sujeto dos es un niño masculino de diez años y el sujeto tres es un adulto masculino de treinta y siete años.

Cabe señalar que la señales no recibieron ningún preprocesamiento antes de usarse en el entrenamiento de la red. Lo único que se hizo fue normalizar la magnitud de la señal a valores en $[0, 1]$.

Al principio se pensó que eran necesarios 30 segundos de señal EEG para identificar a una persona. Sin embargo, después de varias pruebas encontramos que en realidad sólo son necesarios unos segundos. Por lo tanto, para el entrenamiento de la red sólo se usaron 3 segundos.

4.2 Entrenamiento de la Red neuronal

La red se entrenó usando Python específicamente Keras. Lo primero que se hizo fue leer cada archivo de la base de datos (llamada “base”), una vez leída se normalizaron los valores de la señal a valores en $[0, 1]$.

Se experimentó con varias arquitecturas de red neuronal conteniendo capas convolucionales y una capa LSTM. Al final se encontró una red neuronal que presentó los mejores resultados en prueba, ver Tabla 4.1.

Tipo de capa	Tamaño de salida	Número de parámetros
Conv1D	(382, 32)	1280
Dropout	(382, 32)	0
LeakyReLU	(382, 32)	0
Conv1D	(380, 32)	3104
Dropout	(380, 32)	0
LeakyReLU	(380, 32)	0
Conv1D	(378, 32)	3104
Dropout	(378, 32)	0

LeakyReLU	(378, 32)	0
Conv1D	(376, 32)	3104
Dropout	(376, 32)	0
LeakyReLU	(376, 32)	0
LSTM	(150)	109800
Dense	(3)	453

Tabla 4.1 Estructura de la red.

La primera capa es una capa convolucional la cual da una secuencia de salida de tamaño 382, tiene 32 filtros de convolución de tamaño 3×1 y su inicialización de pesos utiliza el algoritmo he_uniform. A esta capa se agregó Dropout, para regularizar los pesos. La función de activación de esta capa es Leaky ReLU la cual ya se explicó en el capítulo 3. El parámetro utilizado para el Leaky ReLU fue de 0.005.

La segunda, tercera y cuarta capas convolucionales son idénticas a la primera capa, pero con tamaños de secuencia de salida de 380, 378 y 376 respectivamente.

Después de las 4 capas convolucionales se agregó una capa de LSTM. Esta capa consiste en 150 celdas LSTM cuyo funcionamiento se explicó ampliamente en el capítulo 3.

Al final se agregó una capa dense la cual es la capa de salida donde se hace la identificación de la persona. Por lo tanto, tiene un tamaño de 3 y se usó una función de activación SoftMax.

Para entrenar la red se utilizó la función de entropía cruzada, y el optimizador Nadam (Nesterov Adaptive Momentum).

Se utilizó la técnica de tasa de aprendizaje cíclica, explicada en el capítulo 3 con el valor mínimo base de oscilación de 0.000005 y el valor o límite máximo de 0.0001 y con el tipo triangular, un tamaño de paso fue 5 veces el número de iteraciones por época. El tamaño de batch fue 64.

En total el entrenamiento se realizó durante 10000 épocas y la base de datos se particionó en cinco partes. Haciendo por cada partición un entrenamiento de la red.

4.3 Pruebas y Resultados

La figura 28 nos muestra el error de la red durante el entrenamiento de la red para la partición 1. Se observa que en el entrenamiento el error mínimo es 0.0241437 y en prueba el error más bajo es 0.137996. Se puede observar un comportamiento caótico del error con picos muy altos, esto es provocado a propósito por el esquema de tasa de aprendizaje cíclica. Después de un pico cada vez vuelve a descender a un mínimo.

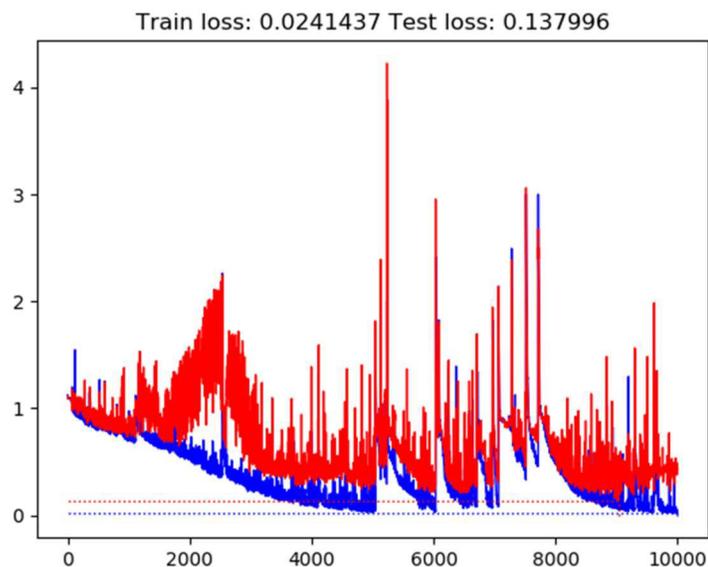


Figura 28 Error de entropía cruzada de la red para la partición 1.

En la tabla 4.2 se muestran los errores obtenidos para las 5 particiones de la base de datos.

Partición	Error en aprendizaje	Error en prueba
1	0.0241437	0.137996
2	0.0331417	0.244698
3	0.000288036	0.106507
4	0.0215909	0.107659
5	0.000635688	0.16219
Promedio	0.01596	0.173

Tabla 4.2 Error de entropía cruzada para las diferentes particiones.

En cuanto al desempeño de la red en clasificación, se obtuvo un resultado en entrenamiento de 99.41% y en prueba un 97.67% de acierto, ver Figura 29.

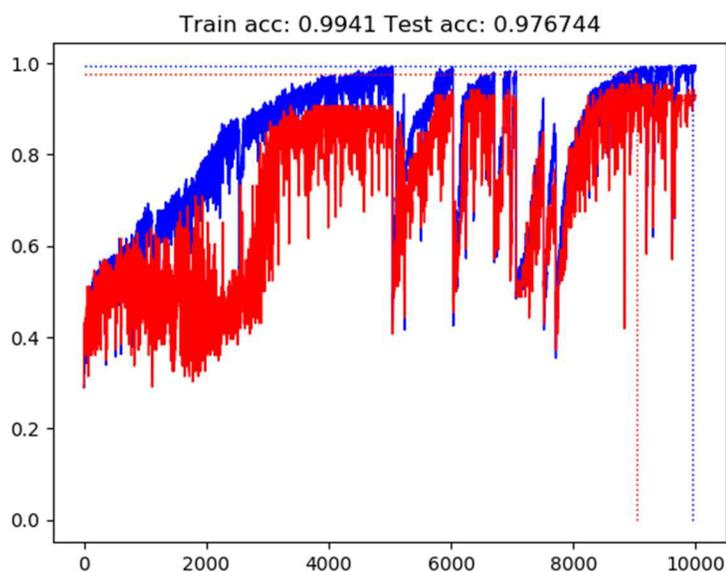


Figura 29 Resultados del entrenamiento para la partición 1.

Los resultados completos del entrenamiento se muestran en la tabla 4.3 y como se puede observar se obtiene en promedio un porcentaje de un 96.7% de aciertos en prueba, con un porcentaje muy alto en entrenamiento.

Partición	Entrenamiento	Prueba
1	99.41	97.6744
2	99.4118	95.2941
3	100.0	97.6471
4	99.4118	96.4706
5	100.0	96.4286
Promedio	99.58796	96.70296

Tabla 4.3 resultados obtenidos.

Para tener una idea más clara del comportamiento del sistema se calculó la matriz de confusión, ver tabla 4.4. Como se puede observar, sólo hay dos errores en la clase 1. Las clases 2 y 3 no presentan ningún error.

Clase correcta	Clase predicha		
	Clase 1	Clase 2	Clase 3
Clase 1	28	1	1
Clase 2	0	25	0
Clase 3	0	0	31

Tabla 4.4 Matriz de confusión para la partición 1.

En este capítulo se hizo la revisión e interpretación de los resultados obtenidos con el identificador de personas por medio de las señales EEG. En el siguiente capítulo se hablará de las conclusiones a las que se llegó, así como de las posibles mejoras y trabajo en un futuro de la investigación.

Capítulo 5

Conclusiones

5.1 Conclusiones

Se desarrolló con éxito un sistema de reconocimiento de personas por medio de señales EEG, se obtuvieron resultados con una clasificación correcta de más del 96%, poniendo como una buena opción para sistemas de reconocimiento personal o como medida antropométrica. Una de las principales ventajas con las que cuenta este sistema comparado con sistemas ya existentes como el identificador de retina, es que a diferencia de esos sistemas el sujeto debe estar si no consiente por lo menos debe estar viva evitando así que el sujeto a identificar pueda ser duplicado.

Una gran ventaja es también que a diferencia de muchos sistemas que no usan la inteligencia artificial, este sistema es capaz de realizar el trabajo de identificación sin un preprocesamiento de las señales EEG. Las señales a utilizarse son las que se obtienen directamente de la diadema. Al principio se pensó en amplificar la señal debido a que es una señal muy pequeña, pero al desarrollarse el sistema se descubrió que no es necesario.

El sistema no es invasivo ni molesto para el sujeto incluso en aplicaciones donde sea necesario estar corroborando continuamente la identificación del sujeto se puede usar la diadema sin que esto perturbe sus actividades. Dándole al sistema una gran movilidad.

A pesar de cómo se explicó en el capítulo 2 las señales del cerebro dependen mucho de la actividad de la persona, el estado de ánimo y muchos factores externos ampliamente descritos en esta tesis, la identificación de las personas se logró con éxito

ya que en realidad no se identifica con las señales recibidas sino con los patrones que son únicos para cada persona a pesar de que no siempre aparecen con las mismas señales y ni con los mismos valores este sistema es capaz de identificar la persona sin perturbaciones externas.

5.2 Trabajos a futuro y mejoras

La visión a futuro es el desarrollo y comparación del sistema en una aplicación en la vida diaria como es el reconocimiento en trámites bancarios o en aplicaciones de recaudación tributaria, activación de instrumentos o aparatos como los celulares donde el reconocimiento facial e identificación por medio de la retina tienen sus principales aplicaciones.

Una gran alternativa sería para reforzar la seguridad en casas inteligentes donde con los grandes avances en el internet de las cosas se ha hecho necesario tener un buen sistema de identificación para el acceso remoto del usuario.

Un reto para el sistema sería determinar cuál es el tiempo mínimo de muestreo para que el sistema pueda funcionar ya que sólo se usaron 3 segundos para la identificación de la persona.

Bibliografía

- [1] P. Rodríguez V, “Diseño de Interfaces Hombre - Máquina (HMI),” *Inst. Electr. y Electrónica – Univ. Austral Chile*, 2013.
- [2] “Biometric Security for Cell Phones,” *Inform. Econ. J.*, 2009.
- [3] Y. Zhu and B. Yuan, “Real-time hand gesture recognition with Kinect for playing racing video games,” in *Proceedings of the International Joint Conference on Neural Networks*, 2014.
- [4] D. Bassily, C. Georgoulas, J. Güttler, T. Linner, and T. Bock, “Intuitive and adaptive robotic arm manipulation using the leap motion controller,” in *Proceedings for the Joint Conference of ISR 2014 - 45th International Symposium on Robotics and Robotik 2014 - 8th German Conference on Robotics, ISR/ROBOTIK 2014*, 2014.
- [5] E. Orozco Coba, “Interfaces Móviles de Realidad Aumentada. Su influencia en la interacción con la información a través de la experiencia de usuario (2014),” *Cuad. del Cent. Estud. Diseño y Comun.*, 2019.
- [6] A. Graves, A. R. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013.
- [7] H.-R. Wiedemann, “Hans Berger,” *Eur. J. Pediatr.*, 1994.
- [8] C. M. Michel and M. M. Murray, “Towards the utilization of EEG as a brain imaging tool,” *NeuroImage*. 2012.

- [9] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013.
- [10] M. Piccolino, "Animal electricity and the birth of electrophysiology: The legacy of Luigi Galvani," *Brain Res. Bull.*, 1998.
- [11] L. Palacios, "Breve historia de la electroencefalografía," *Acta neurol. colomb.*, 2002.
- [12] T. J. La Vaque, "The History of EEG Hans Berger," *J. Neurother.*, 1999.
- [13] F. Ramos-Argüelles, G. Morales, S. Egozcue, R. M. Pabón, and M. T. Alonso, "Técnicas básicas de electroencefalografía: principios y aplicaciones clínicas.," *An. Sist. Sanit. Navar.*, 2009.
- [14] R. Epstein, G. Roberts, and G. Beber, *Parsing the Turing Test: Philosophical and methodological issues in the quest for the thinking computer.* 2009.
- [15] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, 1958.
- [16] B. W. White and F. Rosenblatt, "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms," *Am. J. Psychol.*, 1963.
- [17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, 1986.
- [18] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, 1995.
- [19] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science (80-.)*, 2006.

- [20] G. E. Hinton, S. Osindero, and Y. W. Teh, “A fast learning algorithm for deep belief nets,” *Neural Comput.*, 2006.
- [21] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, 2014.
- [22] D. H. Hubel and T. N. Wiesel, “Shape and arrangement of columns in cat’s striate cortex,” *J. Physiol.*, 1963.
- [23] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput.*, 1997.
- [24] L. N. Smith, “Cyclical Learning Rates for Training Neural Networks,” *2017 IEEE Winter Conf. Appl. Comput. Vis.*, pp. 464–472, Mar. 2017.
- [25] EEG headset. thinker thing. 16 de Agosto 2010. Fotografía. Flickr. <https://www.flickr.com/photos/thinkerthing/8075309856>
- [26] Miguel Ortega. EEG: interpretación para mortales. De Sapiens Medicus. <https://sapiensmedicus.org/eeg-interpretacion-para-mortales/>
- [27] Diego Calvo. Convolución. 20 Julio 2017. Imagen. Sitio web: <https://www.diegocalvo.es/red-neuronal-convolucional/convolucion/>
- [28] Jordi Torres. Sitio web: <https://torres.ai/redes-neuronales-recurrentes/>
- [29] dprogrammer. 6 de abril de 2019. <http://dprogrammer.org/rnn-lstm-gru>
- [30] Figura tomada de <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

[31] figura tomada de: <https://arxiv.org/pdf/1506.01186.pdf>

[32] figura tomada de <https://www.jeremyjordan.me/nn-learning-rate/>