

Análisis de espectrogramas de señales EEG



Presenta: Ricardo Ramos Aguilar

Director de Tesis: Dr. José Arturo Olvera López

Facultad de Ciencias de la Computación
Benemérita Universidad Autónoma de Puebla

Tesis presentada para obtener el grado de:
Doctor en Ingeniería del Lenguaje y del Conocimiento

Diciembre 2020

A toda mi familia por su inmensurable amor, apoyo y comprensión..

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo otorgado a través de la beca no. 457637 durante los estudios de doctorado. Así como a la Facultad de Ciencias de la Computación de la Benemérita Universidad Autónoma de Puebla por brindarme la oportunidad de desarrollar este trabajo de investigación dentro de su programa doctoral.

Agradezco de manera muy particular a mis asesores: Dr. José Arturo Olvera López y Dr. Ivan Olmos Pineda quienes con su amplio conocimiento, experiencia, profesionalismo, consejos y buen humor, me guiaron excelentemente hasta la conclusión del trabajo investigación.

A mi comité revisor: Dra. Susana Sánchez Urrieta, Dr. Jorge de la Calleja y Dr. Manuel Isidro Martin Ortiz, por la importancia y tiempo que le dedicaron a mi trabajo a través de sus observaciones, correcciones, enseñanzas y sugerencias.

Resumen

Una de las formas más comunes para observar el funcionamiento del cerebro es a través de la interpretación de las señales eléctricas emitidas por este órgano. Un proceso usado frecuentemente para el análisis de la actividad eléctrica es la Electroencefalografía (EEG), del cual se obtienen las señales eléctricas. Las señales EEG pueden ser usadas con varios fines, como: medicina, para la identificación de enfermedades o desordenes mentales; en aplicaciones de interfaz cerebro-computadora para dirigir dispositivos usando la actividad cerebral; en el neuromarketing, útil para la exploración de la perspectiva de las personas sobre productos; psicología y neurociencia, para el estudio de procesos cerebrales como la atención, aprendizaje o memoria, entre otros.

Para la utilización de las señales EEG en alguna aplicación, se hace un análisis visual con la finalidad de obtener información. Sin embargo, dependiendo de la actividad a interpretar el resultado puede ser el no esperado. Lo cual representa un problema que ha sido enfrentado con análisis por medio de métodos matemáticos y computacionales para extraer información de las señales. Tradicionalmente, los métodos más usados son en diferentes dominios: tiempo, frecuencia, tiempo-frecuencia y métodos no lineales.

El análisis de las señales EEG por medio de técnicas computacionales se compone de tres fases: adquisición/preprocesamiento, en donde se graban las señales y se filtra el ruido; extracción/selección de características, cuya etapa se encarga de obtener información relevante de una señal; y clasificación, en la cual se construyen modelos descriptivos con el uso de técnicas de aprendizaje computacional. En la última etapa, se identifica algún tipo de señal, sin embargo, esta depende fuertemente de la fase anterior (extracción/selección de características). Si la información obtenida es relevante o de calidad, la identificación de las señales tendrá un buen desempeño. Por lo que, una de las etapas más importante es la segunda, la cual aplica los diferentes dominios ya mencionados; uno de los que más se ha reportado en la literatura es el basado en tiempo-frecuencia, debido a que las señales son analizadas por medio de su representación espectral a lo largo del tiempo, es decir, en análisis es simultáneo en ambos dominios.

La aportación en este trabajo son tres diferentes enfoques genéricos para el análisis de espectrogramas de señales EEG. Por lo que se utilizó la transformada de Fourier

en tiempo corto (STFT) en diferentes conjuntos de datos EEG para la generación de los espectrogramas, con la finalidad de identificar epilepsia y narcolepsia. En el primer enfoque se emplean filtros Gabor para la extracción de características; en los dos últimos se aplican técnicas de extracción de características de textura y algoritmos de aprendizaje no supervisado para la agrupación de picos espectrales y la selección de los más relevantes como características. Adicionalmente, se propuso un método para la selección de un parámetro importante de la STFT (longitud de la ventana), el cual refleja su importancia en la generación de los espectrogramas.

A partir de los resultados obtenidos de diferentes experimentos, los enfoques propuestos muestran evidencia factible de ser soluciones al problema de análisis de señales EEG. El porcentaje promedio de exactitud en este trabajo es superior al 90 % para la mayoría de los experimentos; alcanzando el 100 % para uno de ellos en una clasificación binaria, el cual fue comparado con uno de los autores más relevantes en la literatura. De nuestro segundo enfoque, se encontró que la exactitud obtenida (100 %) es igual a otros métodos, con la diferencia de que nuestra propuesta usa un número menor de características con respecto a los trabajos analizados.

Abstract

One of the most common ways to observe how the brain works is through the interpretation of the electrical signals emitted by this organ. A process that consists of recording electrical activity over time is known as electroencephalography and the signals obtained as electroencephalographic signals or EEG. These signals can be used in several goals, such as: medicine, for the identification of diseases or mental disorders; in brain-computer interface applications, for controlling devices using brain activity; in neuromarketing, useful for exploring people perspective on products; psychology and neuroscience, for the study of brain processes such as attention, learning or memory, among others.

The use of EEG signals in applications involves a visual analysis, which is made in order to obtain information. However, depending on the activity to be interpreted, the result may be unsuccessful. This can be faced with analysis through mathematical and computational methods to extract information from the signals. Traditionally, the most used methods are in different domains: time, frequency, time-frequency and non-linear methods.

The EEG signals analysis through computational techniques consists of three phases: acquisition/preprocessing, where the signals are recorded and noise is filtered; feature extraction/selection, which extracts relevant information from a signal; and classification, in which descriptive models are built by computational learning techniques. In the last stage, some type of signal is identified, however, this hardly depends on its previous phase. If the information obtained is relevant or with enough quality, the identification of the signals will perform well. Therefore, one of the most important stages is the second, which applies the different domains mentioned above. In the literature, that approaches based on time-frequency is widely reported, because the signals are analyzed through their spectral representation over time, that is, analysis is simultaneous in both domains.

The contribution in this research are three different generic approaches for the analysis of EEG signals spectrograms. Therefore, the short-time Fourier transform (STFT) was used in different EEG data sets for the spectrograms generation, in order

to identify epilepsy and narcolepsy. In the first approach, Gabor filters are used for feature extraction; In the last two, techniques for the extraction of texture features and unsupervised learning algorithms are applied for grouping spectral peaks and the selection of the most relevant as features. Additionally, a method was proposed for the selection of an important parameter of the STFT (window length), which reflects its importance in the generation of spectrograms.

From the obtained results in different experiments, the proposed approaches show feasible evidence to solve the EEG signal analysis problem. The average percentage of accuracy in this work is greater than 90 % for most of the experiments; reaching 100 % for a binary classification, which was compared with a relevant author in the literature. From our second approach, it was found that the obtained accuracy (100 %) is similar to other methods, nevertheless, our proposal uses a smaller number of features than the analyzed works.

Índice general

Índice de figuras	XI
Índice de tablas	XIII
1. Introducción	1
1.1. Definición del problema	2
1.2. Objetivos	4
1.3. Pregunta de investigación	4
1.4. Hipótesis	4
1.5. Justificación	5
1.6. Aportación	6
2. Marco Teórico	9
2.1. Sistema nervioso	9
2.1.1. Descripción general	9
2.1.2. Sistema nervioso central	9
2.2. Métodos de neuroimagen	10
2.2.1. Electroencefalografía	11
2.3. Epilepsia	14
2.4. Narcolepsia	15
2.5. Transformada de Fourier de tiempo corto	16
2.6. Aprendizaje automático	17
2.7. Aprendizaje supervisado	18
2.7.1. Redes neuronales	18
2.7.2. Perceptrón multicapa	19
2.7.3. Máquinas de vectores de soporte	20
2.8. Aprendizaje no supervisado	22
2.8.1. K medias	22

2.8.2.	DBSCAN	23
2.8.3.	Técnica de sobre-muestreo minoritaria sintética	24
2.9.	Filtros Gabor	25
2.10.	Envolvente convexo	26
2.11.	Patrón binario local y patrón ternario local	26
2.12.	Métricas de evaluación	27
2.12.1.	Validación cruzada	28
2.12.2.	Coefficiente Kappa	28
2.12.3.	Prueba de significancia estadística Friedman	30
3.	Estado del arte	32
3.1.	Preprocesamiento	32
3.2.	Extracción de características y clasificación	34
3.2.1.	Análisis de textura usando la matriz de co-ocurrencia de Grises	35
3.2.2.	Extracción de características de sub-imágenes	36
3.2.3.	Otras técnicas	38
3.3.	Discusión	40
4.	Extracción de características de espectrogramas de señales EEG	45
4.1.	Señales EEG	47
4.2.	Transformada de Fourier de Tiempo Corto	47
4.3.	Enfoques de extracción de características	50
4.3.1.	Energía y Entropía de Filtros Gabor (EEFGabor)	50
4.3.2.	Características obtenidas de picos espectrales (FOSP)	52
4.3.3.	Picos Espectrales de Vecindades (PEV)	55
4.4.	Evaluación	56
5.	Experimentación y resultados	57
5.1.	Conjunto de datos	57
5.1.1.	Epilepsia - Universidad de Bonn	58
5.1.2.	Narcolepsia - BUAP	59
5.1.3.	Epilepsia - Boston	60
5.2.	Parámetros de la STFT	60
5.2.1.	Longitud de la ventana	61
5.3.	Evaluación del enfoque EEFGabor	65
5.3.1.	Clasificación de epilepsia de Bonn con EEFGabor	66
5.3.2.	Clasificación de narcolepsia con EEFGabor	68

5.3.3. Experimentos complementarios de epilepsia con EEGabor	70
5.4. Evaluación del enfoque FOSP	73
5.4.1. Clasificación de epilepsia de Bonn con enfoque FOSP	74
5.4.2. Clasificación de epilepsia de Boston con enfoque FOSP	80
5.5. Evaluación del enfoque basado en picos espectrales por medio de PEV .	80
5.5.1. Clasificación de epilepsia de Bonn con enfoque PEV	81
5.6. Discusión	82
6. Conclusiones	84
6.1. Trabajo a futuro	86
6.2. Publicaciones	86
Bibliografía	88

Índice de figuras

2.1. Cerebro	10
2.2. Sistema 10-20	12
2.3. Sistema EEG	13
2.4. Ataques epilépticos	15
2.5. Espectrograma	18
2.6. Redes neuronales artificiales	19
2.7. SVM	21
2.8. DBSCAN	24
2.9. convex	26
3.1. Etapas del análisis de señales EEG. Elaboración propia a partir de [3].	32
4.1. Metodología de análisis de señales EEG.	46
4.2. Ilustración de una grabación común de EEG.	48
4.3. Proceso de la STFT al aplicarla a una señal, manipularla espectralmente y aplicar la transformada inversa. Figura tomada de [82].	50
4.4. Enfoque basado en filtros Gabor.	51
4.5. Enfoque basado en picos máximos de energía.	52
4.6. Picos extraídos de cada ventana.	54
4.7. Extracción de características por medio de LTP.	55
4.8. Extracción de características por medio de clusters.	56
5.1. Señales de cada uno de los subconjuntos con una duración de 1 segundo, con un rango de $100 \mu\text{V}$ para las personas sanas y $1000 \mu\text{V}$ para personas con ataques. Figura tomada de [91].	59
5.2. Ventanas (a): (....)-Blackman, (_ _)-Gaussian, (-.-.)-Hamming, (-.-)- Hanning, (_ _ _)-Kaiser, y (___)-Rectangular.	63
5.3. Generación de la respuesta de Gabor a partir de los espectrogramas. . .	66
5.4. Resultados de la clasificación para todo el conjunto de estados de sueño.	69

5.5. Resultados de la clasificación para todo el conjunto de estados de sueño por rata.	70
5.6. Resultados de la clasificación para dos clases del conjunto de epilepsia para diferentes ventanas.	72
5.7. Resultados de la clasificación multiclase para todo el conjunto de epilepsia.	72
5.8. Representación de las 200 señales, 100 por cada clase usando sólo picos máximos por ventanas.	74
5.9. Picos máximos, mínimos y promedio de cada ejemplo.	74
5.10. Enfoque de extracción de picos espectrales basado en PEV.	81

Índice de tablas

2.1. Características extraídas de las representaciones en tiempo-frecuencia. . .	13
2.2. Valores del coeficiente kappa tomado de Landis y Koch [43].	29
3.1. Características de los métodos usados en el preprocesamiento a partir de [47] y [48].	35
3.2. Trabajos que utilizan representaciones en el dominio tiempo-frecuencia.	41
3.3. Características extraídas de las representaciones en tiempo-frecuencia. .	43
5.1. Características de las ventanas con respecto a las clases.	65
5.2. Resultados de la clasificación con características de energía con percep- trones multicapa y SVM.	67
5.3. Resultados de la clasificación con características de entropía con percep- trones multicapa y SVM.	67
5.4. Resultados de la clasificación con características de entropía y energía con perceptrones multicapa y SVM.	68
5.5. Resultados de la clasificación del conjunto de estados de sueño de cada rata.	70
5.6. Nomenclatura del longitud de las ventanas Hanning.	71
5.7. Resultados de la clasificación del conjunto de estados de epilepsia para diferentes tipos de ventana.	71
5.8. Resultados de la clasificación del conjunto de estados de epilepsia con los parámetros de V4 para tres y cinco clases.	72
5.9. Resultados de la clasificación comparados con Sengür et al. [89].	77
5.10. Resultados de la prueba Friedman.	78
5.11. Nomenclatura de experimentos complementarios.	79
5.12. Resultados de clasificación con nuevos subconjuntos para la identificación de epilepsia usando FOSP con picos mínimos y LTP.	79

5.13. Resultados de la clasificación del conjunto de estados de epilepsia de Boston con características de volumen y tiempo.	80
5.14. Resultados de clasificación binaria(Epilepsia) usando DBSCAN con distancia Euclidiana.	82

Capítulo 1

Introducción

El cerebro es uno de los órganos más importantes y complejos que tiene el cuerpo humano, este es el centro de control del sistema nervioso que mantiene funcionando otros órganos y desempeña distintas funciones vitales, tales como: el corazón, la respiración, temperatura interna, secreción de glándulas, entre otras. Se estima que el cerebro está formado de billones de células llamadas neuronas [1], encargadas de producir actividad eléctrica en términos de ondas cerebrales o señales electrofisiológicas, generadas durante el control que mantiene el cerebro.

Actualmente, existen diferentes métodos para medir la actividad cerebral, algunos de los más comunes son: electroencefalografía (EEG), electrocorticografía (EcoG), magnetoencefalografía (MEG), tomografía por emisión de positrones (PET), imagen por resonancia magnética funcional (fMRI). Sin embargo, MEG, PET, fMRI, son métodos que comúnmente no se usan ya que dependen de procesos metabólicos y tienen una respuesta lenta a una comunicación rápida [2–4]. De los métodos mencionados anteriormente, la electroencefalografía es un método de adquisición de impulsos eléctricos generados por la actividad cerebral, esto se realiza por medio de múltiples electrodos o sensores colocados directamente en el cuero cabelludo, los cuales con ayuda de interfaces obtienen las señales en forma digital para ser almacenada en una computadora como señal EEG. Este procedimiento es considerado como el más común para el registro de señales cerebrales, ya que tiene una alta resolución temporal permitiendo tomar mediciones de actividad cerebral a nivel de milisegundos [5, 6], es cómodo y seguro de usar al no ser invasivo [2].

Las señales EEG usualmente se describen en términos del tipo de actividad cerebral, un ejemplo son las actividades rítmicas, las cuales están divididas en distintas bandas de frecuencia que contienen significado biológico; existen seis tipos de ondas formadas a partir de las bandas: delta, theta, alpha, mu, beta y gamma [2], cada una tiene diferente

amplitud y frecuencia, asociadas directamente con distintas propiedades y actividades mentales. Las señales EEG contienen características significativas que nos ayudan a identificar diferentes enfermedades como: epilepsia, pérdida de memoria, Alzheimer, autismo, entre otras [3], estas características son proporcionadas por una o más de las diferentes bandas (delta, teta, alfa, beta y gama). Actualmente, el uso de las señales no solamente está relacionado a aspectos de medicina o biológicos, también hacen uso en las interfaces cerebro computadora (Brain Computer Interface, BCI), neuromarketing, factores humanos, biometría, entre otros [7].

El análisis de señales EEG es relevante para la interpretación de señales eléctricas cerebrales, a partir de este proceso se obtienen las características más representativas que son utilizadas para la identificación de alguna función relacionada con la actividad cerebral. Sin embargo, su análisis es complicado ya que las señales EEG son un tipo de señal de naturaleza no lineal y no estacionaria, debido a que muestran propiedades que no obedecen la superposición y escalamiento, además de tener comportamiento aleatorio y no periódico; comúnmente, las señales con estas propiedades pueden ser tratadas con diferentes métodos de procesamiento de señales (transformada de Fourier, transformada wavelet, filtros, entre otros), que se distinguen por las técnicas en las que procesan a la señal, como son: lineal, dominio de la frecuencia, dominio del tiempo, dominio tiempo-frecuencia y no lineales.

Una técnica comúnmente utilizada para el análisis de señales, es en el dominio del tiempo y la frecuencia; el análisis de señales EEG basados en este principio proporciona una herramienta muy útil, ya que es capaz de caracterizar la energía de la señal EEG, lo que nos ayuda a concentrar componentes de interés en la señal [8]. Existen distintos tipos de representaciones de tiempo-frecuencia que pueden ser aplicados para obtener un mapeo de alguna señal; una representación de las comunes para analizar señales EEG en términos del tiempo y la frecuencia está basada en la Transformada de Fourier de Tiempo Corto (STFT), la cual ha mostrado resultados que indican que es una herramienta de utilidad para el análisis de las señales EEG independientemente de algún tipo de datos en específico.

1.1. Definición del problema

Como ya se mencionó, a pesar de que existen diferentes dominios para el análisis de señales EEG, este es un problema cambiante, debido a que la manera de analizar y procesar la señal depende de la aplicación. Por otra parte, considerando que a estas señales se le suman varios tipos de artefactos y es una señal de tipo no estacionaria,

las clasificaciones de estados mentales es un problema no trivial [9]. Otra problemática que plantea un gran desafío es la dimensionalidad de los datos y esto se debe a los múltiples canales que se usan en la adquisición, ya que las señales obtenidas para analizar, comúnmente no utiliza pocos canales de adquisición [10]. A pesar de esto, no todos los datos de la señal obtenidos son útiles, ya que puede existir redundancia de los datos y esto aumenta el tiempo computacional [3]. La extracción y selección de características pueden reducir la dimensionalidad de los datos, además de mostrar patrones relevantes de la señal EEG asociados a la actividad cerebral, que se reflejarán por medio de la exactitud en la clasificación [11, 12].

El problema de dimensión implica trabajar con cantidades de datos que incrementan en cantidad, ya que por una parte se analiza la información de la actividad cerebral que es constante en el tiempo; por otra parte, el número de canales que se usa es de 8-64, a frecuencias de muestreo de más de 100Hz, resultando datos de grandes dimensiones [13]. Este volumen de datos es preprocesado para la filtración de ruido (eléctrico) o señales adheridas como artefactos, seguido de la extracción de características con la finalidad de seleccionar patrones relevantes para la construcción de modelos de clasificación. La problemática formada de varias etapas, relacionadas una con la siguiente: primero, el preprocesamiento se emplea; después, la extracción y selección de características se aplican; finalizando con la clasificación.

El método que predomina en la literatura para el análisis de señales EEG, es en términos de la frecuencia; lo cual genera más datos, lo que implica un problema como se mencionó anteriormente. Otras técnicas alternativas que son representaciones en tiempo-frecuencia son aplicadas para su análisis, las cuales son útiles. La revisión de trabajos [14–16] sobre análisis de señales EEG en términos del tiempo-frecuencia haciendo uso de STFT muestran exactitudes competitivas en comparación con otras representaciones. Sin embargo, no es la mejor en todos los casos, diferentes resultados muestran que en el mapeo o imagen obtenida representa una mejor concentración de energía que en otras representaciones, además de que la resolución no es la mejor en comparación con otras técnicas. A partir de los resultados mostrados de la literatura con respecto a representaciones de tiempo-frecuencia, se nota que las metodologías implementan distintas técnicas para cada fase del análisis, los métodos utilizados difieren unos de otros a pesar de que todos analizan representaciones. Por lo tanto, las representaciones o transformaciones utilizadas en la literatura no muestran que alguna predomine; los trabajos que aplican la STFT muestran ventajas para el análisis de las señales EEG que pueden ser útiles para una propuesta diferente.

1.2. Objetivos

El planteamiento de la solución en este trabajo de tesis respecto a la problemática mencionada se basa en el siguiente objetivo general:

- Proponer una metodología de extracción de características de espectrogramas de EEG, con la finalidad de utilizarlas para el entrenamiento de clasificadores supervisados y no supervisados, cuyo desempeño de clasificación sea al menos cercano estadísticamente a la exactitud de trabajos relacionados.

Los objetivos específicos son:

- Investigar e identificar métodos existentes en la literatura respecto a señales EEG para la extracción de características de espectrogramas y clasificación.
- Seleccionar, proponer y/o adaptar un método de extracción de características sobre espectrogramas de señales EEG.
- Evaluar las características extraídas en términos de la exactitud con al menos tres clasificadores.
- Seleccionar al menos tres corpus de señales EEG para aplicarlos a los enfoques propuestos.

1.3. Pregunta de investigación

Considerando la problemática mostrada en la sección anterior, surge la siguiente pregunta de investigación:

¿A partir de las características obtenidas de espectrogramas de señales EEG, es posible extraer patrones relevantes en términos de técnicas de aprendizaje computacional?

1.4. Hipótesis

El análisis de las señales EEG es aplicado a grandes volúmenes de datos, en donde se encuentran características relevantes relacionadas directamente con la actividad cerebral, independientemente de la acción que se realice. Estas características pueden ser extraídas usando métodos de procesamiento digital de imágenes y extracción de características en los espectrogramas de señales EEG. A partir de esto y de la pregunta de investigación formulada, las hipótesis de investigación son las siguientes:

- A partir de la información descriptiva de las señales EEG en el dominio de la frecuencia, particularmente con la representación de espectrogramas, es posible extraer características relevantes que denoten modelos descriptivos respecto a un conjunto de señales EEG.
- Mediante estos modelos puede llevarse a cabo el análisis de señales EEG por medio de técnicas de aprendizaje computacional.

1.5. Justificación

Actualmente, el análisis de las señales EEG resulta importante, ya que de este es posible obtener una perspectiva del funcionamiento del cerebro, por medio de la información que resulte útil contenida en la señal obtenida de un proceso previo de adquisición. El estudio de la señal EEG es un proceso que muestra características relevantes o representativas de la señal, asociadas con la actividad cerebral; estas características se obtienen al aplicar técnicas computacionales a la representación digital de la actividad eléctrica.

En la literatura, actualmente, existen diferentes métodos para el análisis de señales EEG, predominando el análisis en el dominio de la frecuencia, seguido del dominio del tiempo. En el dominio del tiempo se muestra dificultad para obtener información acerca de la distribución de energía sobre diferentes frecuencias y las diferentes frecuencias sobre el tiempo. Por otra parte, en el dominio de la frecuencia, se puede procesar y analizar información de contenido en frecuencias que puede ser extraído de la señal, aquí la desventaja consiste en reconocer los componentes de frecuencia ocurridos en el tiempo [8]. Para describir el contenido espectral de la señal ocurrido en el tiempo, es necesario usar una técnica en términos del tiempo y la frecuencia, que proporciona un mapeo el cual permite analizar la señal EEG por medio de distintos métodos de procesamiento digital de imágenes y usar diferentes métodos de extracción de características a los aplicados en dominios de frecuencia o temporales.

Comúnmente, diferentes aplicaciones de distintas áreas como comunicaciones, topografía sísmica, medicina, entre otras; muestran señales de naturaleza no estacionaria y variantes en el tiempo. Las señales EEG muestran características con esta naturaleza, además de tener una alta dimensión (miles-millones de muestras) de datos y baja relación señal a ruido (SNR) en un rango de 1-10 dB, lo que dificulta su análisis. Distintos métodos para el análisis han sido implementados; sin embargo, técnicas en el dominio tiempo-frecuencia se han aplicado a señales con las características mencionadas obteniendo resultados aceptables [8] con respecto a la exactitud de la clasificación; estos

métodos hacen uso de distintas representaciones de tiempo-frecuencia para su análisis, mostrando como diferencia principal la resolución de la imagen o mapeo resultante [17], lo que se ve reflejado en la clasificación.

Las representaciones de tiempo-frecuencia son transformaciones variantes en el tiempo, que describen cómo el contenido espectral de una señal está cambiando con el tiempo. La mayoría de las representaciones transforman una señal de una dimensión a una función de dos dimensiones (tiempo-frecuencia); tales representaciones son útiles para el análisis, modificación, síntesis y detección de variedad de señales no estacionarias con contenido espectral variante en el tiempo [18] como las señales EEG. Una de las representaciones más comunes para el análisis de señales EEG es la Transformada de Fourier de Tiempo Corto (STFT), esta proporciona una implementación computacional rápida y simple, debido a las características de la función matemática que la describen.

Así mismo en la literatura se han reportado resultados aceptables (exactitud mayor al 90 %) en términos de clasificación, derivados de una extracción de características de espectrogramas obtenidos de aplicar la STFT a la señal EEG. Sin embargo, a pesar de los resultados obtenidos mostrados en la literatura con esta técnica, aún se tiene un alcance limitado considerando las características de la señal anteriormente mencionadas, lo que provee diferentes retos de investigación como son: el trabajo computacional generado, la exactitud de la clasificación obtenida, reducción de grandes dimensiones de datos, filtrado de señales no deseadas, entre otras [19, 17]. Por lo tanto, es necesario proponer una metodología diferente para el análisis, la cual permita descartar señales no deseadas de las señales crudas EEG para tener una señal más limpia independientemente del enfoque, posteriormente aplicar STFT a la señal (de dimensión alta), para extraer características del espectrograma obtenido, las cuales deben mostrar patrones relevantes de la señal, con la finalidad de analizar automáticamente comportamientos asociados con la actividad cerebral por medio de clasificación supervisada o no supervisada. Con esta metodología propuesta, se espera mostrar resultados competentes con los reportados en trabajos relacionados; un análisis con resultados competitivos, puede ser útil en diferentes aplicaciones como: área médica, interfaces cerebro computadora (BCI), biometría, entre otros.

1.6. Aportación

Este trabajo se enfoca en el análisis de espectrogramas obtenidos de señales EEG, usando datos de distintos repositorios con registros de mediciones, los cuales fueron obtenidos a través de neurosensores en términos de tiempo-frecuencia. Algunos de los

datos recopilados ya fueron objeto de estudio en la literatura, por lo que los resultados obtenidos se comparan directamente con los de ellos. El desarrollo de los enfoques, analiza las fases de distintas metodologías de análisis, con la finalidad de diseñar y/o proponer un enfoque diferente.

Durante el desarrollo de este trabajo, se analizaron señales EEG a través de las distintas fases ya mencionadas; obteniendo diferentes particularidades importantes de las etapas, además de la exactitud en términos de la clasificación. A continuación se enlistan los aportes realizados:

- Inicialmente, se hace un análisis de las diferentes técnicas aplicadas a señales EEG basadas en espectrogramas, mostrando una comparativa breve de lo que proponen diferentes autores.
- Posteriormente, se propone un enfoque de extracción de características para la identificación de epilepsia, aplicando filtros Gabor a espectrogramas y obteniendo características de la respuesta.
- Se muestra un análisis de los parámetros de la STFT basados en la energía de los espectros de las ventanas más comunes. De los diferentes valores utilizados se observaron los resultados (espectrogramas) y en términos de exactitud de un experimento de clasificación, mostrando la relevancia de los parámetros de la STFT.
- A partir del punto anterior, se propone un método para encontrar la longitud de las ventanas dedicadas a las señales durante la aplicación de la STFT, basado en frecuencias conocidas de las señales analizadas.
- Con la justificación de algunos parámetros de la STFT, se desarrollan tres diferentes enfoques utilizando la información que proporcionan los espectrogramas: tiempo, frecuencia y energía.
- Finalmente, experimentos complementarios muestran el desarrollo de otro enfoque basado en picos espectrales.
- Los enfoques desarrollados, son alternativas para la identificación de epilepsia debido a que los experimentos en su mayoría se desempeñaron con dicha finalidad.

La organización del presente documento es la siguiente. En el capítulo 2, se muestran los fundamentos del análisis de EEG, así como las características de los conjuntos utilizados; se presenta una breve descripción de STFT y aprendizaje automático; el

capítulo 3, describe los trabajos relacionados; el capítulo 4, el desarrollo del modelo propuesto se describe, explicando las fases utilizadas en la metodología: preprocesamiento, extracción de características y clasificación; en el capítulo 5 se presentan los resultados experimentales obtenidos con la propuesta; finalmente, en el capítulo 6, se exponen las conclusiones y el trabajo a futuro.

Capítulo 2

Marco Teórico

En este capítulo se presentan los conceptos fundamentales que son utilizados en el documento y para el desarrollo de la propuesta metodológica, tales como: el sistema nervioso, adquisición de señales, métodos de neuroimagen, epilepsia, narcolepsia, STFT, aprendizaje automático, entre otros.

2.1. Sistema nervioso

2.1.1. Descripción general

Una neurona está compuesta de un cuerpo celular y un axón, que termina en un área sináptica. La información que viaja a través de ella es una señal eléctrica que corresponde a una despolarización de la membrana axonal. De esta forma, el axón transmite el potencial de acción hasta la sinapsis, por medio de un área de comunicación entre neuronas [2]. Las neuronas están organizadas en rutas, extensiones o redes cuyas conexiones determinan sus roles. Diferentes sistemas (motor, somatosensorial, sensorial) pueden tener vías ascendentes o descendentes, que van desde el receptor periférico al área del cerebro involucrada en la interpretación de la señal, o que va desde la corteza todo el camino hasta el efector (por ejemplo, el músculo) [20].

2.1.2. Sistema nervioso central

El sistema nervioso central incluye el encéfalo, que se encuentra en el cráneo y la médula espinal; este puede dividirse en dos partes principales: la corteza cerebral y las regiones subcorticales. Las regiones subcorticales son filogenéticamente más antiguas e incluyen áreas asociadas con el control de funciones vitales básicas como: respiración,

ritmo cardíaco y regulación de temperatura, respuestas emocionales, reflejos, así como aprendizaje y memoria [21]. La superficie de cada hemisferio tiene un aspecto plegado, que permite individualizar los lóbulos (figura 2.1): frontal, parietal, occipital y temporal, en la superficie y el lóbulo insular en el interior. Estos lóbulos están separados por surcos: central (también conocido como la fisura de Rolando), lateral, parieto-occipital y temporal-occipital [20].

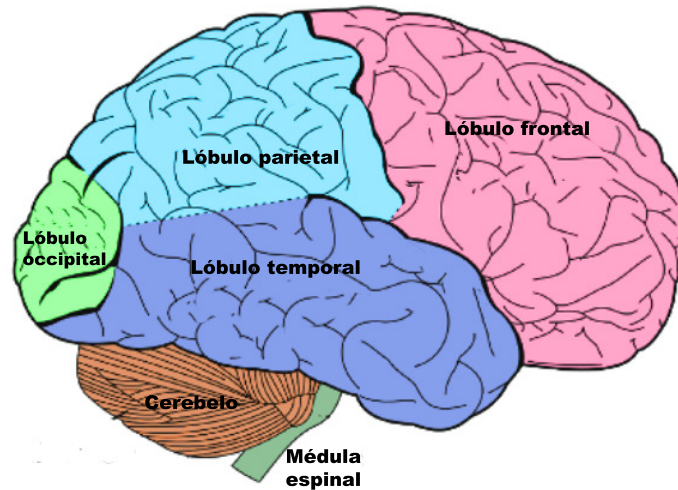


Figura 2.1 Vista general de la superficie de la corteza, obtenido de [20].

El papel funcional de las diferentes áreas de la corteza cerebral son descritas de la siguiente manera:

- Las áreas principales, que incluyen la corteza motora primaria y las áreas que reciben estímulos sensoriales: corteza somatosensorial primaria (lóbulo parietal) para información sensorial, corteza auditiva primaria (lóbulo temporal) y primaria lóbulo visual (lóbulo occipital);
- Las áreas secundarias, que corresponden a información elaborada del procesamiento que puede ser plurimodal, y áreas asociativas, cuyas funciones son funciones cognitivas y atencionales, las cuales prestan atención a los estímulos para identificarlos. Las funciones cognitivas son procesadas en tales áreas.

2.2. Métodos de neuroimagen

Muchas actividades cerebrales son generadas por acciones potenciales de una persona o con el cambio en el flujo sanguíneo; el registro de estas actividades puede

hacerse mediante el monitoreo de señales electrofisiológicas [2]. Existen dos clases comunes de tecnologías de imágenes cerebrales: las tecnologías invasivas, en las que los sensores se implantan directamente en el cerebro y las tecnologías no invasivas, que miden la actividad cerebral de una manera superficial sobre el cráneo. Las tecnologías invasivas proporcionan una resolución temporal y espacial alta, por lo general cubren solo regiones muy pequeñas del cerebro [21]. Los métodos más utilizados son: 1) electroencefalografía (EEG), 2) electrocorticografía (ECoG), 3) registros de una sola neurona, 4) magnetoencefalografía (MEG), 5) tomografía por emisión de positrones (PET, Positron Emission Tomography), 6) imagen de resonancia magnética funcional (fMRI, functional Magnetic Resonance Imaging), y 7) imágenes ópticas (es decir, funcional Near InfraRed (fNIR)). Sin embargo, MEG, PET, fMRI, y fNIR no son comúnmente utilizados porque dependen del metabolismo y son menos receptivos a la comunicación rápida; asimismo son caros y técnicamente exigentes. Con base en las recientes actividades de investigación de Interfaces Cerebro Computadora (BCI, Brain Computer Interfaces), solo electroencefalografía (EEG), electrocorticografía (ECoG) y las grabaciones de neurona única son los tres métodos válidos en aplicaciones basadas en estas interfaces. Actualmente, para sistemas BCI se usan equipos baratos y ofrecen buenos canales de comunicación entre el cerebro y el software para el control del sistema [2].

2.2.1. Electroencefalografía

La electroencefalografía es el registro de la actividad eléctrica a través del cuero cabelludo debido al disparo eléctrico de algunas de las neuronas en el cerebro. Esta actividad es grabada en un corto período de tiempo a través de múltiples electrodos ubicados en el cuero cabelludo directamente en la corteza. EEG se considera como el método más común para la grabación de señales cerebrales por alta resolución temporal (cantidad de datos registrados a lo largo del tiempo con una tasas de muestreo de entre 250 y 20000 Hz), fácil de usar, seguro y factible [22].

Los electrodos usados son de dos tipos, electrodos activos o pasivos; los electrodos pasivos requieren de un amplificador para incrementar las señales medidas, y los electrodos activos generalmente tienen un amplificador integrado. El objetivo principal del uso de amplificadores integrados o externos es la atenuación del ruido ambiental, señales adheridas debido al movimiento del cable o presencia de artefactos. Uno de los problemas en algunos dispositivos de EEG, es la conducción, provocando problemas de comunicación por parte de los dispositivos o electrodos, por lo que, para facilitar el

contacto se necesita un gel o líquido salino para reducir la impedancia del contacto del cable [2].

Para que una medición de señales sea confiable, la distancia entre electrodos debe tener entre 1 cm y 2 cm de separación para una señal con ruido bajo [2] (en promedio 2.3 a partir de la división de la señal por la desviación estándar de la señal de ruido) [23]. Además, EEG es capaz de detectar cambios en las señales cerebrales en milisegundos, y una acción requiere casi 0.5-130 ms para propagarse a través de una neurona [2].

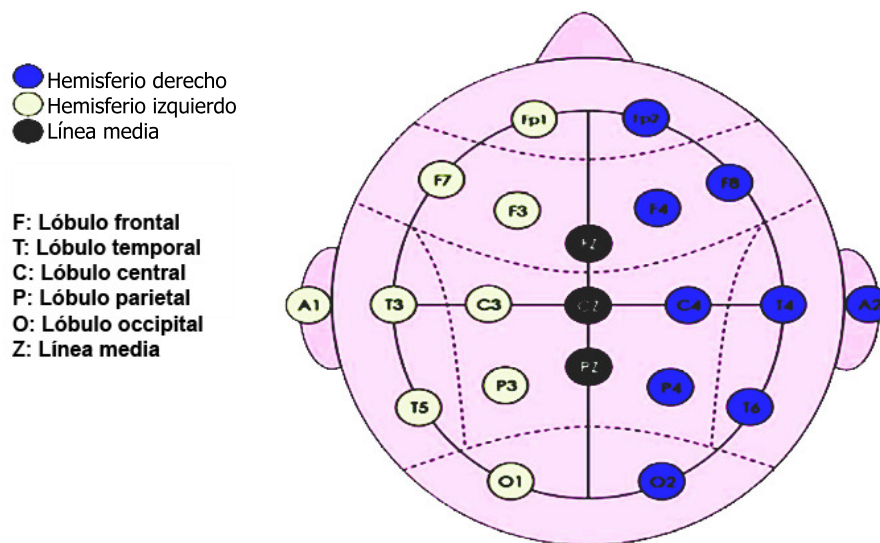


Figura 2.2 Sistema 10-20 y posiciones de los electrodos, tomada de [2].

Un estándar para la grabación de señales EEG, es el sistema 10-20, que se muestra en la figura 2.2, la numeración 10-20 se refiere a puntos de referencia anatómicos específicos o distancia entre electrodos; representando el 10-20 % de los perímetros de la cabeza de adelante hacia atrás o de derecha a izquierda. Otra alternativa, son los sistemas 10-10 y 10-5 que según Kumar y Bhuvaneshwari son un estándar válido [3]. Durante la toma de muestras el EEG es digitalizado normalmente con un convertidor analógico a digital (ADC) de 12 bits con frecuencia de muestreo que va desde 100 Hz a varios cientos de Hertzios. La figura 2.3 muestra un diagrama de bloques de uno de los ajustes de grabación de un solo canal EEG donde los filtros de paso alto y bajo se usan antes de las amplificaciones de las señales [3].

El filtro pasa bajas se usa para efectos de antialiasing y el filtro pasa altas se usa para eliminar artefactos de baja frecuencia. Las señales EEG generalmente se detallan en términos de tipos de actividad que son rítmicas y transitorias. Las actividades rítmicas se dividen en ciertas bandas de frecuencia, las cuales se ha demostrado que

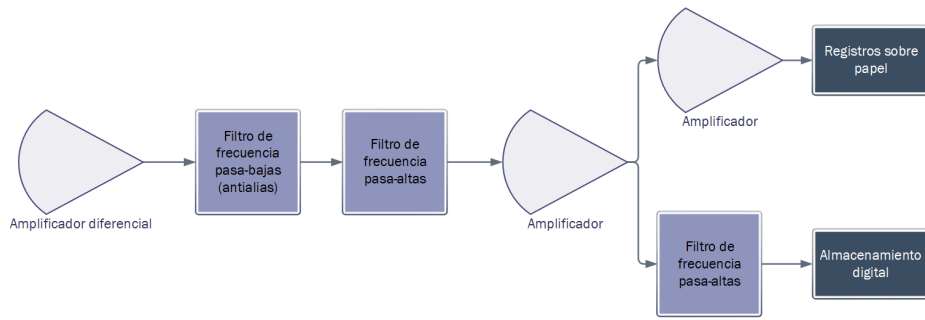


Figura 2.3 Diagrama de bloque de la grabación de un solo canal EEG.

Tabla 2.1 Características extraídas de las representaciones en tiempo-frecuencia.

Tipo	Rango de frecuencia (Hz)	Forma de onda	Propiedades	Actividad mental
Delta(Δ)	1-3		Frontalmente en adultos, posteriormente en niños; ondas de alta amplitud	Sueño de onda lenta (profundo) en bebés
Tetha(θ)	4-7		Línea media frontal (Fz a Cz)	Al ralentí, inconsciente, meditación y somnolencia
Alpha(α)	7-12		Regiones posteriores de la cabeza, ambos lados, mayor en amplitud en el lado dominante.	Relajación y concentración
Mu(μ)	8-13		Corteza Sensorimotora	La supresión indica que las neuronas motoras están funcionando
Beta(β)	12-30		Corteza sensoriomotora, entre C3 y C4, distribución simétrica, más evidente frontalmente ondas de baja amplitud	Alerta, pensamiento y concentración activa
Gamma(γ)	>30		Corteza somatosensorial	Procesamiento somatosensorial se muestra durante la coincidencia de memoria a corto plazo de objetos reconocidos, sonidos o sensaciones táctiles

tienen cierta importancia biológica o cierta distribución sobre el cuero cabelludo [26]. Las bandas de las señales EEG son nombradas con cierta nomenclatura; según su osculación las que caen debajo de 1 Hz y por encima de 20 Hz son consideradas como artefactos. La Tabla 2.1 compara entre todas las bandas de EEG en términos de frecuencia, forma, propiedades y actividades mentales. Como se indica en la tabla, hay seis tipos de señales que son Delta (δ), Theta (θ), Alpha (α), Mu (μ), Beta (β) y Gamma (γ). Delta (δ) tiene una amplitud variable y cae en el rango de 1-4 Hz [24]. Delta parece ser la más alta en amplitud y la onda más lenta asociada con los estados del sueño profundo y despertar. Theta (θ) tiene una amplitud que es mayor que $20 \mu V$ y está dentro del rango de 4-7 Hz. Theta se genera con la inspiración inactiva, creativa, inconsciente material, somnolencia y meditación profunda. Alpha (α) es una onda con amplitud de $30-50 \mu V$ y velocidad de cambio entre 8 y 13 Hz, es usualmente asociada con la relajación, concentración y algunas veces atención. Mu (μ), se encuentra en el rango de frecuencia de ondas alfa donde la amplitud registrada sobre la corteza motora es máxima, ordinariamente, es asociada con la supresión, la cual indica que las neuronas motoras están funcionando. Beta (β) se asocia con alerta, pensamiento y concentración activa y cae en el rango entre 12 y 30 Hz. Finalmente, Gamma (γ) podría ser detectada en la corteza somatosensorial con frecuencia mayor a 30 Hz. También se muestra durante la coincidencia de memoria a corto plazo de objetos reconocidos, sonidos o sensaciones táctiles [17, 2, 25].

2.3. Epilepsia

La epilepsia es una de las enfermedades neurológicas más comunes y devastadoras experimentadas en todo el mundo. La enfermedad se distingue en las personas que la padecen porque presentan convulsiones recurrentes denominadas “ataques epilépticos”. Las convulsiones se definen como cambios repentinos en el funcionamiento eléctrico del cerebro, lo que resulta en alteraciones y comportamientos tales como perder la conciencia, movimientos espasmódicos, pérdida temporal del aliento y de la memoria; estos generalmente ocurren en la corteza o fuera del borde del cerebro. La epilepsia puede desarrollarse debido a una anomalía en la red de interconexión cerebral, desequilibrio de sustancias químicas de señalización nerviosa llamadas neurotransmisoras o alguna combinación de estos factores [26, 25].

La figura 2.4 muestra la zona del cerebro donde se presentan los impulsos eléctricos anormales durante una convulsión. Las neuronas normalmente generan impulsos electroquímicos que actúan sobre otras neuronas, glándulas y músculos para producir

pensamientos, sentimientos y acciones humanas. En la epilepsia el patrón normal de actividad neuronal se altera, causando sensaciones extrañas, emociones y comportamientos o a veces convulsiones, espasmos musculares y pérdida de la conciencia. Puede haber una especie de “tormenta” eléctrica breve que surja de neuronas que son inherentemente inestables debido a un defecto genético (como en los diversos tipos de epilepsia hereditaria) o de neuronas que se vuelven inestables debido a anomalías metabólicas, como niveles bajos de glucosa. Por otro lado, la descarga anormal puede provenir de un área localizada del cerebro (este es el caso de pacientes con epilepsia causada por una lesión en la cabeza o un tumor cerebral). Durante una convulsión, las neuronas pueden disparar hasta 500 veces por segundo, mucho más rápido que lo normal (1-100 μV). En algunas personas, esto sucede solo de vez en cuando; para otros, puede suceder hasta cientos de veces al día[22, 26].

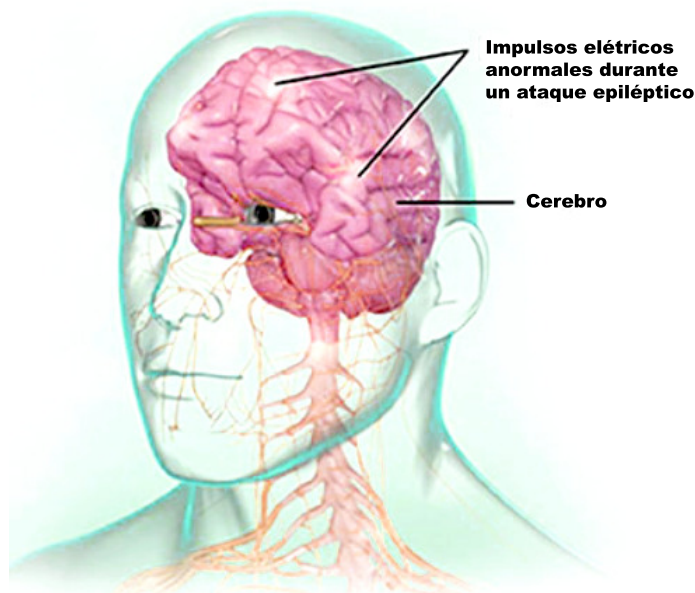


Figura 2.4 Impulsos anormales durante ataques tomada de [22].

2.4. Narcolepsia

En los humanos, la narcolepsia es un trastorno del sueño debilitante e importante que fue descrito por primera vez en 1877 por Westphal, un médico alemán, basándose en su observación de un paciente con ataques repentinos de sueño asociados con síntomas de incapacidad motora y afasia. Él lo consideró una forma de epilepsia; tres años más tarde, el neurólogo francés Gélinau describió el trastorno por primera vez como una

entidad clínica distinta en lugar de un síntoma de otra condición y propuso el término narcolepsia, que significa “convulsión del sueño”[27].

La narcolepsia se caracteriza por dos síntomas principales, somnolencia diurna excesiva y cataplejía, y tres síntomas suplementarios, alucinaciones hipnagógicas o hipnopómpicas, parálisis del sueño y sueño nocturno fragmentado. Desde la perspectiva del trabajo realizado en roedores, el síntoma más importante es la cataplejía. Un episodio catapléjico se describe como una disminución o pérdida abrupta y reversible en el tono muscular sin alteración en el nivel de conciencia. Este es un síntoma definitorio de la narcolepsia, aunque debe tenerse en cuenta que una definición precisa de cataplexia sigue siendo debatible ya que también se describe como una transición entre un estado mixto de vigilia y sueño de movimientos oculares rápidos (REM). Esencialmente, todos los episodios en ratones cambian muy rápidamente al sueño REM, para los ratones, los términos episodio catapléjico y episodio narcoléptico son intercambiables [27].

2.5. Transformada de Fourier de tiempo corto

El análisis de las señales en el dominio de tiempo-frecuencia, comúnmente es usando espectrogramas; por medio de la aplicación de la transformada de Fourier de tiempo corto a una señal dividida. Para la generación de los espectrogramas, la STFT se aplica a una señal, dividiéndola en varios segmentos de señales de tiempo corto desplazando una ventana de tiempo con algunas superposiciones, un proceso llamado ventaneo. Dependiendo de la función de ventana de tiempo $w[n]$, un espectrograma se clasifica como banda estrecha o banda ancha. Si la ventana de tiempo es corta, entonces su transformada de Fourier será una banda ancha y un tiempo más largo devuelve un espectrograma de banda estrecha [28]. La ecuación general STFT de una señal S está dada por la ecuación 2.1:

$$S(m, k) = \sum_{i=0}^{n-1} s(n + mN')w(n)e^{-j\frac{2\pi}{N}nk} \quad (2.1)$$

Donde $k = [0 : K]$ es el K^{th} coeficiente de Fourier.

$K = N/2$ es el índice de frecuencia correspondientes a la frecuencia de Nyquist.

$S(m, k)$ indica el índice-m tiempo-frecuencia(frame) del espectrograma.

N = longitud del segmento de la ventana.

N' =el paso de desplazamiento de la ventana de tiempo.

N' debe ser menor que N para producir una superposición entre las ventanas de

tiempo.

S depende de la función de ventana; en la práctica, se usan diferentes formas de ventanas, tales como: simétrico, unimodal y gaussiano.

El espectrograma presenta una relación entre la resolución de tiempo y la resolución de frecuencia: una ventana grande proporciona menor resolución en el tiempo y una mejor visualización en la frecuencia. La ventana obtiene un segmento de tiempo de la señal, durante el cual las características espectrales son casi constantes; los segmentos obtenidos cambian la ventana de tiempo con cierta superposición. El espectrograma se define como la magnitud de $S(m, k)$, representada como $A(m, k)$, como se muestra en la ecuación 2.2:

$$A(m, k) = \frac{1}{N} |S(m, k)|^2 \quad (2.2)$$

La resolución del espectrograma se puede mejorar modificando la longitud de la ventana; una ventana ancha proporcionará una mejor resolución de frecuencia, pero baja resolución de tiempo. Sin embargo, una ventana angosta da una resolución de tiempo mejor con una resolución baja en la frecuencia. Una buena visualización en el espectrograma depende de la selección de una longitud de ventana apropiada. La figura 2.5 muestra el espectrograma de una señal, que es una representación espectral variable en el tiempo de una señal. El diseño de un espectrograma suele ser el siguiente: el eje horizontal representa el tiempo, el eje vertical representa la frecuencia, y la tercera dimensión es la amplitud (contenido espectral) de un par frecuencia-tiempo, codificado por colores. Estos datos tridimensionales también pueden crear una trama 3D, donde la intensidad de energía se representa por medio de los colores, en donde los más claros representan los valores más altos de energía y los más oscuros corresponden a valores menores, una gráfica 2D proporciona una alternativa de visualización [18].

2.6. Aprendizaje automático

El aprendizaje automático es una rama de la inteligencia artificial centrada en entender cómo las computadoras pueden aprender (o mejorar su rendimiento) en función de datos. Un área de investigación principal está enfocada en que programas computacionales aprendan automáticamente a reconocer patrones complejos y tomen decisiones inteligentes basadas en datos [29].

En el desarrollo de esta investigación se propusieron tres enfoques, cada uno de estos evalúa la calidad de las características obtenidas por medio de tres diferentes algoritmos de aprendizaje automático: redes neuronales, máquinas de vectores de

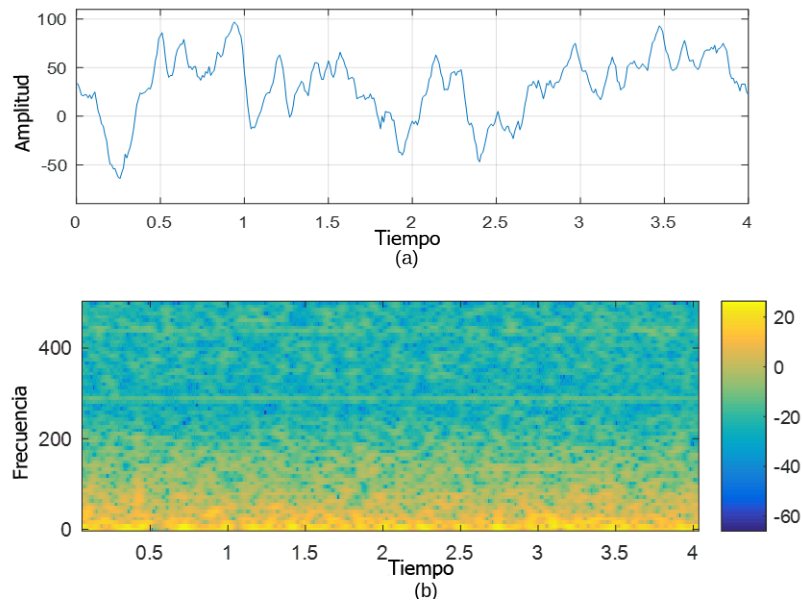


Figura 2.5 a) Señal EEG en el dominio del tiempo, b) Espectrograma de la señal EEG.

soporte y k-vecino cercanos; cada uno de estos seleccionados a partir de lo reportado en la literatura analizada y mostrada en capítulo 3. Y por otro lado, el último enfoque desarrollado, además de usar clasificadores para la evaluación de las características, hace uso de un clasificador no supervisado para la agrupación de puntos en común y así elegir uno de ellos como el más relevante el cual se usa como característica. Estos algoritmos de aprendizaje automático, son descritos en los apartados siguientes.

2.7. Aprendizaje supervisado

El aprendizaje supervisado son un conjunto de técnicas para realizar predicciones basadas en comportamientos. La supervisión en el aprendizaje proviene de los ejemplos etiquetados en el conjunto de datos de entrenamiento. Por ejemplo, en el problema de reconocimiento de código postal, se utilizan un conjunto de imágenes de códigos postales escritas a mano y sus correspondientes traducciones legibles por máquina como ejemplo de entrenamiento, que supervisan el aprendizaje del modelo de clasificación [29].

2.7.1. Redes neuronales

Una neurona artificial es una unidad de procesamiento de información que es fundamental para el funcionamiento de una red neuronal artificial (ANN, Artificial

Neural Network). Este es un modelo de una neurona artificial, que consta de tres elementos básicos:

1. Un conjunto de enlaces de conexión de diferentes entradas x_i (o sinapsis), cada una de las cuales se caracteriza por un peso o fuerza w_{ki} . El primer subíndice se refiere a la neurona en cuestión y el segundo se refiere a la entrada de la sinapsis a la que se refiere el peso. En general, los pesos de una neurona artificial pueden estar en un rango que incluye valores negativos y positivos.
2. Un sumador de las señales de entrada x_i ponderadas por las respectivas intensidades sinápticas w_{ki} . La operación descrita aquí constituye un combinador lineal.
3. Una función de activación f para limitar la amplitud de la salida y_k de una neurona. El modelo de la neurona mostrado en la figura 2.6 también incluye un sesgo aplicado externamente, denotado por b_k . El sesgo tiene el efecto de aumentar o disminuir la entrada neta de la función de activación, dependiendo de si es positiva o negativa.

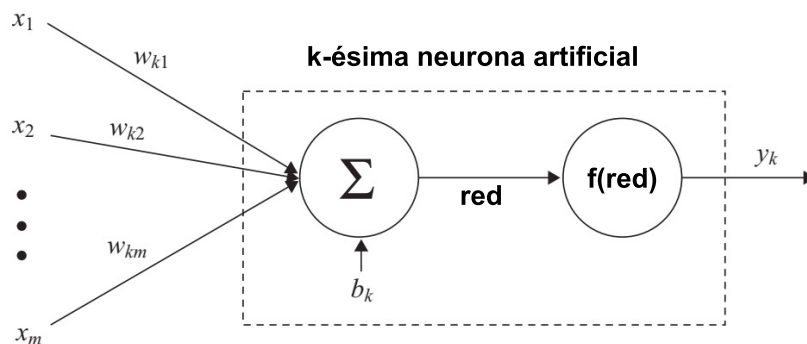


Figura 2.6 Modelo de una neurona artificial.

2.7.2. Perceptrón multicapa

El perceptrón multicapa (MLP, MultiLayer Perceptron) es un tipo de red neuronal artificial, con distinto número de capas ocultas, está compuesta por una capa de entrada para recibir la señal, una capa de salida que toma una decisión o predicción sobre la entrada y entre ambas capas, un número arbitrario de capas ocultas para los cálculos del MLP. Los perceptrones multicapa comúnmente se aplican a problemas

de aprendizaje supervisado: entrenan en un conjunto de pares de entrada-salida y aprenden a modelar la correlación (o dependencias) entre esas entradas y salidas. La definición de sus parámetros implica los pesos y sesgos del modelo para minimizar el error. La retropropagación se utiliza para hacer que los ajustes de ponderación y sesgo se relacionen con el error, y el error en sí se puede medir de varias formas, incluido el error cuadrático medio (RMSE) [29].

Las redes prealimentadas como las MLP son como el tenis o el ping pong, están principalmente involucrados en dos movimientos, un constante ir y venir. En la programación hacia adelante, el flujo de señal se mueve desde la capa de entrada a través de las capas ocultas hasta la capa de salida y la decisión de la capa de salida se mide con respecto a las etiquetas o función objetivo [29].

Por otro lado, utilizando una retropropagación, la regla de cadena de cálculo, derivadas parciales de una función de error, los diversos pesos y sesgos se propagan a través del MLP. Este proceso de diferenciación proporciona ajustes en pesos que resultan en un paso más cerca del mínimo de error. Esto se puede hacer con cualquier algoritmo de optimización basado en gradiente, como el descenso de gradiente estocástico. La red sigue jugando ese juego de tenis hasta que se alcanza un umbral de error, lo cual se conoce como convergencia [29].

2.7.3. Máquinas de vectores de soporte

Las máquinas de vectores de soporte (SVM, support vector machine) son un algoritmo de aprendizaje automático supervisado para clasificación o regresión. En este algoritmo, cada elemento de datos se considera como un punto en el espacio de n dimensiones (donde n es el número de características que tiene) con el valor de cada característica que es el valor de una determinada dimensión. Luego, se realiza la clasificación al encontrar el hiperplano de separación que diferencia las clases de una mejor forma [29].

Los vectores de soporte son las coordenadas de la observación individual, las cuales son una frontera que mejor segrega las dos clases (hiperplano/línea).

La definición dada por Mohammed, M. et al., en [30], menciona que dados N vectores x_1, x_2, \dots, x_N , con cada vector x_j con m características $a_{j1}, a_{j2}, \dots, a_{jm}$ y pertenecen a una de dos clases C_1 o C_2 . Como se mencionó anteriormente, se dice que los datos son *linealmente separables* si es posible encontrar un hiperplano en el espacio característico que pueda separar las instancias de la clase C_1 y C_2 . Lo cual se hace posible, a través de un mapeo de las características a un espacio dimensional más alto, donde los datos se vuelven linealmente separables. A partir de los datos separables

linealmente, habrá una brecha entre las instancias de las dos clases dadas, esta se conoce como margen. Sería una ventaja si el margen entre las instancias de las dos clases es lo más amplio posible como muestra la figura 2.7. Los N vectores de características dados se seleccionan aleatoriamente para ser un conjunto de entrenamiento. El objetivo principal de la SVM es encontrar un hiperplano que clasifique todos los vectores de entrenamiento en las dos clases. Podría haber muchos hiperplanos que separan las instancias de las dos clases, pero se desea que el hiperplano elegido maximice el margen entre las dos clases.

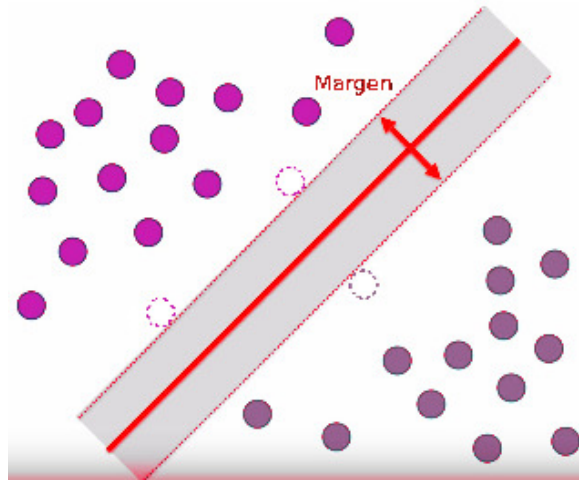


Figura 2.7 Margen de las máquinas de vectores de soporte.

La SVM, $S(x)$, es una función linealmente discriminante de la forma:

$$S(x) = w^T x + b \quad (2.3)$$

donde x es un vector de características, w es un vector m -dimensional de pesos, y el escalar b es un sesgo. El vector de peso es ortogonal al hiperplano y controla su dirección, mientras que el sesgo controla su posición. Ahora, si se le da un vector de características \bar{x} , entonces $S(\bar{x}) = w^T \bar{x} + b$ satisface:

1. $S(\bar{x}) = w^T \bar{x} + b > 0$ si \bar{x} es una instancia en C_1
2. $S(\bar{x}) = w^T \bar{x} + b < 0$ si \bar{x} es una instancia en C_2

La etapa de entrenamiento de la SVM incluye ajustar los pesos y el sesgo, de tal manera que todas las instancias de C_1 se encuentren en un lado del hiperplano y las instancias de C_2 se encuentren en el otro lado del hiperplano.

El objetivo principal de la SVM es maximizar el margen, que es la distancia entre el hiperplano y los vectores más cercanos a él desde ambas clases (vectores de soporte),

de manera que las instancias de las clases C_1 y C_2 están igualmente separadas del hiperplano.

2.8. Aprendizaje no supervisado

El aprendizaje no supervisado es esencialmente un sinónimo de agrupamiento. El proceso de aprendizaje no está supervisado ya que los ejemplos de entrada no están etiquetados por clase. Por lo general, se puede usar el agrupamiento para descubrir clases dentro de los datos. Por ejemplo, un método de aprendizaje no supervisado puede tomar, como entrada, un conjunto de imágenes de dígitos escritos a mano, supongamos que encuentra 10 grupos de datos, estos clústeres pueden corresponder a los 10 dígitos distintos de 0 a 9, respectivamente. Sin embargo, dado que los datos de entrenamiento no están etiquetados, el modelo aprendido no puede decirnos el significado semántico de los grupos encontrados [29].

En las subsecciones siguientes, se presentan los conceptos de los métodos no supervisados que respectan a esta tesis, en los cuales se basa el enfoque propuesto en este trabajo de investigación.

2.8.1. K medias

El algoritmo K medias (K-means partitional-clustering) o K means es un algoritmo de partición basado en optimización de errores cuadrados respecto a la pertenencia de elementos a los centroides. Comienza con una partición inicial aleatoria y sigue reasignando las muestras a los clusters (por su definición en inglés), en función de la similitud entre las muestras y los clusters, hasta que se cumple un criterio de convergencia. Normalmente, este criterio se cumple cuando no hay reasignación de ninguna muestra de un cluster a otro que cause una disminución del error cuadrado total. Un problema importante con este algoritmo es que es sensible a la selección de la partición inicial y puede converger a un mínimo local de la función de criterio si la partición inicial no se elige correctamente [29]. Básicamente, la ejecución de K medias empieza con la especificación del número K de cluster deseados; posteriormente, se asigna cada observación a exactamente uno de los clusters K.

Fundamentalmente, los pasos del algoritmo K media se dividen en dos. Inicialmente, se selecciona un aleatoriamente un número de 1 a K , para la asignación de cada una de las observaciones; posteriormente, se itera hasta que las asignaciones de cluster dejen de cambiar. Para cada uno de los cluster K, se calcula el centroide de estos. El

k-ésimo centroide del cluster es el vector de las medias de la característica p de las observaciones en el k-ésimo cluster. Finalmente, asignar cada observación a el cluster cuyo centroide sea el más cercano (donde la distancia más cercana se define por la Euclídeana)

2.8.2. DBSCAN

El algoritmo de agrupamiento espacial basado en densidad de aplicaciones con ruido (DBSCAN, Density-Based Spatial Clustering of Applications with Noise) es un algoritmo de agrupamiento basado en densidad. El algoritmo hace crecer regiones con una densidad suficientemente alta en grupos y descubre grupos de formas distintas en los conjuntos de datos. Define un grupo como un conjunto máximo de puntos densamente conectados [31]. Las ideas básicas que envuelven un cluster basado en densidad son las siguientes:

- La vecindad dentro de un radio ϵ de un objeto dado se llama la vecindad ϵ del objeto.
- Si la vecindad ϵ de un objeto contiene al menos un número mínimo, $MinPts$, de objetos, entonces el objeto se llama un objeto central.
- Dado un conjunto de objetos, D , se dice que un objeto p es directamente denso-alcanzable a la densidad desde el objeto q si p está dentro del vecindario ϵ de q , y q es un objeto central.
- Un objeto p es denso-alcanzable a la densidad desde el objeto q con respecto a ϵ y $MinPts$ en un conjunto de objetos, D , si hay una cadena de objetos p_1, \dots, p_n , donde $p_1 = q$ y $p_n = p$ tal que $p_i + 1$ es directamente denso-alcanzable desde p_i con respecto a ϵ y $MinPts$, para $1 \leq i \leq n$, $p_i \in D$.
- Un objeto p es densamente conectado al objeto q con respecto a ϵ y $MinPts$ en un conjunto de objetos, D , si hay un objeto $o \in D$, de manera que tanto p como q sean accesibles a la densidad desde o con respecto a ϵ y $MinPts$.
- Si el objeto está densamente conectado, este y los puntos previos son asignados a un cluster. Dentro de estos entran los puntos borde, que son los que usualmente se encuentran en los extremos, aunque no cumplen el criterio de núcleo son puntos directamente alcanzables por la densidad.
- Aquellos objetos no densamente conectados, son considerados como ruido.

La figura 2.8 ejemplifica el resultado del algoritmo, en donde cierto número de elementos forman tres clusters, a partir de un ε y un mínimo de puntos $MinPts = 3$. Los clusters los conforman los elementos dentro de los círculos con la línea punteada, los círculos más pequeños de color blanco son elementos de ruido y los de borde son los círculos negros (cercanos a la circunferencia punteada).

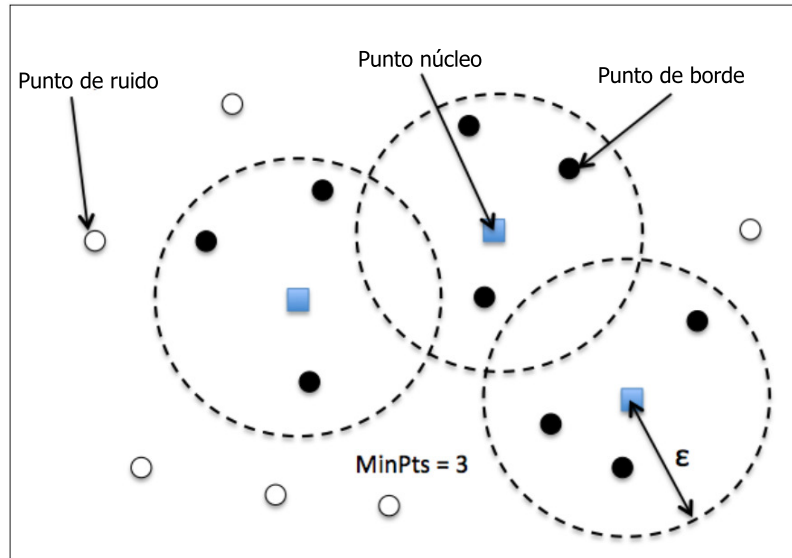


Figura 2.8 Puntos de DBSCAN, tomado de [32].

2.8.3. Técnica de sobre-muestreo minoritaria sintética

Uno de los problemas comunes dentro de la clasificación se presenta cuando los datos con los que se trabaja no están balanceados, es decir, los ejemplos por clase son diferentes en cantidad. Este problema se refleja en los resultados de la clasificación, comúnmente, en un aprendizaje sesgado, perjudicando a la clase minoritaria. Para enfrentar esta problemática existen técnicas de balanceo que generan datos sintéticos con la finalidad de tener un número muy parecido entre los ejemplos de cada clase.

La técnica de balanceo SMOTE (por sus siglas en inglés, Synthetic Minority Over-sampling Technique) propuesta por Chawla et al. [33], consiste en que la clase minoritaria se sobremuestra (proceso en el que se muestran elementos de una clase con la finalidad de ajustar su distribución) creando ejemplos sintéticos. Se forman datos de entrenamiento adicionales al realizar ciertas operaciones en datos reales; la clase minoritaria se sobremuestra tomando cada muestra de clase minoritaria e introduciendo ejemplos sintéticos a lo largo de los segmentos de línea que se unen

a cualquiera o todos los vecinos más cercanos de la clase minoritaria. Dependiendo de la cantidad de sobremuestreo requerida, los vecinos más cercanos son elegidos aleatoriamente, la técnica SMOTE solo son cinco. Si la cantidad de sobremuestreo necesaria es del 200 %, solo se eligen dos vecinos de los cinco vecinos más cercanos y se genera una muestra en la dirección de cada uno. Las muestras sintéticas se generan de la siguiente manera: inicialmente, se toma la diferencia entre el vector de características (muestra) y su vecino más cercano; posteriormente, se multiplica esta diferencia por un número aleatorio entre 0 y 1 y se agrega al vector de características en consideración. Esto provoca la selección de un punto aleatorio a lo largo del segmento de línea entre dos características específicas. Este enfoque obliga efectivamente a la región de decisión de la clase minoritaria a ser más general.

2.9. Filtros Gabor

Los filtros Gabor son una técnica introducida por Daugman para analizar texturas en una imagen. Estos modelan la respuesta de las células corticales simples en el sistema visual y extraen los contornos de la imagen y las texturas [34]. Los filtros presentan diferentes ventajas, ya que enfrentan diferentes escenarios ante las imágenes como: la rotación, escala, traslación, perturbaciones fotométricas, iluminación y ruido de la imagen [35]. Un filtro de Gabor 2D es una función con núcleo gaussiano, modulada por una onda plana sinusoidal compleja [34] y puede definirse como:

$$G(x, y) = \frac{f^2}{\pi\gamma} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp(j2\pi f x' + \phi), \quad (2.4)$$

donde

$$x' = x \cos\theta + y \sin\theta \quad (2.5)$$

$$y' = -x \sin\theta + y \cos\theta \quad (2.6)$$

f es la frecuencia de la sinusoide.

θ representa la orientación de las bandas normales en relación con las bandas paralelas de la función de Gabor.

ϕ es el desplazamiento de fase.

σ es la desviación estándar de la envolvente gaussiana.

γ es la relación de aspecto espacial que especifica la elipse de soporte de la función de Gabor.

2.10. Envolverte convexo

Un subconjunto U del plano se llama convexo si y solo si para cualquier par de puntos p, q en U el segmento de línea \bar{pq} está completamente contenido en U . El casco convexo $CH(U)$ de un conjunto U es el conjunto convexo más pequeño que contiene U . Para ser más precisos, es la intersección de todos los conjuntos convexos que contienen U como se muestra en la figura 2.9 [36].

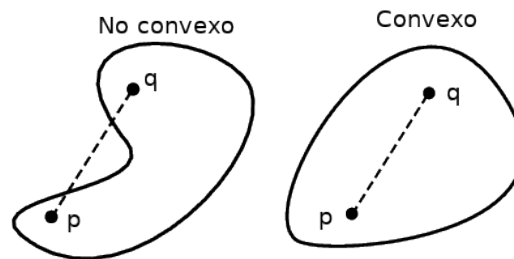


Figura 2.9 Puntos convexos y no convexos.

El envolverte convexo de un conjunto de puntos como el conjunto convexo más pequeño que contiene los puntos o, más precisamente, como la sección intermedia de todos los conjuntos convexos que contienen los puntos. El casco convexo es una construcción fundamental para las matemáticas y la geometría computacional y su uso particular es en el análisis de datos de espectrometría. Otros problemas pueden reducirse al casco convexo, por ejemplo, intersección de medio espacio, triangulación de Delaunay, diagramas de Voronoi y diagramas de potencia [36].

2.11. Patrón binario local y patrón ternario local

El patrón binario local (LBP) por sus siglas en ingles Local Binary Pattern es la característica de textura más popular en el rubro de las imágenes digitales, es robusto a las variaciones de iluminación y contraste, ya que solo considera los signos de las diferencias de píxeles. Los códigos de histograma LBP hacen que el descriptor sea resistente a las codificaciones dentro del vecindario de histograma. Sin embargo, es sensible al ruido y pequeñas fluctuaciones de los valores de píxeles; para contrarrestar esto, Tang y Triggs propusieron un Patrón Ternario Local (LTP) [37], en comparación con LBP, tiene 2 umbrales que crean 3 estados diferentes en comparación con 2 en LBP. Es más resistente al ruido y pequeñas variaciones de valor de píxeles en comparación con LBP. Al igual que LBP, también se ha utilizado para la clasificación de texturas y

la detección de rostros [38]. El cálculo de la codificación LBP en la coordenada espacial (x, y) se realiza como se indica en la ecuación 2.7:

$$LBP_{x,y} = \sum_{b=0}^{B-1} s(p_b - p_c)2^b, s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \quad (2.7)$$

donde p_c es el valor del pixel en (x, y) , p_b es el valor del pixel estimado usando interpolación bilineal de los vecinos en la localización b del círculo de radio R alrededor de p_c y B es el número total de pixeles vecinos. Un bloque de histograma $2^B bin$ se calcula a partir del umbral z .

LBP es invariante a los cambios de intensidad monotónicos. Por lo tanto, es robusto a las variaciones de iluminación y contraste. Sin embargo, es sensible al ruido y pequeñas fluctuaciones de valor de pixeles. Por lo tanto, LTP propuso para considerar estas sensibilidades. El código LTP en la ubicación (x, y) se calcula de la siguiente manera:

$$LTP_{x,y} = \sum_{b=0}^{B-1} s'(p_b - p_c)3^b, s'(z) = \begin{cases} 1, & z \geq T \\ 0, & -T < z < T \\ -1, & z \leq -T \end{cases} \quad (2.8)$$

donde T es un umbral definido por el usuario. Según lo definido por $s'(z)$, LTP tiene 3 estados mientras que LBP tiene dos. Se calcula un histograma de bloque $3^B bin$. Para $B = 8$, el histograma tiene 6561 contenedores, que es de muy alta dimensión. Por lo tanto, diferentes autores proponen dividir el código LTP en sus códigos LBP superiores e inferiores [38].

2.12. Métricas de evaluación

En esta sección se describen algunas de las métricas utilizadas para la evaluación de los resultados obtenidos. Debido a que el objetivo de este trabajo consiste en el desarrollo de una metodología de extracción de características genérico con diferentes aplicaciones, es decir, las clases tienen la misma importancia, por lo tanto, la exactitud se usó como métrica principal. La exactitud de un clasificador en una prueba dada, es el porcentaje de predicciones que un modelo previamente entrenado realizó correctamente. La exactitud es un complemento de la medida de tasa de error [39], la cual se define como el porcentaje de las predicciones hechas incorrectamente. La siguiente ecuación

(2.9) define la exactitud de una clasificación binaria.

$$Exactitud = \frac{TP + TN}{P + N} \quad (2.9)$$

En donde TP y TN son los ejemplos clasificados correctamente, a partir de una clase positiva y negativa, respectivamente. P corresponde al número total de ejemplos de la clase positiva clasificados correcta e incorrectamente y N son el mismo total como en P , de la clase negativa. Cuando existen más de dos clases, se lleva a cabo el mismo procedimiento, el numerador ($TP + TN$) de la ecuación 2.9 son los ejemplos que se clasificaron correctamente para cierta clase y el denominador ($P + N$) representa el total de ejemplos. Los valores obtenidos a partir de la fórmula están en un rango de 0 a 1, donde el valor más bajo 0 corresponde a predicciones nulas y 1 cuando todas las predicciones se acertaron; para obtener el valor en términos de porcentaje, el resultado obtenido de la ecuación 2.9 es multiplicado por 100.

2.12.1. Validación cruzada

Una de las técnicas para hacer más confiable un clasificador es la validación cruzada, la cual estima la exactitud a partir de particiones muestreadas aleatoriamente de un conjunto de datos.

Han y Kamber [40] mencionan que en la validación cruzada k-fold, los datos iniciales se dividen aleatoriamente en k subconjuntos mutuamente excluyentes, D_1, D_2, \dots, D_k , cada uno de aproximadamente el mismo tamaño, las pruebas se realizan k veces. En la iteración i , la partición D_i se reserva como el conjunto de prueba y las particiones restantes se usan para entrenar el modelo. En la primera iteración, los subconjuntos D_2, \dots, D_k son el conjunto de entrenamiento para obtener un primer modelo, que se prueba con D_1 ; la segunda iteración se entrena en los subconjuntos D_1, D_3, \dots, D_k y se prueba con D_2 ; y así sucesivamente. Para la clasificación, la estimación de la exactitud se obtiene por medio de la ecuación 2.9 para cada iteración k . Finalmente, se calcula el promedio con las k exactitudes obtenidas.

2.12.2. Coeficiente Kappa

La validación cruzada en ocasiones suele presentar una desventaja, ya que durante el muestreo de los subconjuntos, algunas instancias o ejemplos pudieron no haber sido seleccionados o el caso contrario, seleccionados más de una vez. Un método que verifica los datos a través de sus exactitudes de clasificación es el coeficiente kappa, el cual

muestra qué tanta concordancia entre dos observadores, por lo que al aplicarlo, se encuentra cuanto coinciden en los resultados. Torres [41] menciona que los valores que puede tomar el coeficiente kappa son entre -1 y +1; si se acerca a 1 el grado de concordancia observada será mayor. Por otro lado, si el coeficiente tiene un valor negativo si el porcentaje observado es menor que el dado por azar [41]. Al comparar el resultado entre dos clasificadores, el comportamiento puede obtenerse por medio de lo siguiente como lo menciona Benítez [42]. A partir de los resultados de dos clasificadores F y G , la concordancia observada se calcula como

$$Pa = \frac{A}{N} \quad (2.10)$$

donde A es el número de ejemplos para los que ambos dieron la misma predicción y N es el número total de ejemplos.

El estadístico *kappa* se obtiene como:

$$k(F, G) = \frac{Pa - Pe}{1 - Pe} \quad (2.11)$$

donde Pe (la concordancia esperada) se calcula como:

$$Pe = \sum_{i=1}^{ns} \frac{P_f(i) \times P_g(i)}{N^2} \quad (2.12)$$

donde ns es el número de clases y $P_x(i)$ es el número de predicciones del clasificador X (con $X \in F, g$) de la clase i .

Para interpretar el coeficiente kappa se debe observar la tabla 2.2, en donde el coeficiente kappa obtenido se compara con el de la primera columna, el cual corresponde a un nivel de fuerza de la concordancia.

Tabla 2.2 Valores del coeficiente kappa tomado de Landis y Koch [43].

Estadístico kappa	Fuerza de la concordancia
≤ 0.00	Pobre
0.01 - 0.20	Leve
0.21 - 0.40	Aceptable
0.41 - 0.60	Moderada
0.61 - 0.80	Considerable
0.81 - 1.00	Casi perfecta

2.12.3. Prueba de significancia estadística Friedman

El análisis experimental del desempeño de algún modelo computacional es necesario, sin embargo, métricas al respecto tales como: la exactitud, sensibilidad, precisión, entre otras, no son suficientes por si solas. La comparación con otros algoritmos u otros clasificadores es importante, ya que refuerza los resultados obtenidos, verificando si existe diferencia estadísticamente significativa entre ellos, validando que los resultados obtenidos no son debidos al azar. Comúnmente, estas pruebas se hacen sobre la validación cruzada, usando las mismas particiones para los diferentes modelos, cada experimento de validación cruzada produce una estimación de error independiente diferente, el objetivo es la exactitud media en todos los conjuntos de datos posibles del mismo tamaño, y si esta media es mayor para un esquema u otro [44].

Existen diferentes pruebas estadísticas como la prueba de *Student* o *Mcnemar*, sin embargo, en específico para experimentos computacionales Garcia et al. [45], muestra un análisis experimental para múltiples tests con objetivos de inteligencia computacional y minería de datos. Una test de los más comunes es el de Friedman descrito en el trabajo de Garcia [45], el cual es un análogo no paramétrico del análisis de varianza paramétrico bidireccional. El objetivo de esta prueba es determinar si se puede concluir de una muestra de resultados que existe una diferencia entre los resultados mostrados. El primer paso para calcular el estadístico de prueba es convertir los resultados originales en rangos. Por lo tanto, se clasifican los algoritmos para cada problema por separado, el algoritmo de mejor desempeño debe tener el rango de 1, el segundo mejor rango 2 y así sucesivamente. En caso de empate, se calculan los rangos promedio.

Sea r_i^j el rango del j -ésimo de k algoritmos en el i -ésimo conjunto de datos. La prueba de Friedman requiere del cálculo de los rangos promedio de algoritmos a evaluar, $R_j = \frac{1}{n} \sum_i r_i^j$. Bajo la hipótesis nula, que establece que todos los algoritmos se comportan de manera similar y por lo tanto, sus rangos R_j deben ser iguales, la estadística de Friedman (ecuación 2.13) se distribuye de acuerdo con χ_F^2 con $k-1$ grados de libertad, cuando n y k son lo suficientemente grandes (como regla general, $n > 10$ y $k > 5$). Para un número menor de algoritmos y conjuntos de datos, se han calculado valores críticos exactos. Por otro lado se mostró [45] que Friedman χ_F^2 presenta un comportamiento moderado y propuso una mejor estadística (ecuación 2.14)

$$\chi_F^2 = \frac{12n}{k(k+1)} \left[\sum_j R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (2.13)$$

$$F_f = \frac{(n-1)\chi_F^2}{n(n-1)\chi_F^2} \quad (2.14)$$

que se distribuye de acuerdo con la distribución- F con $k-1$ y $(k-1)(n-1)$ grados de libertad. Para encontrar los valores críticos y determinar los resultados de hipótesis (resultados significativamente diferentes o no) la tabla descrita por Sheskin [46] debe ser revisada. La hipótesis que se suele probar es que todos los clasificadores obtienen el mismo error. Sin embargo, si se rechaza la hipótesis nula, se puede concluir que existen diferencias estadísticamente significativas entre los clasificadores.

Capítulo 3

Estado del arte

El análisis de una señal EEG comúnmente se componen de tres etapas: preprocesamiento, extracción/selección de características y clasificación, como se muestra en la figura 3.1 [4, 47, 3]. En la primera fase de preprocesamiento, se eliminan artefactos y ruido, en la segunda fase se extraen las características de la representación de STFT y se seleccionan las características más relevantes o significativas que construirán modelos de clasificación para la última etapa. En esta sección, se presentan los trabajos más relevantes relacionados con el tema propuesto. Las secciones están divididas en dos, una para preprocesamiento y otra para las dos fases restantes del análisis.

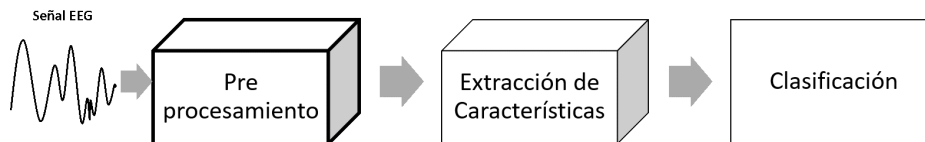


Figura 3.1 Etapas del análisis de señales EEG. Elaboración propia a partir de [3].

3.1. Preprocesamiento

La fase inicial del análisis de las señales EEG es el preprocesamiento; durante la adquisición de la señal, se pueden sumar señales no deseadas o que no estén relacionadas con la señal de objetivo. Esto se debe al instrumento con el que se adquiere la señal o a las señales provenientes de fuentes fuera del cerebro, como es el movimiento de músculos, latidos del corazón o movimiento de los ojos, estas señales son llamadas artefactos. Algunos métodos mencionados en la literatura para la resolución de este problema son: detección de artefactos, filtros espaciales, filtros espectrales, entre otros.

La tabla 3.1 muestra las características de los métodos mostrados en los artículos analizados.

El artículo [48] muestra los métodos que se llevan a cabo para el procesamiento de una señal EEG, que va desde la adquisición hasta la clasificación. En la parte del preprocesamiento se describen métodos como: Referencia de Promedio Común (CAR) el cual es un método no costoso computacionalmente para remover ruido, éste referencia la señal a un promedio común de todos los electrodos para substraer de cada muestra el valor promedio de la señal de todos los neurosensores; el promedio de toda la actividad representa un estimado de la actividad en el sitio de referencia y después se extrae este promedio que produce en principio una solución desreferenciada [20]. Al contrario de CAR, el método Superficie Laplaciana (SL) en una posición de electrodo específica se deriva restando una combinación de las señales de un conjunto de electrodos circundantes de la señal de electrodo central [20].

El análisis de Componentes Independientes (ICA) es una técnica estadística y computacional para revelar componentes ocultos que subyacen a conjuntos de variables, mediciones o señales aleatorias; una suposición común hecha cuando se aplica la técnica ICA es que las fuentes subyacentes desconocidas son independientes entre sí y se han combinado linealmente para formar una señal mixta, ICA devuelve los componentes independientes cuando esta suposición de independencia es correcta [20]. Patrones Espaciales Comunes (CSP) es una técnica de filtrado discriminativa específica para mejorar la señal, la cual detecta los patrones dentro de las señales complejas de EEG mediante la incorporación de información espacial de la EEG; ésto se emplea por medio de matrices de covarianza como base y buscando un subespacio discriminativo de tal modo que la varianza para una clase es maximizada y la varianza para otra clase es minimizada al mismo tiempo [20].

EL análisis de Componentes Principales (PCA) es una técnica estadística para identificar y volver a referenciar los datos mediante mapeo lineal, que transforma un número de variables posiblemente correlacionadas en un número menor de variables no correlacionadas conocidas como componentes principales; el eje primario se calcula de tal manera que representa la mayor cantidad de variabilidad en los datos. Los componentes subsiguientes se calculan similarmente y se tiene en cuenta la dirección de la variabilidad restante en orden decreciente con respecto a la cantidad de variabilidad en los datos [20].

Otra técnica utilizada es el filtro de Kalman Robusto, este es un enfoque de dos etapas de predicción y corrección, en la etapa de predicción, el estado se predice usando un modelo dinámico, mientras que en la segunda etapa, el estado predicho es corregido

usando un modelo de observación tal que la covarianza de error del estimador es minimizada; siendo recursivas, las nuevas mediciones pueden ser procesadas a medida que llegan [20].

Filtros Adaptativos es una técnica que tiene como objetivo modelar la relación entre dos señales de una manera iterativa, éste consiste de cuatro partes: la señal siendo procesada, la estructura que describe relación de entrada o salida, el parámetro que puede ser iterativamente variado para alterar la entrada del filtro o la relación de salida [47]. Finalmente, filtros digitales como pasa-baja, pasa-banda, pasa-alta, entre otros, han sido implementados para remover artefactos y ruido, los cuales ocupan distintas bandas de frecuencia para su diseño [47]. De los métodos mencionados [48] reporta que los más usados son ICA, CAR, SL, PCA, CSP y Filtros Adaptativos.

En [47] se proponen métodos muy parecidos a [48], sin embargo, ellos mencionan que el método de pre-procesamiento depende de: el tipo de señal, número de electrodos en el instrumento de adquisición, computacionalmente eficiente o no, entre otras desventajas y ventajas. En el trabajo se realiza una propuesta para clasificación de movimientos imaginarios, obteniendo resultados competitivos, para un conjunto de datos de dos clases, se obtiene una exactitud de 96 % y 100 % utilizando Análisis de Discriminante Lineal (LDA) y Máquinas de Vectores de Soporte (SVM) respectivamente, utilizando en su metodología standard Windowed Mean, que es un método de BCILAB que consiste en una herramienta para MALTAB por medio de filtros espectrales para el preprocesamiento de la señal, por lo tanto, se trabaja en términos de la frecuencia.

3.2. Extracción de características y clasificación

La fase de extracción de características es aplicada a la respuesta de la STFT, que comúnmente es representada como la magnitud del espectrograma o una imagen. En esta sección se presentan algunos trabajos relevantes para la extracción de características de diferentes representaciones en tiempo-frecuencia, así como la clasificación resultante de usar las características. Una de las metodologías iniciales basadas en el enfoque de espectrogramas de EEG, fue aplicado por Tzallas et al., [49] en el 2009, en donde mostró una alternativa para la detección de epilepsia, extrayendo de los espectrogramas densidad de potencia espectral y energía fraccional; a partir de ésta, surgieron otros métodos aplicados al enfoque de la STFT, los cuales son descritos en las siguientes subsecciones.

Tabla 3.1 Características de los métodos usados en el preprocesamiento a partir de [47] y [48].

Método	Características
ICA	Eficiente computacionalmente en tiempo de ejecución. Alto desempeño para datos de gran tamaño. No se aplica para todos los casos. Requiere muchos cálculos para la descomposición.
CAR	Mejora la SNR. Su desempeño es superior sobre todos los métodos de referencia. La cobertura completa causa errores.
SL	Robusto contra los artefactos generados en regiones no cubiertas por el electrodo. Sensible a artefactos y patrones ranura.
PCA	Reducción de la dimensión de las características. ICA se desempeña mejor en comparación de PCA.
CSP	Buen desempeño en imágenes motoras. No requiere selección de bandas específicas. Requiere gran cantidad de electrodos (>64).
Filtros adaptativos	Trabajan bien para señales con traslapes de espectro. Necesitan dos señales y una de referencia
Filtros digitales	Remueven fácilmente ruido eléctrico. Requieren múltiples frecuencias de la señal EEG.

3.2.1. Análisis de textura usando la matriz de co-ocurrencia de Grises

Un método estadístico para examinar texturas considera la relación de la matriz de co-ocurrencia de escala de grises (GLCM, gray level co-occurrence matrix), el cual es aplicado a los espectrogramas de las señales EEG. Un trabajo realizado por Mustafa et al., [1] muestra la clasificación de imágenes, en esta investigación los autores aplican STFT a la señales EEG provenientes de la resolución de cuestionario a diferentes sujetos; a partir de la imagen hacen un análisis de textura con el método de matriz de co-ocurrencia en escala de grises. En la extracción de características se obtuvieron 20 características con cuatro diferentes orientaciones, seguido de una reducción de características por medio de análisis de componentes principales (PCA), utilizando k -vecinos cercanos (kNN) para la clasificación. Se menciona que sus resultados son competentes al usar GLCM y kNN en comparación con otros. Los resultados de clasificación obtenidos fueron de 70.83 % para $k=4$ con 8 componentes principales para tres distintas clases. Otra propuesta similar se realizó posteriormente [50], comparando SVM y ANN como punto de interés; siguiendo un metodología similar a su anterior trabajo y complementando la extracción de características con una propuesta más (características estadísticas), teniendo como resultado una mejora en exactitud de un 87.5 % para ANN y 90 % para kNN con una distancia Euclidiana.

La investigación [10] propone un método para clasificar EEG con dos diferentes categorías (estados de sueño y ataques epilépticos), aplicando STFT a la señales EEG; las características se obtienen a partir de GLCM por medio de valores estadísticos, posteriormente éstas son codificadas por medio de vectores de Fisher. La clasificación es llevada a cabo con una máquina de aprendizaje extremo, resultando una exactitud de 95.17%,95.38 % para el conjunto de datos del sueño y 96.40 % para el conjunto de datos de ataques epilépticos.

En [51] se muestra una metodología en la que se obtiene un espectrograma aplicando STFT a la señal EEG para el estudio de ataques epilépticos en personas; las imágenes obtenidas son convertidas a escala de grises de 8 bits y divididas en sub-imágenes correspondientes a ciertas bandas de frecuencia. Matriz de co-ocurrencia en escala de grises (GLCM), método de codificación de la característica de textura (TFCM) y patrón binario local (LBP) son métodos empleados para la extracción de las características. Finalmente, para la clasificación con SVM y LIBLINEAR obtienen 100 % de exactitud para las características de GLCM (80), LBP (1280) y el peor resultado con TFCM (60) con una exactitud de entre 87 %-92 %.

La autenticación de las personas usa distintos biométricos (huella digital, voz, retina, entre otros). Una propuesta metodológica aplicada a esta área por [52] consiste en aplicar la STFT a señales EEG, posteriormente los espectrogramas resultantes son divididos en sub imágenes a diferentes frecuencias. La extracción de características es a través de energía y estadísticas de las imágenes; una selección de características es aplicada por medio de suma de distancias. Finalmente se aplica ANN y SVM para la clasificación, obteniendo resultados promedio con un porcentaje de exactitud de entre 90 %-98 %.

El trabajo presentado por [53] propone un sistema BCI que consiste en detectar imágenes motoras del movimiento de la mano izquierda y derecha. La metodología usada aplica STFT a la señal EEG, al espectrograma resultante se le aplica GLCM, para obtener diferentes descriptores estadísticos de textura, los cuales son ocupados para la clasificación utilizando un clasificador de regresión lineal con una exactitud del 87.6 %. Este sistema hace uso de un solo canal, está aplicado en tiempo real y es probado directamente con personas (no con datos previamente obtenidos).

3.2.2. Extracción de características de sub-imágenes

La señal EEG está definida por los tipos de ondas con las que cuenta, estas son diferenciadas por la frecuencia, las bandas están relacionadas a diferentes eventos que ocurren en el cuerpo humano. Diferentes autores hacen uso de estas frecuencias para

realizar representaciones en el tiempo y la frecuencia y así obtener sub-imágenes de las cuales se extraen las características como son: energía, datos estadísticos, entropía, entre otras. En la investigación [3] detectan ataques epilépticos usando transformada S la cual es una generalización de la transformada de Fourier, a la cual aplican una optimización del ancho de ventana; se calcula la densidad espectral y dividen el mapeo en bandas de frecuencia. La energía de cada sub-banda es sumada y se obtienen un vector de 15 características; posteriormente, se aplica PCA, seleccionando solo cuatro características. La clasificación se realiza con redes neuronales, dando como resultados 82.36 %, 99.33 % y 99.5 % para tres diferentes problemas.

En el trabajo desarrollado por [54] para la detección de ataques epilépticos, se aplica la Transformada de Hilbert-Huang (THH) para el mapeo en el tiempo y la frecuencia de la señal EEG, posteriormente, se hace una segmentación basados en las bandas de frecuencia de los ritmos de una señal EEG, obteniendo distintas subimágenes; la extracción de características es por medio de estadísticas del histograma de la imagen en escala de grises. Finalmente, para la clasificación se aplica SVM obteniendo una exactitud del 99.12 %.

El artículo [24] muestra cómo se detectan los estados somnolientos de las personas. La metodología propuesta consiste en extraer características estadísticas en el dominio del tiempo, se agrega también la potencia de distintas bandas de frecuencia, provenientes de aplicar STFT, dando un total de 52 características; Random forest se aplica para seleccionar características. Finalmente, la clasificación se lleva a cabo con kNN con una función de árbol kd, obteniendo diferentes resultados de exactitud, con un mínimo de 88.57 % y un máximo de 91 % para distinta número de características.

Diferentes representaciones se aplican a la señal EEG para su análisis; el trabajo desarrollado por [14] emplea una Distribución de Tiempo Frecuencia Adaptativa (ADTFD), Transformada Discreta de Wavelet (WVD), Transformada de Fourier de Tiempo Corto (STFT), Distribución B Modificada extendida (EMBD), para cada representación se extraen características estadísticas y se aplica SVM, obteniendo un 97.5 % para la función adaptativa sobre las demás distribuciones. Otra comparativa desarrollada por [16] aplica: espectrograma, distribución Choi-Williams y distribución Wigner Vile Pseudo Suavizada (SPWVD). Las representaciones son aplicadas a señales EEG de ataques epilépticos; ésta metodología hace uso de entropía de Rényi marginal para detectar picos de ataques, además por medio de una división para distintas bandas de frecuencia. Concluyendo que la mejor distribución es SPWVD por la concentración de energía que muestra basados en la entropía de Rényi.

3.2.3. Otras técnicas

Los métodos propuestos por sus autores en las secciones anteriores llevan una secuencia de pasos muy similar en el análisis de la señal. Los trabajos presentados en esta sección utilizan otras técnicas que se aplican directamente a la señal EEG para después hacer el mapeo, o directamente a la representación para extraer coeficientes que no son estadísticos como en la sección anterior. Un experimento realizado por [55] para clasificar el ritmo con el cual sílabas imaginadas se producen, procesaron una alta densidad de datos, a los que se les aplicó un algoritmo de identificación ciega de segundo orden modificada (SOBI) para reducir su dimensión y remover artefactos. Además de usar estructura temporal consistente para descomponer ciegamente los datos. Estas características son unidas con las de tiempo-frecuencia que provee la transformación de Hilber-Huang. Se menciona que los experimentos realizados son prometedores para aplicaciones de habla imaginada, ya que obtienen una exactitud de 68 % - 73.63 % usando un clasificador bayesiano basado en análisis de discriminante lineal.

Una investigación realizada por [56] compara los componentes obtenidos durante la escritura física y la escritura imaginada que muestra una señal EEG. La señal es filtrada como parte del preprocesamiento, se le aplica STFT y se modifica el espectrograma para solo quedar con los componentes en términos de la frecuencia, finalmente concluyen los componentes de la escritura en ambas representaciones mantienen una fuerte correlación.

El trabajo realizado por [25] propone una metodología para la identificación y monitoreo de actividad cerebral, específicamente para el análisis de señales EEG de personas con ataques epilépticos. Este experimento aplica un filtro pasa baja para remover ruido y artefactos, en seguida se aplica STFT para la representación en el tiempo-frecuencia. La extracción de características es por medio de coeficientes cepstrales y un análisis de relevancia estocástico. La clasificación es por medio de una SVM con un kernel de base radial, obteniendo una exactitud del 95.78 % al 100 % para todos los subconjuntos de datos.

Una aplicación para detectar ataques epilépticos por medio de la STFT discreta es realizada por [26], las características extraídas son los coeficientes de Malmquist-Takenaka además de valores estadísticos, el clasificador utilizado es un árbol de decisión alternativo, obteniendo una exactitud de 96.7 %, 98.36 % y 99.7 % para diferentes subconjuntos de la base de datos usada.

En [57] la señal EEG se descompone en múltiples funciones de modo intrínseco (IMFs) que es una descomposición de modo empírico multimodal (MEMD), a los modos más significantes la STFT es aplicada; como característica se usa un pico y la entropía

de la magnitud. Se hace uso de un test Kruskal- Walls, gráficos, histogramas para justificar sus características. Finalmente, se obtiene una exactitud del 90 % con un clasificador kNN.

Un trabajo de [15] para la clasificación de señales de estados con alcohol y control, aplica la STFT a señales EEG; los espectrogramas obtenidos se convierten a escala de grises para extraer características. Técnicas de histogramas de co-ocurrencia de gradientes orientados (CoHOG) y eigenvalores de Matrices Hessianas de CoHOG son aplicados, las características de cada método y una combinación de éstas son utilizadas en un clasificador de mínimos cuadrados no negativo (NNLS), obteniendo exactitudes del 91.67 % al 95.83 % con una combinación de características de ambos métodos para el mejor resultado.

El preprocesamiento de las señales EEG por lo general se aplica en términos de la frecuencia, con la finalidad de eliminar ruido de fondo y mejorar la señal. Un método propuesto por [58] para resolver esta problemática se aplica directamente a una señal transformada al dominio del tiempo-frecuencia, basados en la descomposición a subimágenes por medio de descomposición de valor singular (SVD). El ruido es eliminado aplicando una organización de las imágenes y aplicando la inversa de STFT. Se aplica la permutación del valor de entropía para modificar el ordenamiento y se repiten los pasos anteriores. Los autores concluyen que es un buen método para la detección de potenciales evocados visuales de estado estable y supresión de ruido usando un solo canal.

Anteriormente, se mostraron metodologías en donde se hace uso de técnicas para la extracción de características sobre espectrogramas de señales EEG, mostrando resultados aceptables para cada una de sus aplicaciones. A pesar de mostrar un buen desempeño, se han explorado otro tipo de enfoques que de igual forma hacen uso de los espectrogramas, son los basados en aprendizaje profundo; comúnmente, con el uso de redes convolucionales, las cuales intentan imitar el cortex visual del ojo humano para identificar distintas características y poder identificar patrones. Biswal et al., en el 2017 [59], presentan una metodología para identificar estados del sueño, en la que a partir de espectrogramas de señales EEG, se obtienen características que son utilizadas para el entrenamiento de distintos modelos de clasificación como son regresión logística, árboles de decisión, perceptrón multicapa, redes neuronales convolucionales y recurrentes, seleccionando el mejor a partir de este último, y obteniendo en promedio de exactitud el 85.76 % para la clasificación de cinco clases. En 2018, Yuan et al., [60], proponen una metodología basada en aprendizaje profundo para la identificación de epilepsia; aplicando la STFT a múltiples canales EEG generando los espectrogramas

correspondientes, a cada uno de ellos se les extraen características, las cuales son agrupadas en un canal codificador (channel-encoder), generando un vector global se evalúa con una función score; logrando identificar epilepsia con una exactitud de entre el 96 % y 98 %. Un enfoque propuesto por Ruffini et al., en 2019, es descrito en [61], con la finalidad de reconocer el comportamiento de desordenes en movimientos oculares en el sueño. Su metodología consiste en el entrenamiento de una red neuronal convolucional con cinco capas, a través de espectrogramas de señales EEG, alcanzando exactitudes de entre 73 % y 80 % de la clasificación de dos clases. Otra metodología muy parecida basada en una arquitectura de red neuronal convolucional es propuesta por Vrbancic y Podgorelec [62] para la identificación automática de trastornos neurales de deterioro motor, en donde se clasifican dos clases (personas con y sin desordenes), obteniendo resultados de exactitud del 69.23 %. Siguiendo la misma metodología, Yuan y Cao [63], detectan pacientes en estado de coma y muerte cerebral, obteniendo exactitudes de entre 88 % y 96 %.

3.3. Discusión

Los artículos analizados en esta sección muestran diferentes metodologías a seguir; la primera fase del análisis consiste en remover artefactos y ruido, la mayoría de los experimentos no hacen uso de algún método para esta fase, ya que algunos conjuntos de datos ya están preprocesados. Sin embargo, para los conjuntos de datos que contienen señales no deseadas, el método que predomina son filtros, ya que al analizar las señales y las distintas bandas de frecuencia, se tienen parámetros generales del comportamiento de las señales, lo que facilita el diseño de un filtro que elimine las señales que no pertenecen a distintas señales EEG.

La tabla 3.2 muestra un resumen de las metodologías de los trabajos analizados con respecto a representaciones tiempo-frecuencia, en donde la primera columna, representa los autores y el año de publicación; la segunda, muestra los objetivos de los artículos; la tercera, la información encontrada de los datos usados en cada trabajo; la cuarta, el número de clases empleado; la quinta, el tipo de preprocesamiento si es que hubo; finalmente, se muestra la clasificación alcanzada en porcentaje de exactitud y el tipo de clasificador empleado. En ésta se nota que el número de clases que predominan para el análisis sólo son dos, empleando un máximo de tres clases. Los datos son de distintos enfoques y con distinta aplicación, a pesar de esto, el preprocesamiento es aplicado en pocos trabajos, predominando los filtros como método principal. Para la extracción de características, se extrajeron valores estadísticos en escala de grises

de la representación en tiempo-frecuencia; se observó que no se emplean métodos de selección de características para la mayoría de enfoques, la minoría hace uso de análisis de componentes principales. Por último, la clasificación fue desempeñada en general por máquinas de vectores de soporte, seguidos de vecinos cercanos y redes neuronales.

Tabla 3.2 Trabajos que utilizan representaciones en el dominio tiempo-frecuencia.

Referencias	Objetivo	Datos	Clases	Pre procesamiento	Extracción de características	Clasificación
Deng, et al., 2010 [55]	Clasificar ritmos de sílabas	6 sujetos, 128 canales a 1024Hz.	2	-	Hilbert, binarización de regiones de interés	68 % - 73.63 % con clasificador Bayesiano con LDA.
Mustafa et al., 2010 [1]	Clasificar imágenes	Conjunto propuesto	3	-	STFT: estadísticos y PCA	70.83 % con kNN
Mustafa et al., 2012 [50]	Clasificar balanceo cerebral	Conjunto propuesto	3	-	STFT: estadísticas y PCA	90 % con kNN (distancia Euclidiana) y 87.5 % con ANN
Duque, et al., 2014 [25]	Detección de ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn	2	Filtro pasa bajas	STFT: análisis aleatorio y coeficientes cepstrales	95.78 %-100 % con SVM
Fu, et al., 2014 [54]	Detección de ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn	2	-	HHT: escala de grises, estadísticos e histograma	99.12 % con SVM
Kovács, et al., 2014 [26]	Detección de ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn	2	-	STFT: coeficientes de Malmquist-Takenaka y estadísticos	96.7 %, 98.36 % y 99.7 % con árbol de decisión alternativo
Khairul, et al., 2015 [57]	Detección de movimientos imaginarios	1 persona, 2 canales, 128 Hz.	2	Filtro pasa banda	STFT: umbrales	90 % con kNN
Kumar y Sharma, 2015 [64]	Detección de ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn	2	-	STFT: suma de energías y PCA.	99.5 %, 99.33 % y 92.36 % con ANN
Nieves y Manian, 2016 [52]	Autenticación biométrica	32 canales, C3,CZ,C4. Movimientos imaginarios	2	-	STFT: energía por frecuencias, estadísticas y suma de distancias	98 % con ANN y 6 canales. 90 % y 96 % con SVM (2~y 3 canales)
Alcin, et al., 2016 [65]	Detección multiclase: estados de sueño y ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn Adquisición propia, 100 Hz, 30s, 8 sujetos.	2	-	STFT: GLMC y vectores de Fisher	95.17 % y 95.38 % para sueño y 96.40 % para epilepsia con máquina de aprendizaje extremo
Camacho, et al., 2016 [53]	Detección de imágenes motoras en tiempo real	1 canal, 400Hz, 20 sujetos.	2	Filtro pasa banda	STFT: GLMC	87.6 % con clasificador de regresión lineal
Jalilifard, et al., 2016 [24]	Detección de estados somnolientos	10 sujetos	2	-	STFT: potencia y estadísticos sobre el tiempo	88.57 %-91 % con kNN-Kd Trees
Khan, et al., 2016 [14]	Detección de ataques epilépticos	5 neonatos, 32 Hz.	2	-	Distribuciones TF: estadísticos	97.5 % con SVM
Bajaj, et al., 2016 [15]	Clasificación de señales de estados alcohólicos y de control	120 archivos, 64 electrodos, a 256Hz.	2	-	STFT: escala de grises, CoHOG y Eig(Hess)-CoHOG y estadísticos	91.67 %, 91.67 % y 95.83 % con NNLS
Sengur, et al., 2016 [51]	Detección de ataques epilépticos	100 señales, 5 subconjuntos, 173.61Hz, Universidad de Bonn	2	-	STFT: escala de grises, TFCM, LBP y GLCM.	100 % con LBP y TFCM y 87 %-92 % con GLCM, con SVM

De los diferentes artículos que hacen uso de espectrogramas y otras distribuciones, se notó que, para el análisis de enfermedades de epilepsia, existe un conjunto de datos que diferentes autores [16, 57] hacen referencia, es un conjunto de datos realizado por la Universidad de Bonn un conjunto de datos con cinco diferentes subconjunto A-E, con 100 señales de un solo canal, con una duración de 26.3 s y una frecuencia de muestreo de 173.61 Hz. Esta base de datos contiene información de alta calidad (número de estudios, tiempo y número de ataques).

Independientemente de la distribución utilizada para representar las señales EEG, diferentes trabajos usan la representación del mapeo en la matriz de co-ocurrencia en escala de grises para extraer características. Las características comúnmente están basadas en los valores estadísticos que presenta dicha matriz, no siguiendo un estándar ya que cada autor propone y justifica las características a usar. Las características usadas se evalúan con distintos clasificadores, de los que predominan las máquinas de vectores de soporte y algoritmo de k vecinos cercanos, seguidos de redes neuronales. Los resultados mostrados reflejan que el mejor desempeño en cuanto a exactitud es para SVM y ANN.

Los trabajos se enfocan a aplicaciones médicas, relacionadas al sueño, a aspectos de autenticación, a la detección de imágenes motoras imaginarias, entre otras. Las aplicaciones médicas están dirigidas a la detección de anomalías dentro de las señales EEG como son: trastornos neurológicos provocados por ataques epilépticos, detección de los efectos del alcohol en el cerebro. En el análisis del sueño, se hacen experimentos para la detección de los estados del sueño para identificar problemas de somnolencias. Por otra parte, aplicaciones dirigidas a aspectos de autenticación, muestran resultados que pueden ser competitivos en comparación de otros biométricos, mostrando que la señal EEG tiene propiedades suficientes que lo hacen ser seguro, único, no visible y no modificable. Entre las aplicaciones mencionadas y otras como son: clasificación de imágenes motoras, detección de palabras imaginadas, los trabajos no muestran tiempo de ejecución y es debido a la aplicación ya que no se requiere de ejecución en tiempo real. El tiempo de ejecución también está relacionado con las distribuciones utilizadas, se concluye que el uso de la STFT para la generación de mapeos en tiempo-frecuencia es rápida (con respecto a otras distribuciones como Wavelet o Hilbert), debido a su complejidad computacional $O(n \log n)$ donde n es el número de muestras [66], por otro lado, con respecto a la resolución se tienen mejores resultados para otras representaciones [8]. Para los objetivos en la literatura, la resolución obtenida con STFT es suficiente para cumplirlos, eso implica menor trabajo computacional y menor tiempo de ejecución. Sin embargo, se mostró en trabajos

como [1, 50, 15] que la dimensionalidad es una problemática a resolver, ya que los autores no mencionan el tiempo de ejecución a pesar de aplicar análisis de componentes principales para la selección de características; el último autor no aplica reducción de características, aun así menciona que hace falta reducir los vectores, ya que se tienen miles de valores.

Tabla 3.3 Características extraídas de las representaciones en tiempo-frecuencia.

Trabajo	Extracción de características	Características
Deng et al. 2010 [55]	Hilbert: binarización, normalización y regiones de interés usando (SOI) el algoritmo de identificación ciega de segundo orden modificado (SOBI)	20 características: número de SOI's; distancia media de tiempo entre cada dos SOI's vecinos; duración media de los SOI; frecuencia central media de SOI's y frecuencia media de SOI. Se usaron cinco características de cuatro componentes SOBI.
Mustafa et al. 2010 [1]	STFT: GLMC, características de textura y PCA	Características de textura propuestas por Haralick y Soh; auto correlación, contraste, correlación, progresión de clúster, tonalidad de clúster, disimilaridad, energía, entropía, homogeneidad, probabilidad máxima, varianza, promedio de sumas, varianza de suma, entropía de suma, varianza de resta, entropía de restas, información de correlación 1, información de correlación 2, diferencia inversa normalizada y momento de diferencia inversa normalizada.
Mustafa et al. 2012 [50]	STFT: GLMC, Estadísticas y matrices de huge, PCA	Características de textura, combinación de técnicas de Haralick, Coh y Clausi: autocorrelación, contraste, correlación, progresión de clúster, tonalidad de clúster, disimilaridad, energía, entropía, homogeneidad, probabilidad máxima, varianza, promedio de sumas, varianza de suma, entropía de suma, varianza de resta, entropía de restas, información de correlación 1, información de correlación 2, diferencia inversa normalizada y momento de diferencia inversa normalizada.
Duque, et al. 2014	STFT	Cinco coeficientes cepstrales en frecuencia asociados con los ritmos alfa, beta theta y delta.
Fu et al. 2014 [54]	HHT: Imagen en escala de grises: Estadísticos, histograma.	Características extraídas de los histogramas de intensidad de píxeles a partir de segmentos del mapeo en escala de grises. las características: medio, varianza, sesgo y curtosis.
Kovács et al, 2014 [26]	HTFT: Coeficientes de Malmquist-Takenaka, estadísticos.	Características de las STFT discreta, a partir de coeficientes racionales y cinco valores estadísticos absolutos de los coeficientes: medio, mediano, máximo, mínimo, desviación estándar.
Khairul et al, 2015 [57]	STFT: Umbrales.	Pico y entropía de la magnitud espectral.
Kumar y Sharma, 2015 [64]	STFT: Suma de energías, PCA.	Energía fraccional para segmentos EEG.
Sengür et al, 2015 [51]	STFT: GLCM, método de codificación de la característica de textura (TFCM) y patrón binario local (LBP).	GLCM: contraste, correlación, energía y homogeneidad.
Nieves y Manian, 2016 [52]	STFT: Energía por frecuencias y estadísticas. Suma de distancias.	Media, varianza y sesgo de la distribución de energía
Alcín et al. 2016 [65]	STFT: GLMC y VF.	Descriptores de textura: Contrastes, correlación, energía y homogeneidad.
Camacho et al. 2016 [53]	STFT: GLMC.	Descriptores de textura: correlación, energía, contraste, homogeneidad y disimilitud.
Jalilifard et al., 2016 [24]	STFT: potencia y características estadísticas sobre el tiempo. Random Forest para la reducción.	Media del valor de la potencia del espectrograma y características en el dominio del tiempo (Desviación estándar y entropía de Shanon).
Khan et al, 2016 [14]	Distribuciones TF: Estadísticas.	Frecuencia instantánea (IF) y Amplitud instantánea (AI) se usan como componentes para extraer características: Media IF, media IA, varianza IF, varianza IA, curtosis IF y curtosis IA.
Bajaj et al, 2016 [15]	STFT: Escala de grises.	Características de textura local con: co-ocurrencia de histogramas de gradientes orientados (CoHOG) y Eig(Hess).

Los trabajos analizados extraen diferentes características de espectrogramas de señales EEG, a pesar de que algunas son similares el cambio de alguna de estas muestra resultados significativos en términos de la exactitud (de 1 % a 3 %). La tabla 3.3 presenta

las características que usan los autores, la primera columna (trabajo) corresponde a los autores, la segunda (extracción de características) consiste de los métodos utilizados y la última columna (características) son algunas de las que se extrajeron en los trabajos mostrados. Estos trabajos en general extraen características a partir de la textura o de tipo estadísticas de las representaciones de tiempo-frecuencia, otros trabajos se basan en la energía y unos pocos utilizan diferentes métodos como LBP, TFCM, CoHOG, coeficientes cepstrales, entre otros. Por otra parte, se mostraron enfoques que no hacen uso de la etapa de extracción de características como son los basados en aprendizaje profundo, sin embargo, el uso de estos algoritmos implica el entrenamiento con al menos miles de ejemplos y el uso de equipos dedicados con tecnología específica.

Capítulo 4

Extracción de características de espectrogramas de señales EEG

El análisis de señales EEG es un proceso que involucra diferentes etapas, como se mencionó en los capítulos previos: inicialmente, es preprocesamiento es aplicado, en el cual se filtran señales adheridas a las señales que se analizarán; posteriormente, se extraen las características que puedan ser relevantes y denoten patrones; finalmente, se evalúan estas señales por medio de clasificadores.

La extracción de características es un proceso en el que se obtiene información o características relevantes de las señales EEG, con la finalidad de interpretarlas de una forma más rápida o fácil respecto al uso de una señal completa [67]. Se dice que una característica es una medida distintiva o componente estructural extraído de un patrón [68]; por lo tanto, un esquema de extracción de características significa elegir las características más útiles o información, la cual es la más importante para un análisis de señales EEG, ya que la clasificación es fuertemente dependiente de las características obtenidas [69, 70].

La metodología desarrollada en esta investigación consiste en el análisis de espectrogramas de señales EEG a través de diferentes enfoques de extracción de características, con diferentes objetivos, es decir enfoques genéricos, por lo que a partir de ciertas características relevantes que están contenidas en las señales EEG como son frecuencia y potencia se desarrollaron las propuestas. Thatcher et al., realizaron estudios sobre la potencia de las señales EEG en pruebas de inteligencia y anormalidades y Sponheim et al., en esquizofrenia, encontrando que existe relación entre la frecuencia y la potencia que presentan las señales EEG [71, 72]. Por otro lado, Tiernet et al., realizó un estudio sobre autismo en donde describe diferentes niveles de potencia en varios rangos de frecuencias (bandas) de las señales EEG, identificando que la potencia varía por edad

[73]. Park presenta un estudio para la predicción de ataques epilépticos basado en la potencia espectral de las señales EEG y por medio de máquinas de vectores de soporte [74]; lo que denota la importancia de la potencia. Sin embargo, una característica importante de la potencia espectral es el pico, debido a que muestra de una forma más clara la frecuencia de las señales EEG en un análisis espectral de frecuencia o tiempo-frecuencia [75, 76]; por lo que, se seleccionaron características fundamentales de las señales EEG que son las bandas, potencia y los picos para la propuesta de los diferentes enfoques.

Recordemos que la STFT es una alternativa con resultados competentes en el análisis de las señales EEG como se describió en el capítulo 3 y esto se debe a la ventaja que la transformada muestra: la representación espectral de las señales a lo largo del tiempo. Por consiguiente, las metodologías o enfoques propuestos parten de los datos que proporciona la STFT y las características de las señales EEG; es decir, información de los picos espectrales en determinadas frecuencias a lo largo del tiempo.

Un estudio realizado por Allen y MacKinnon muestra la relevancia de los picos en diferentes representaciones tiempo-frecuencia incluyendo la STFT para señales EEG durante movimientos repetitivos, sin embargo, los picos, específicamente en los espectrogramas dependen del ajuste de los parámetros de la STFT [77]. En consecuencia, durante esta investigación, se propusieron diferentes enfoques con la finalidad de obtener características relevantes de picos espectrales en espectrogramas, además de un método para el ajuste de los parámetros más importantes de la STFT como es la longitud de la ventana. La figura 4.1 muestra las etapas de las metodologías propuestas: partiendo de una base de datos de señales EEG, la STFT se aplica a éstas, obteniendo como respuesta espectrogramas por cada señal; posteriormente, se aplican los enfoques desarrollados durante esta investigación, los cuales son descritos en la sección 4.3.

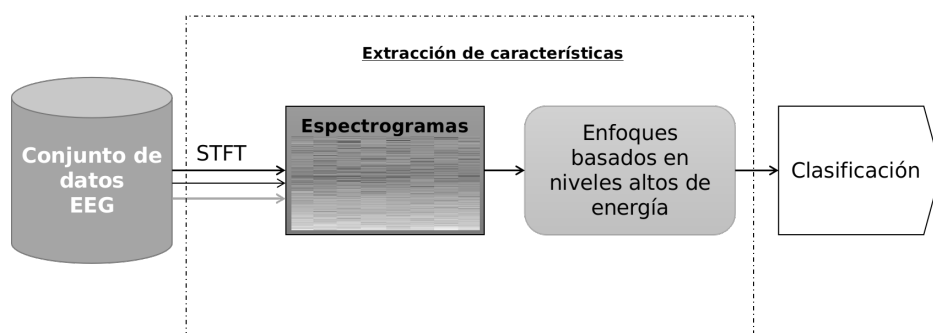


Figura 4.1 Metodología de análisis de señales EEG.

4.1. Señales EEG

En la sección 2.2.1, se mencionó que las señales EEG son registros de la actividad que se genera en las neuronas, las cuales vienen descritas principalmente por amplitudes de voltaje a lo largo del tiempo. Sin embargo, otra característica importante de estas señales es la frecuencia, que se refiere a la actividad repetitiva rítmica descrita en Hertz (Hz), estos potenciales de EEG se manifiestan como oscilaciones impredecibles no periódicas; dependiendo de la edad de las personas, las frecuencias y amplitudes de voltajes pueden cambiar [22].

La figura 4.2, muestra una grabación común de señales EEG, dependiendo del número de canales que sean usados para el registro se tiene un número diferente de señales y de datos. Como se mencionó en la sección 2.2.1, existe un sistema estándar para el número de canales a usar. Sin embargo, dependiendo del tipo de dispositivo se tendrá una mejor resolución a lo largo del tiempo, esto lo proporciona la frecuencia de muestreo de cada dispositivo, comúnmente, se hacen registros cada segundo de 128 valores hasta 1000 [78]. Mientras el número de muestras se más grande, el detalle de las señales es mas claro, sin embargo, la dimensión de los datos con los que se trabajaría también sería mayor. El manejo de estos datos son administrados por distintos formatos de archivos, uno de los más comunes es el EDF (European Data Format), el cuál almacena como metadatos, información del paciente y organiza todos los canales, la desventaja de este archivo, es cuando los datos provienen de dispositivos con tasas de muestreo altas, el manejo se hace complejo, una solución a esto, es la grabación en formato MEF (Multiscale Electrophysiology Format) [79], a pesar de la existencia de estos estándares, se encontraron bases de datos con formatos CSV y archivos de texto. Los archivos de texto, generalmente, son generados por canal; los CSV y MEF es un solo archivo pero internamente tiene todos los canales, así que estos tienen que ser separados y analizados por canal.

4.2. Transformada de Fourier de Tiempo Corto

Como se menciona en la sección 3, Tzallaz [49] presentó una distribución de frecuencia de tiempo frecuentemente utilizada para analizar los EEG de pacientes que experimentan una crisis epiléptica; mostrando a la STFT como una alternativa para extraer características de los espectrogramas. Adicionalmente, la STFT también se ha utilizado para analizar electrocardiogramas (ECG), que mapean la señal en una función bidimensional de frecuencia y tiempo [1]. Para recuperar información de las

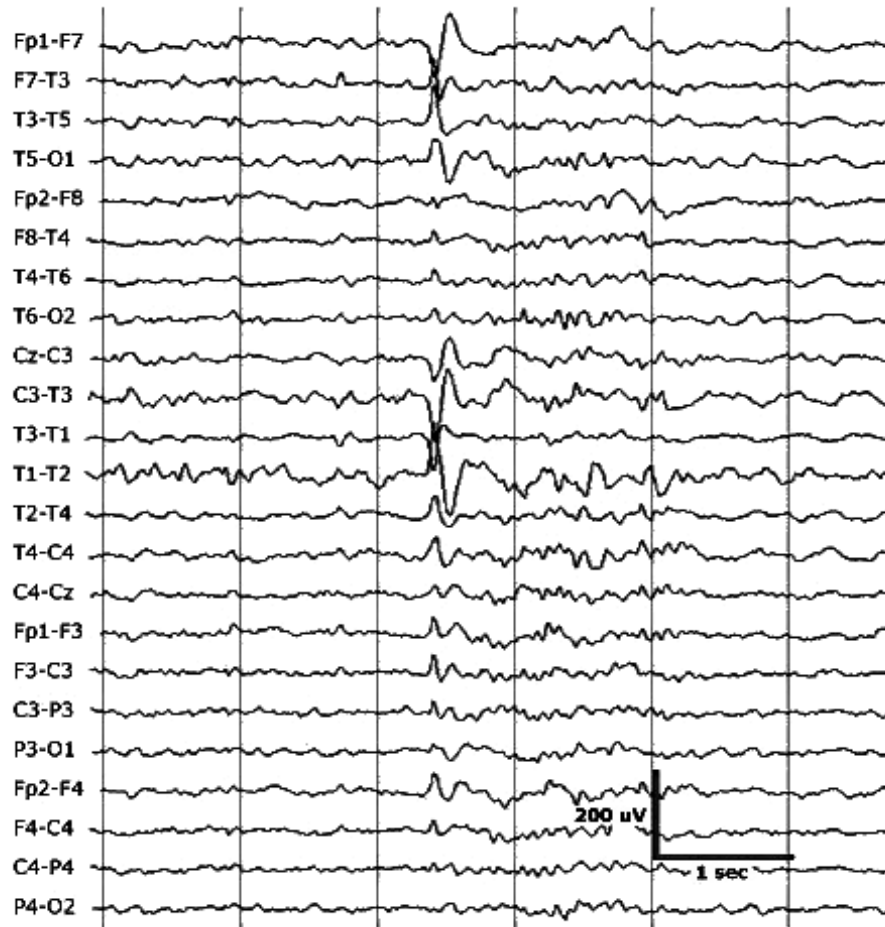


Figura 4.2 Ilustración de una grabación común de EEG.

EEG, la STFT se aplica a la señal, dividiéndola en distintos segmentos de tiempo y por traslapes [56], el proceso se conoce como ventaneo. Dependiendo de la función de ventana de tiempo w_n , el espectrograma resultante se clasifica como de banda estrecha o banda ancha. Cuando las ventanas de tiempo son cortas, la STFT será de banda ancha; para un tiempo más largo dará como resultado un espectrograma de banda estrecha [80]. Los segmentos (frames) de la señal se pueden expresar como [81]:

$$x_l[n] = w_n * x[n + lL], \quad 0 \leq n \leq N - 1, \quad (4.1)$$

donde N es el tamaño de la ventana, l es el índice del frame y L es el tamaño del desplazamiento, n es un índice del instante que corresponde al comienzo del desplazamiento de la ventana. La expresión $x[n + lL]$ representa una posición sobre la señal, n es la fase de desplazamiento de la ventana, $*$ representa una modulación entre

dos términos. Finalmente, la Transformada discreta de Fourier (DFT) [81] se aplica a cada frame de la señal, de la siguiente manera:

$$X[k, l] = \sum_{n=0}^{N-1} x_l[n] e^{-i2\pi nk/K} = \sum_{n=0}^{N-1} w_n * x[n + lL] e^{-i2\pi nk/K}, \quad (4.2)$$

donde K es la longitud del DFT y k es un índice de frecuencia o índice bin. La STFT $X[k, l]$ representa el comportamiento de un segmento de la señal a lo largo del tiempo lL y el número de bins k para una frecuencia de muestreo de F_s . Estos índices discretos corresponden a la longitud en tiempo de la muestra, que se denota por lL/F_s y la frecuencia está dada por kF_s/K [81]. Para simplificar la notación, una frecuencia radial de

$$\omega_k = 2\pi k/K \quad (4.3)$$

así, la expresión STFT se convierte en:

$$X[k, l] = \sum_{n=0}^{N-1} w_n * x[n + lL] e^{-i\omega_k n}. \quad (4.4)$$

Aplicando la ventana a lo largo del tiempo, la STFT resulta en una representación espectral de un segmento de tiempo de la señal de entrada; mostrando a $X[k, l]$ como una función de la frecuencia k para cada valor del índice de tiempo l , la STFT corresponde a una serie de espectros ubicados en el tiempo. Alternativamente, la STFT puede verse como una función de tiempo para cada frecuencia; representando a $X[k, l]$ como una serie de tiempo que es una función de l para cada bin k (puntos para calcular el DFT), la STFT corresponde a un banco de filtros que descompone la señal de entrada sobre canales de frecuencia o bandas [81]. Estas dos interpretaciones del STFT se representan con respecto al plano de frecuencia de tiempo. Los cambios en el tiempo y la frecuencia se han vuelto comunes en la literatura moderna sobre el procesamiento de señales como una forma de representar el tiempo de la señal y las resoluciones de frecuencia.

La STFT presenta un inconveniente relacionado con la resolución de los espectrogramas, provocado principalmente por el tamaño de la ventana. Una ventana estrecha ofrece una mejor resolución en el tiempo, pero no en frecuencia, mientras que una ventana ancha proporciona una buena resolución en frecuencia, aunque no tiene el mismo desempeño para la resolución en el tiempo. Por lo tanto, es difícil lograr imparcialidad en los términos de tiempo y frecuencia simultáneamente, ya que la STFT depende de un solo tipo de ventana que divide a la señal (estacionaria); en la siguiente sección se muestra una forma de calcular el tamaño de la ventana.

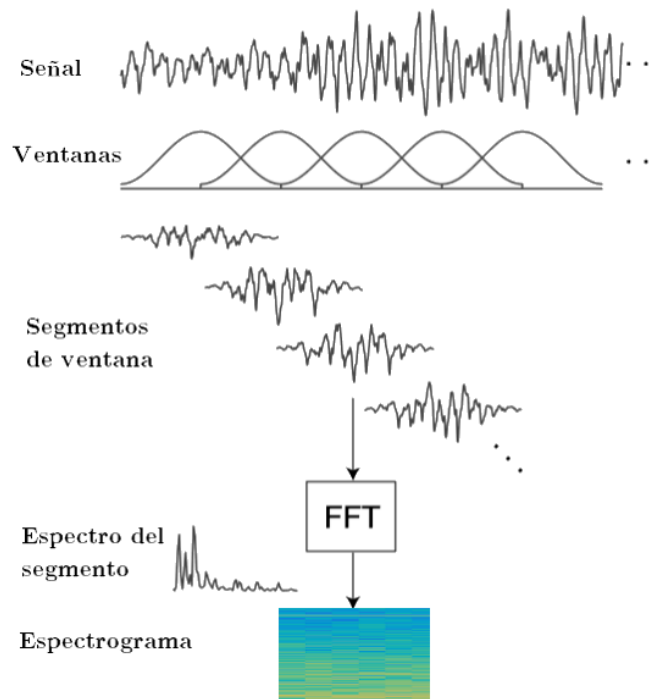


Figura 4.3 Proceso de la STFT al aplicarla a una señal, manipularla espectralmente y aplicar la transformada inversa. Figura tomada de [82].

El proceso de la STFT comienza con la aplicación de la ventana a una señal en segmentos más cortos, es decir, cada ventana se aplica a las partes de la señal, donde las ventanas pueden superponerse o traslaparse; posteriormente, la transformada rápida de Fourier (FFT) se aplica a cada segmento por separado [82] y cada resultado se une para formar el espectrograma. En esta etapa la resultante puede manipularse y luego se aplica la FFT inversa para devolver cada segmento al dominio del tiempo como se muestra en la figura 4.3. En consecuencia, para aplicar la STFT a las señales, se deben considerar las características de la transformación; tales como: tipo y longitud de la ventana, longitud del lóbulo principal (correspondiente a las ventanas del espectro), puntos FFT y superposición. Estos parámetros afectan la resolución del espectrograma y las etapas siguientes de los enfoques de extracción de características.

4.3. Enfoques de extracción de características

4.3.1. Energía y Entropía de Filtros Gabor (EEFGabor)

Inicialmente, la base de datos de epilepsia de la Universidad de Bonn fue utilizada para el análisis de las señales EEG. Este conjunto de datos cuenta con señales de

personas sufriendo ataques epilépticos durante los registros; este tipo de desorden se presenta de forma recurrente [83], por lo que la representación de las EEG por medio de espectrogramas muestra formas (representadas por colores) aparentemente repetitivas [84]. Análogamente, la textura se refiere a la caracterización de regiones en una imagen, representadas por la repetición de estructuras básicas en áreas de una imagen. Por lo que, a partir de esto se aplicaron filtros Gabor a señales EEG, con la finalidad de encontrar patrones parecidos a la textura dentro de los espectrogramas (ver figura 4.4). La respuesta $R(x,y)$ de un filtro de Gabor $G(x,y)$ es la convolución del filtro de

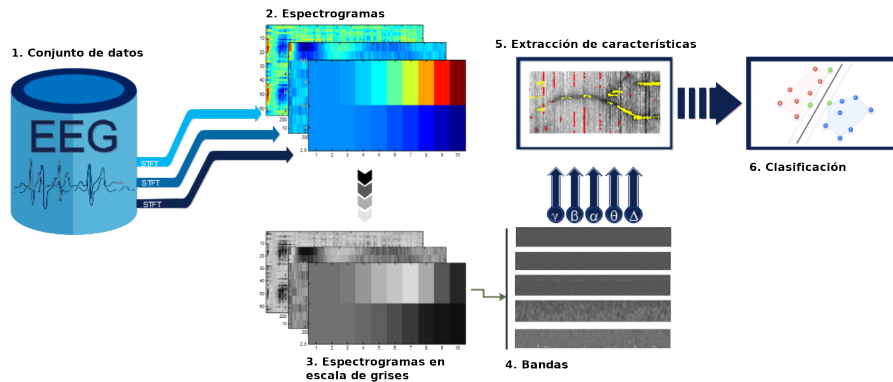


Figura 4.4 Enfoque basado en filtros Gabor.

Gabor con un espectrograma de EEG $A(x,y)$. Por lo tanto, la respuesta del filtro de Gabor se obtiene como:

$$R(x,y) = A(x,y) \otimes G(x,y), \quad (4.5)$$

Donde \otimes es el operador de convolución.

Usando un banco de filtros con diferentes parámetros, es posible obtener características de imágenes de espectrograma con diferentes escalas y orientaciones, y usarlas para recuperar características locales y discriminatorias [85]. Una característica de textura común es la entropía, que es una medida de la cantidad de información (incertidumbre) en una fuente; en las imágenes, mide la aleatoriedad de la distribución de los valores del coeficiente sobre los niveles de intensidad [86]. En este caso, la fuente es una imagen (espectrograma), vista como una matriz de intensidades (energía) a la que se le aplica la ecuación 2.4 para encontrar la entropía de Shannon.

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i), \quad (4.6)$$

En la ecuación anterior, $P_r[X = x_i] = p(x_i)$ es la distribución máxima de probabilidad de la fuente y n es el número de probabilidades $p(x_i)$, La energía es otra característica de textura de uso frecuente [87], que mide la uniformidad de la distribución del nivel de intensidad [86, 88] y propone la energía en un análisis utilizando la textura del filtro de Gabor. La energía se calcula a partir de la siguiente ecuación donde M y N son el tamaño de la subimagen $a(M, N)$:

$$e(x) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |a(i, j)|, \quad (4.7)$$

4.3.2. Características obtenidas de picos espectrales (FOSP)

El enfoque EEFGabor descrito en la subsección anterior (4.3.1), se basa en la identificación de patrones formados por la potencia de las señales EEG en los espectrogramas, haciendo referencia principalmente a la dispersión de la energía. Sin embargo, como se mencionó al principio de este capítulo, las características importantes de la potencia son sus picos y su banda de frecuencia; en consecuencia, se propone FOSP que es un enfoque que extrae características a partir de picos espectrales, que ocurren en las bandas de EEG para ciertos instantes de tiempo. Se aplica una función de máximo para encontrar picos espectrales, los cuales son usados para generar superficies y obtener volúmenes, debido a que estos denotan la energía o potencia a partir de los picos (puntos más relevantes con respecto a su nivel energía), se aplica LBP y LTP con la finalidad de abstraer un sólo elemento a partir de tiempo, frecuencia y energía. Alternativamente, se aplica K-means, con la finalidad de generar cierto número de cluster de picos con características en común como energía, frecuencia y tiempo, y así seleccionar un elemento por cluster el cual representa la característica. En general, cada una de las técnicas utilizadas para la extracción de características fueron seleccionadas por su adaptación con las características (simultáneas) que nos proporcionan los picos espectrales (tiempo, frecuencia y energía).

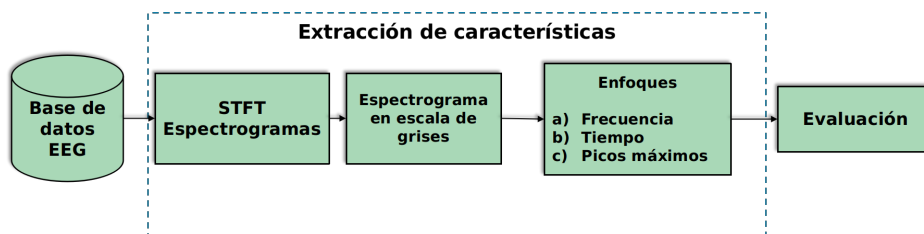


Figura 4.5 Enfoque basado en picos máximos de energía.

La metodología comienza con la identificación de los picos espectrales dentro de la magnitud del espectrograma, a diferencia del enfoque descrito en la sección 4.3.1, los espectrogramas no se convierten en imágenes; posteriormente, las características se extraen de los picos, aplicando las técnicas anteriormente mencionadas y descritas las subsecciones siguientes; finalmente, la evaluación se realiza por medio de tres diferentes clasificadores: MLP, SVM y kNN.

4.3.2.1. Picos espectrales

La ecuación 4.4 muestra la respuesta dada por la STFT o el espectrograma, que puede verse como una matriz X con k filas y l columnas, que representan las frecuencias y el instante de tiempo respectivamente. Por lo tanto, los picos espectrales se calcularon de la siguiente manera:

$$E[l] = \max[X(l)] \quad (4.8)$$

donde E es la energía máxima en el índice de tiempo l sobre todas las frecuencias. Como se muestra en la Fig. 4.6: el eje y representa la frecuencia de la señal en el tiempo (eje x); por lo tanto, si el espectrograma se representa como una matriz, cada columna es la transformada de Fourier de un segmento de la señal EEG. En consecuencia, el espectrograma se analiza por columnas, encontrando el máximo de cada una, representando el pico con la energía más alta. La cantidad de picos depende del ventaneo, es decir la misma cantidad de picos por columna es la misma con respecto al número de ventanas; finalmente, cuando todos los picos se hayan calculado, se obtiene uno: mínimo, medio o máximo.

4.3.2.2. FOSP con superficies, frecuencia y patrón binario local

Este método utiliza diferentes características obtenidas del espectrograma como el volumen, frecuencia y la codificación de LBP. Primero, el volumen se basa en un enfoque que considera las intensidades del espectrograma en una representación tridimensional. Todos los puntos de intensidad y las coordenadas de tiempo y frecuencia se usaron para obtener el volumen del espectrograma; así mismo, todas las intensidades se combinaron para obtener la energía aproximada. Los picos espectrales se usaron para calcular el área de superficie y se agregaron las intensidades de los picos más altos; el área formada por picos y el volumen de todos los puntos del espectrograma se calcularon a través de cerco convexo.

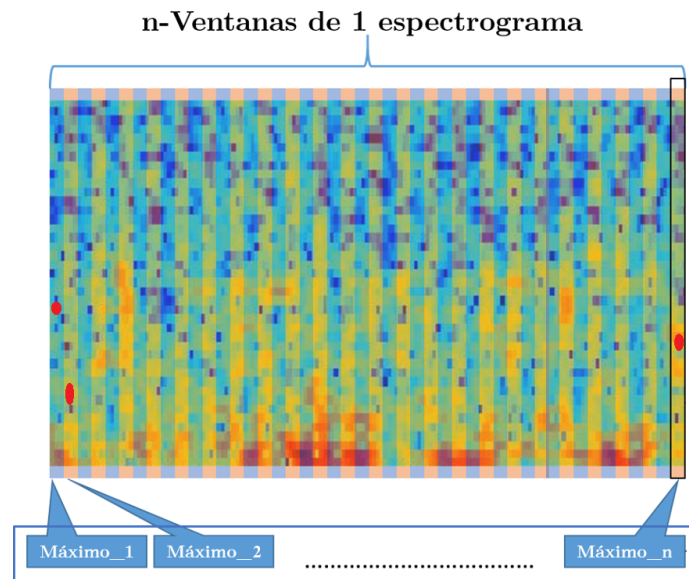


Figura 4.6 Picos extraídos de cada ventana.

En segundo lugar, las características de frecuencia se extrajeron del espectrograma de los picos máximos descritos en la subsección anterior; el número de picos depende del número de ventanas propuestas. La figura 4.6 representa esta etapa, donde el eje x es el tiempo y el eje y (columnas) correspondiente a las frecuencias. Los puntos rojos representan los picos máximos, que se utilizan para encontrar la frecuencia de ese pico en un instante determinado. Esta frecuencia es una aproximación obtenida de la frecuencia máxima de la señal EEG, la cual se encuentra en la parte superior del espectrograma. Finalmente, se utilizó una adaptación de LBP para obtener una característica de tiempo, que compara el nivel de frecuencia pico de cada ventana a lo largo del tiempo y genera un código.

4.3.2.3. FOSP con energía, K-means y patrón ternario local

La energía es una característica obtenida solo con la suma de los puntos de intensidad del espectrograma; con el algoritmo K-means, se agruparon los picos, se utilizaron experimentalmente diferentes números de K y la característica extraída fue la energía en los centroides. El número de grupos K es el número de características extraídas, que son los puntos máximos más relevantes. Las características de tiempo se toman de las ventanas y las 5 bandas conocidas del EEG [89], formando secciones o celdas en las que se ubican los picos máximos, como se muestra en la figura 4.7. Estas características se obtienen a través del método LTP que es una variante LBP, el cual consiste en extraer

características de textura utilizando umbrales o valores para la comparación. La técnica LTP descrita en [90] se usa para extraer características de textura para la identificación de humanos, en donde se consideran tres comparaciones de intensidades, las cuales fueron útiles para nuestro método, donde se adaptaron a la frecuencia, tiempo y energía. LTP se ajustó a los picos encontrados en el espectrograma, comparando la celda con un pico sobre la ventana n con la celda del siguiente pico $n + 1$ (a la derecha); si están en los mismos rangos de frecuencia, es decir, las posiciones de las celdas que contienen los picos son las mismas con respecto a las frecuencias o al eje *vertical*, entonces se genera un valor de 0; si la celda del pico en la ventana n está en una posición inferior a la del pico de la ventana $n + 1$, se genera un valor de 1; finalmente, si la posición máxima de la ventana n es mayor que la posición de la celda de ventana $n + 1$, se obtiene un valor de -1. Los valores se crean a través del desplazamiento de todos los picos a lo largo del tiempo, desde la primera ventana hasta la última. El vector de valores (-1, 0 y 1) obtenido, se divide en otros dos vectores como se menciona en la en le sección 2.11; en uno, los valores de -1s se convierten en 0 y en el otro los 1s se convierten en 0s y los -1s en 1s, ambos vectores binarios son convertidos a valores decimales y se utilizan como características como se muestra en la parte derecha de la figura 4.7.

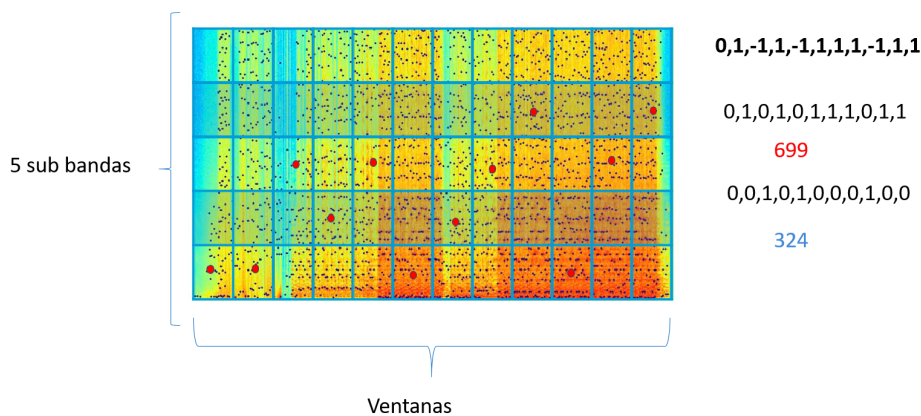


Figura 4.7 Extracción de características por medio de LTP.

4.3.3. Picos Espectrales de Vecindades (PEV)

El último enfoque desarrollado se muestra en la figura 4.8, el cual se propone con la finalidad de complementar a FOSP, debido a que los picos obtenidos con dicho enfoque son locales o encontrados por instante de tiempo y no contempla picos vecinos; a pesar de que K-means se aplica y contempla vecindades, éste depende del parámetro K ; DBSCAN es utilizado debido a que los picos son agrupados por su densidad y esta

nos indica que ocurre un evento importante por su nivel de energía, en un instante de tiempo y a una frecuencia dada. Inicialmente, a partir de las bandas comunes de la señales EEG en un espectrograma, se obtienen cinco segmentos correspondientes a *alpha*, *beta*, *gamma*, *delta* y *theta*. Cada uno de estos son preprocesados por medio de cuartiles correspondientes a los niveles de energía que contiene cada sub-espectrograma; posteriormente, mediante un rango definido del cero al cuartil tres, los puntos son descartados, quedando solo los puntos con mayor energía. A estos, se le aplica DBSCAN con la métrica *Manhattan* para agruparlos, a partir de la frecuencia, tiempo y energía, formando distintos cluster; tomando en cuenta el cluster más denso se ubica el punto con mayor energía y se selecciona como característica la energía y la frecuencia. La métrica Mahattan fue elegida dado que la distancia total es resultado de la suma de las distancias absolutas de los tres parámetros, los cuales fueron normalizados y la importancia de estos es la misma debido a la ocurrencia de algún parámetro a lo largo de tiempo, frecuencia y energía.

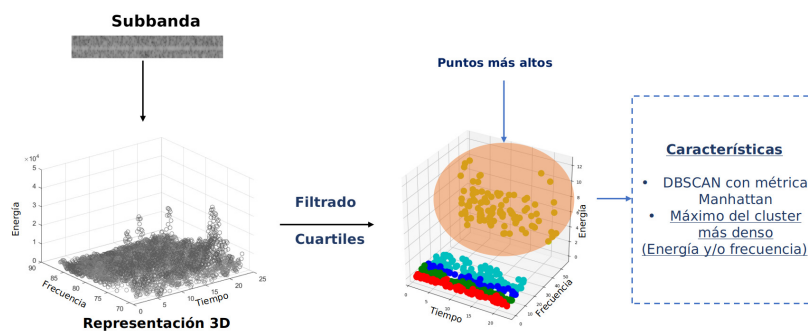


Figura 4.8 Extracción de características por medio de clusters.

4.4. Evaluación

Para la evaluación de las características se usaron diferentes algoritmos de clasificación como el perceptrón multicapa, máquinas de vectores de soporte (lineal, polinomial y RBF) y k vecinos cercanos, según lo analizado y mostrado en la sección 3.3. La metodología experimental fue evaluada de formas diferentes, debido a que los conjuntos de datos fueron distintos, así como los enfoques. Todos los resultados se presentan en términos de exactitud, sin embargo, para la metodología FOSP, la cual presenta los mejores resultados, se hicieron diversos experimentos y los resultados se reforzaron con un análisis de coeficiente Kappa y un análisis de significancia estadística. En el siguiente capítulo se detallan la experimentación y se muestran los resultados.

Capítulo 5

Experimentación y resultados

Los resultados de los enfoques EEFGabor, FOSP, y PEV detallados en la sección 4.3 se muestran en este capítulo, el cual se divide en cinco secciones: el primer, presenta los conjuntos de señales EEG con los que se trabajó; en seguida, se describe la caracterización de los parámetros de la STFT; finalmente, las últimas tres secciones muestran los resultados de los experimentos.

Para estos experimentos se consideraron parámetros de las ventanas propuestas y combinaciones de éstas, lo cual se detalla en las secciones siguientes. Como se mencionó en la sección 4.4, diferentes algoritmos de clasificación se usaron para la evaluación de las características; a partir de la literatura se mostró que se obtienen resultados aceptables para la clasificación de señales EEG [50, 1, 24, 57] superiores al 90 % de exactitud para diferentes tareas. Los resultados se muestran en términos del porcentaje de la exactitud, usando diferentes métodos de evaluación: validación cruzada en todos los experimentos de los tres enfoques. Para el enfoque FOSP se agregó la medida estadística de coeficiente kappa y una prueba de significancia estadística Friedman, debido a que este enfoque obtuvo los mejores resultados con respecto a la exactitud de clasificación (100 %). Por otra parte, para un experimento con uno de los conjuntos de datos (estados de sueño) se recurrió a aplicar una técnica de balanceo, debido a que el número de instancias es diferente por clase; esto provocó mejores resultados en comparación con el conjunto sin balancear, en donde se obtenían valores menores al 80 % con respecto a la exactitud.

5.1. Conjunto de datos

Para el análisis de las señales EEG se utilizaron tres conjuntos de datos: 1) Epilepsia-Bonn, 2) Narcolepsia-BUAP y 3) Epilepsia-Boston. El primer conjunto se nombra

así por el objetivo de este, seguido de Bonn quien es la Universidad que realizó las grabaciones; este conjunto contiene datos preprocesados, así como la selección de canales propuesta por expertos y es uno de los más trabajados en la literatura. La segunda base de datos son registros tomados de ratas, los cuales son usadas para experimentos de epilepsia y narcolepsia en los laboratorios de Fisiología de la BUAP; las señales contenidas en los registros están preprocesadas y los canales usados son cuatro que incluye uno de EMG. La última base de datos usada se seleccionó con el objetivo de observar las diferencias que existen con la primera, este conjunto capturado por el Hospital de Boston fue a través un número mayor de pacientes y canales de EEG a diferencia del primer.

5.1.1. Epilepsia - Universidad de Bonn

La metodología propuesta, se probó con un conjunto de datos EEG de acceso abierto sobre ataques epilépticos, elaborado por la Universidad de Bonn, en [91] se describe el conjunto de datos compuesto por cinco subconjuntos o series (denotadas A – E), cada serie contiene 100 segmentos de EEG de un solo canal con una duración de 23.6 segundos. Los segmentos fueron seleccionados y recortados de grabaciones continuas de EEGs multicanal después de la inspección visual de artefactos. Además, los segmentos tenían que cumplir con un criterio estacionario. Los subconjuntos A y B consisten de segmentos tomados de grabaciones de EEG que se llevaron a cabo en cinco voluntarios sanos utilizando una estandarización de las ubicaciones de los electrodos de superficie de acuerdo al sistema internacional 10-20. Los voluntarios se relajaron en estado de vigilia con los ojos abiertos (A) y los ojos cerrados (B), respectivamente. Los conjuntos C, D y E se originaron a partir de un archivo EEG de diagnóstico prequirúrgico, tomados de cinco pacientes. Los segmentos del conjunto D se registraron desde la zona epileptogénica y los del conjunto C de la formación del hipocampo del hemisferio opuesto del cerebro. Los subconjuntos C y D contiene actividad medida durante los intervalos sin convulsiones, el conjunto E contiene actividad de los ataques. En este subconjunto se seleccionaron segmentos de todos los sitios de grabación que exhiben actividad ictal. Todas las señales de EEG se registraron con el mismo sistema amplificador de 128 canales, utilizando una referencia común promedio, omitiendo electrodos que contienen actividad patológica (C, D y E) o artefactos de movimiento ocular fuerte (A y B). Después de la conversión de analógico a digital de 12 bits, los datos se obtuvieron continuamente a una frecuencia de muestreo de 173.61 Hz. La configuración del filtro pasa banda fue de 0.53–40 Hz (2 dB / oct.), la figura 5.1 muestra un segmento de 1 segundo de cada uno de los subconjuntos.

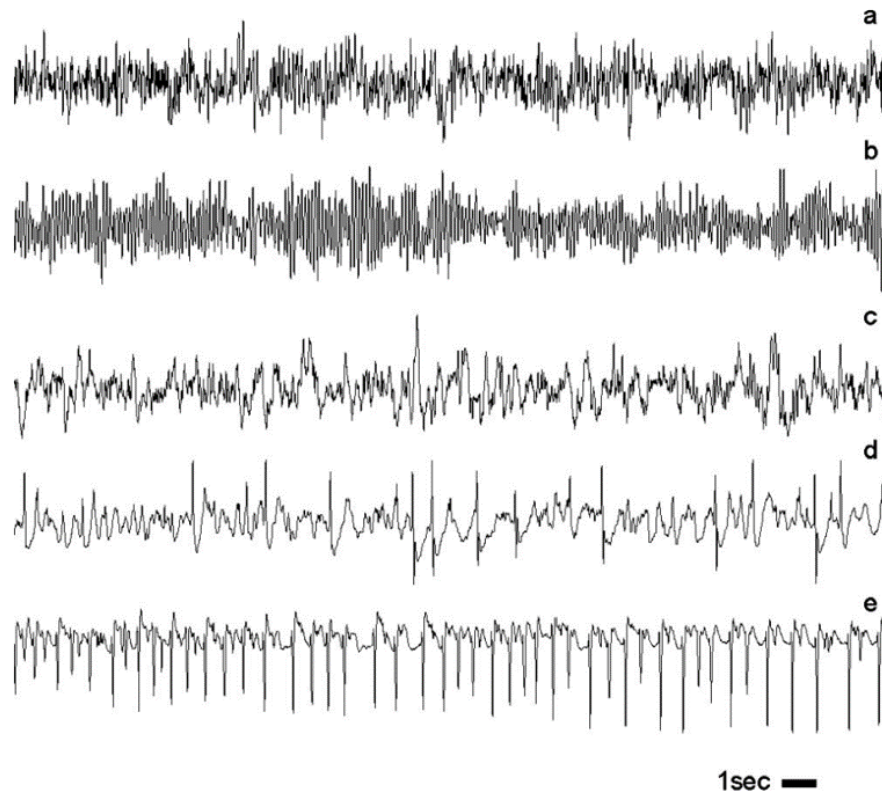


Figura 5.1 Señales de cada uno de los subconjuntos con una duración de 1 segundo, con un rango de $100 \mu\text{V}$ para las personas sanas y $1000 \mu\text{V}$ para personas con ataques. Figura tomada de [91].

5.1.2. Narcolepsia - BUAP

La base de datos de Narcolepsia, corresponde a un conjunto de señales EEG y ECG tomadas de ratas, las cuales tienen este padecimiento; los registros fueron etiquetados por expertos durante los diferentes eventos presentes. En el trabajo de Eguibar y Cortés [92] describen que los registros fueron tomados con una tarjeta de conversión analógico-digital con una velocidad de muestreo de 200 Hz, durante los eventos de sueño-vigilia. Las llamadas épocas de grabación sucedieron durante 10 segundos, es decir, el etiquetado de eventos es durante este periodo de tiempo. Por clase se cuenta con aproximadamente 2000 muestras preprocesadas de cuatro canales EEG y una señal ECG. Los datos contienen cuatro diferentes clases, REM1, REM2, wake y sleep, de cinco ratas distintas. A diferencia de otros conjuntos de datos, estas señales provienen de EEG implantados directamente (invasivos) a los animales.

5.1.3. Epilepsia - Boston

El conjunto de datos de Boston sobre epilepsia es más extenso en comparación del de Bonn, en este se registran 916 horas de grabación continua a un frecuencia de muestreo de 256 Hz y con el uso del estándar internacional 10-20 para la colocación de los electrodos. Los datos son de 22 pacientes del Hospital de niños de Boston (5 hombres, edades 3-22 y 17 mujeres, edades 1.5-19). Las grabaciones están divididas en carpetas que contienen archivos .edf, con épocas de grabación de 10 segundos o menos; la mayoría de los archivos contienen 23 señales de los canales, algunos con 24 o 26. Todos el conjunto conforma un total de 664 instancias de las cuales 198 son de casos de epilepsia [93].

5.2. Parámetros de la STFT

La STFT es representada por medio de espectrogramas en un plano de tiempo-frecuencia, mostrando los niveles de energía en colores. Sin embargo, dependiendo de la selección de los parámetros descritos en la sección 2.5, los valores en tiempo-frecuencia-energía pueden cambiar. Estos cambios en el tiempo y la frecuencia se han vuelto comunes en la literatura moderna sobre el procesamiento de señales como una forma de representar el tiempo de la señal y las resoluciones de frecuencia.

La STFT presenta un inconveniente relacionado con la resolución de los espectrogramas, provocado principalmente por el tamaño de la ventana. Una ventana estrecha ofrece una mejor resolución en el tiempo, pero no en frecuencia, mientras que una ventana ancha proporciona una buena resolución en frecuencia, pero no resulta de la misma forma para la resolución en el tiempo. Por lo tanto, es difícil lograr un buen ajuste en los dominios de tiempo y frecuencia simultáneamente, porque la STFT depende de un solo tipo de ventana con una misma longitud, encargada de segmentar alguna la señal; en este trabajo, se propuso una forma de calcular el tamaño de la ventana.

El proceso de la STFT comienza con la ventana que se aplica a segmentos más cortos de una señal, es decir, donde la fase de las ventanas pueden superponerse; la transformada rápida de Fourier (FFT) se aplica a cada segmento por separado [82] a través de las ventanas y cada resultado se une para formar el espectrograma, que puede manipularse y luego la FFT inversa se aplica para devolver cada segmento al dominio del tiempo.

Por lo que, para aplicar el STFT a las señales, se deben considerar las características de la transformación, tales como: tipo y longitud de la ventana, longitud del lóbulo

principal (correspondiente a las ventanas del espectro), puntos FFT y superposición. En las siguientes subsecciones se describen la selección de algunos de estos parámetros.

5.2.1. Longitud de la ventana

Los parámetros de la STFT son importantes para obtener una buena resolución del espectrograma, durante la experimentación inicial se analizaron parámetros a partir de la revisión de la literatura. En un inicio Van G., et al., para el análisis de señales EEG, propusieron ventanas del tamaño de 1s; Tzallas uno de los autores que comenzó a usar la STFT en la identificación de epilepsia con EEGs, aplicó en un trabajo una ventana Hanning con una longitud de 64 [49]; Nicolae realizó un trabajo sobre los efectos de la STFT para la estimación espectral en EEG y empleó diferentes ventanas (Hamming, Hanning, Gaussina y Barlett), con longitudes de 128, 256, 512 y 1024 y traslapes de 95 % [94]; Zabidi en su investigación sobre el análisis de EEG durante escritura imaginada usó ventanas tipo Hamming con longitudes de entre 0.1 y 0.8 segundos [56]; Duque L., et al., en un trabajo para la identificación y monitoreo cerebral con EEGs propone el uso de una ventana Gaussiana de longitud de 2.9 segundos con un traslape de 40 % [25]; en un estudio más reciente de Gerla V., et al., sobre el análisis de sueño con señales EEG usaron ventanas de longitudes de 1, 3 y 60 segundos [95]. Cada uno de los enfoques aplican diferentes parámetros, en general estos provienen de la experimentación, sin justificar su uso.

En esta investigación se propuso una forma para la justificación de los parámetros que tienen mayor efecto sobre el espectrograma: tipo de ventana y tamaño. Esto se realizó por medio de un análisis visual de los espectros y las señales, así como la experimentación de diferentes valores de los parámetros.

Inicialmente, se seleccionaron diferentes tipos de ventanas: Bartlett, Blackman, Hamming, Gauss, Kaiser y rectangular, con una longitud de 64 y 128, como ventanas de prueba. Con un análisis visual de los lóbulos (principal y secundarios) de sus espectros específicamente se puede seleccionar una. Un lóbulo principal cumple la función de representar la frecuencia fundamental y los lóbulos secundarios son representados por los armónicos restantes; si un lóbulo principal es estrecho, la frecuencia es más legible, sin embargo, este tipo de comportamiento ocurre con un espectro con lóbulos secundarios con niveles de energía altos, la cual se distribuye en ellos; por el contrario, si el espectro de una ventana contiene lóbulos secundarios con energía baja el lóbulo principal será ancho como se muestra en la figura 5.2, la cual será referenciada en el resto del párrafo. La ventana rectangular es la ideal en el tiempo ya que divide la señal sin modularla, esto es lo que ocurre cuando se aplica la FFT a una señal completa, se

aplica la STFT con una ventana con una longitud del mismo tamaño que la señal. En la figura 5.2 se observa un espectro con un lóbulo principal muy delgado, sin embargo, sus lóbulos secundarios son muy altos. Lo ideal es tener algo balanceado entre tiempo y frecuencia, lo cual se logra con una ventana muy parecida a su espectro. La ventana Gaussiana se puede considerar como la más óptima por el hecho de tener un espectro muy parecido a su ventana, sin embargo, computacionalmente no podemos obtenerla, por eso se recurre a aproximaciones.

La ventana Blackman muestra los lóbulos secundarios más pequeños, seguido de Hanning, que tiene una ventana con un lóbulo principal angosto: estas ventanas tienen la característica que sus espectros son parecidos a ellas, lo cual se debe a que son variantes del coseno alzado. Tomando en cuenta todas las ventanas, el orden que toman con respecto al tamaño del lóbulo principal (del más ancho al más angosto) es el siguiente: Blackman, Gaussian, Hamming, Hanning, Kaiser y Rectangular; con respecto a los lóbulos secundarios el nivel de energía que presentan de mayor a menor quedan de la siguiente forma: Rectangular, Kaiser, Hamming, Gaussian, Hanning y Blackman.

A partir de las características de los espectros, se nota que no hay una que cumpla los requerimientos de: lóbulo principal angosto y lóbulos secundarios pequeños. Una opción justa es una ventana con un término medio como la tipo Hanning, debido a que tiene un lóbulo principal medio y sus lóbulos principales son de energía media en comparación de las otras ventanas. Debido a sus propiedades, una reconstrucción perfecta y alias de los lóbulos laterales solo es posible si la superposición de ventanas es 75% [96] con respecto a la longitud.

La longitud es el otro parámetro que interviene en la resolución del espectrograma; en [97], Mateo et al., muestran un método para ajustar una ventana a partir de tres características principales: frecuencia, periodo y número de periodos o ciclos de una señal dentro de una ventana. En este trabajo para la aplicación de los enfoques, se propuso algo semejante a al trabajo de Mateo, sin embargo, no se toma en cuenta periodo y se utiliza frecuencia de muestreo, ya que es un parámetro que se relaciona con la visualización de los componentes de Fourier.

Por lo tanto, se parte del parámetro de frecuencia de muestreo de las señales EEG de las bases de datos (ataques epilépticos y la de estados de sueño) se tiene que estas son de 173.61 Hz y 200 Hz respectivamente. Por otra parte, las señales EEG tienen diferentes bandas, las cuales en los seres humanos se perciben a partir de 1 Hz [24] y para los estados de sueño en las ratas también es de 1 Hz [92], entonces se identificaron estas frecuencias desde 1 Hz para no discriminar alguna frecuencia de la señal presente. Si las frecuencias de las señales son de 1Hz la duración en tiempo es de 1 segundo de

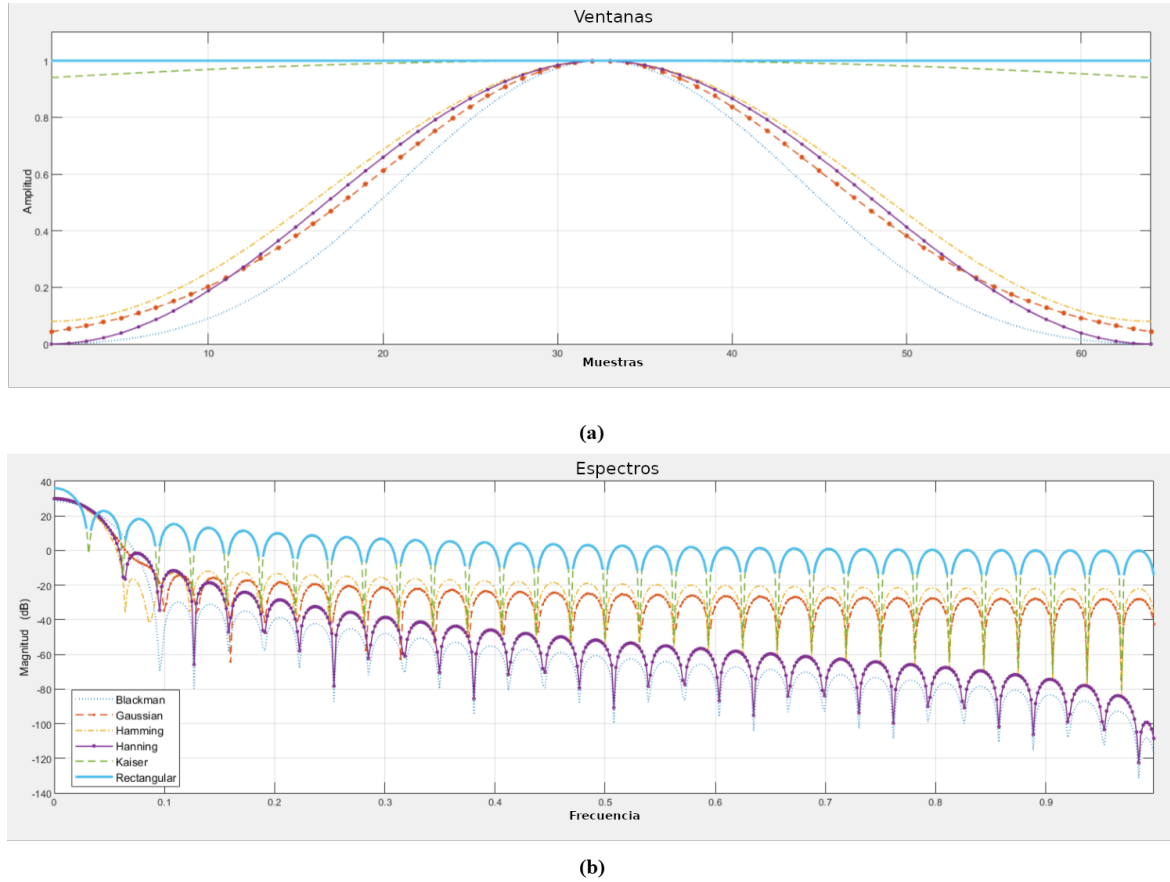


Figura 5.2 Ventanas (a): (....)-Blackman, (_ . _)-Gaussian, (-.-.)-Hamming, (-.-)-Hanning, (_ _ _)-Kaiser, y (___)-Rectangular.

un ciclo completo, sin embargo, no se tiene la certeza del comienzo y fin de un ciclo, por lo que se sugiere que las ventanas sean del doble (2 segundos para este ejemplo). Por lo tanto, los componentes de frecuencia en la transformada de Fourier discreta (DFT) son espaciados a partir de las siguientes relaciones [98]:

$$\Delta f = \frac{F_s}{N} \quad \text{ó} \quad \Delta f = \frac{1}{NT} = \frac{1}{T_0}, \quad (5.1)$$

Donde F_s es la frecuencia de muestreo, T es el intervalo muestreado, $T = \frac{1}{F_s}$ y N es el periodo de la secuencia resultante, Δf es la frecuencia de espacio (también llamada frecuencia de resolución) y T_0 es la longitud grabada ($T_0 = NT$). Para tener frecuencias espaciadas cercanamente o para incrementar la resolución, Δf necesita ser pequeño [98].

Los parámetros elegidos fueron a partir de cierta resolución en frecuencia, indicando desde que frecuencia se notan los cambios en las señales; la cual es determinada por

el lóbulo principal de la ventana, que está relacionado con la forma y longitud de la ventana [99]. Entonces la ecuación 5.1 por la parte izquierda se multiplica por un factor m correspondiente al lóbulo principal por el tipo de ventana, $m = 2$ para la rectangular, $m = 4$ para Hanning, Hamming y Barlett, y $m = 6$ para Blackman [99, 98], quedando la ecuación 5.1 como la ecuación 5.2:

$$\Delta f = m \frac{F_s}{N}, \quad (5.2)$$

Como se mencionó, se propone el uso de la ventana Hanning, donde el factor es de $m = 4$ y para frecuencias de 1 Hz, las ventanas deben ser de 1 segundo, ya que el inverso de las frecuencias nos dice cuanto dura el periodo de las señales. Por lo tanto, para la base de datos de ataques epilépticos se tiene que para una duración de 23.6 segundos, hay 4096 muestras, entonces para frecuencias de 1 Hz, el tamaño de la ventana en número de muestras es de 173.55. Anteriormente, se mencionó que se tenía que elegir al menos el doble, es decir, el nuevo tamaño es de 347.11, esta ventana no se puede hacer más grande, debido a que: empleando ventanas grandes se detectan mejor las frecuencias bajas y las bandas del EEG contienen señales con frecuencias muy variables. Para la selección del tamaño con respecto a un número de muestras, solo se ponderó el resultado obtenido ya que comúnmente contiene decimales.

Lo descrito anteriormente, es un método para la selección de parámetros de la STFT de cualquier señal. En los experimentos realizados, se aplicó una ventana Hanning con diferentes longitudes, las cuales se calcularon con respecto a las frecuencias de las señales analizadas (epilepsia y narcolepsia), estos resultados se muestran en la tabla 5.1. La primera columna corresponde a los tipos de señales; la segunda, muestra la clase del conjunto de datos utilizado y la frecuencia en la que oscilan las señales. Para una persona saludable en estado relajado son al rededor de 1 Hz [4, 91]; con respecto a un ataque epiléptico presente la frecuencia es mayor o igual a 2.5 Hz [91, 100] y para un estado despierto(wake) de la base de datos de sueño en ratas se tomó al menos 12 Hz [92]. La tercera columna presenta el tamaño de la ventana en segundos y número de muestras, como se describió anteriormente, el valor viene dado por el doble del periodo (basados en las frecuencias mostradas de la segunda columna). La cuarta columna, corresponde al valor de la longitud del traslape que es de un 75 % de la ventana propuesta, este valor se propone a partir de las características del lóbulo principal y secundarios de una ventana Hanning [101]. Durante la evaluación de los enfoques, se usaron las longitudes de las ventanas mostradas en la tabla 5.1, así como el traslape para la ventana tipo Hanning.

Tabla 5.1 Características de las ventanas con respecto a las clases.

Tipo de datos	Clase y frecuencia	Longitud de vector EEG (tiempo y número de muestras)	Ventana propuesta (tiempo y número de muestras)	Longitud del traslape a un 75 %
Epilépsia	Saludable ≥ 1 Hz	23.6s y 4097	2s y 347	260
Epilépsia	Ataques ≥ 2	23.6s y 4097	1s y 174	131
Estados de sueño	stage1 ≥ 1 Hz	10s y 1997	2s y 399	300
Estados de sueños	wake ≥ 12	10s y 1997	0.16s y 33	25

5.3. Evaluación del enfoque EEFGabor

La metodología presentada en esta sección se enfoca en extraer características de los espectrogramas EEG a través de filtros 2D Gabor, como se señaló en la sección 4.3.1. En este apartado se muestran los resultados de la aplicación de ese enfoque con las bases de datos de Epilepsia-BONN y Narcolepsia-BUAP; para el primer conjunto se consideraron dos subconjuntos: S (personas sufriendo un ataque epiléptico) y Z (personas sanas), y para el segundo stage1 y awake. Las características extraídas con EEFGabor se emplearon para entrenar clasificadores SVM y ANN.

La figura 5.3 muestra las etapas de este enfoque: inicialmente, los espectrogramas obtenidos fueron convertidos en imágenes en escala de grises de 8 bits y luego se dividieron en cinco subimágenes relacionadas con las bandas de frecuencia de los ritmos para localizar estructuras significativas. Esto se basó en las bandas propuestas por [89]; otros trabajos, como Duque et al., Mustafa et al., y Kovacs et al., [25, 1, 102], propusieron dividir el espectrograma. Para este enfoque, solo se usaron cuatro bandas, ya que los niveles de energía de gamma fueron nulos. Los rangos empleados fueron: Delta (0-4Hz), Theta (4-8Hz), Alpha (8-12Hz), Beta (12-30Hz) y Gamma (30-50Hz).

Los filtros Gabor empleados, se configuraron con cinco frecuencias y ocho orientaciones según lo sugerido por Haghighat et al. [35], obteniendo 40 imágenes por espectrograma, éste fue el resultado de la convolución de los filtros con cada imagen; se usó sólo la representación en magnitud y las características obtenidas fueron energía y entropía, teniendo vectores de 40 y 80 elementos por señal debido a dos características obtenidas.

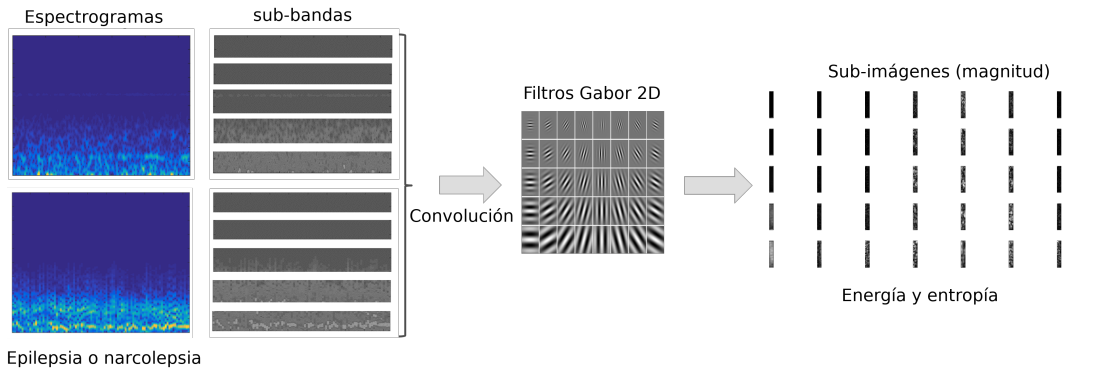


Figura 5.3 Generación de la respuesta de Gabor a partir de los espectrogramas.

5.3.1. Clasificación de epilepsia de Bonn con EEFGabor

Durante la aplicación de este enfoque se obtuvieron dos tipos de características de textura (energía y entropía) y se generaron diferentes vectores reuniendo las características extraídas para tres diferentes experimentos. Para evaluar estas características, se aplicó SVM con diferentes núcleos (lineal, polinomial y RBF) y ANN, según lo reportado en la literatura [25, 51, 64] para la clasificación de epilepsia. El rendimiento de los clasificadores se evaluó utilizando la exactitud como métrica en términos de porcentaje; los valores presentados se obtuvieron de un promedio de 10 experimentos, ya que se utilizó la validación cruzada con 10 pliegues estratificados.

En el primer experimento, se utilizó la energía como característica; por lo tanto, el vector de características tuvo una dimensión de 200 unidades. Se entrenaron SVM y ANN, obteniendo diferentes resultados, como se muestra en la tabla 5.2. La mejor exactitud se obtuvo mediante un SVM-Polinomial, con un 98%; para el SVM-lineal y el perceptrón se obtuvo el mismo resultado. La desviación estándar de SVM fue menor, con 0.024, lo que significa que los 10 experimentos fueron similares; con el perceptrón, se cambió el número de épocas (100 y 500) y el resultado fue el mismo.

Como segunda prueba, se empleó la entropía como característica única y se utilizaron los mismos clasificadores que en el experimento anterior (tabla 5.2); el SVM polinomial mostró la mejor exactitud con 97.5% y la peor con SVM RBF de 93.0%; uno de los perceptrones obtuvo un 94% como mejor resultado. Los resultados son mostrados en la tabla 5.3.

Para el último experimento se combinaron la entropía y la energía, obteniendo un vector de características de 400 unidades. La exactitud fue mejor que en los dos experimentos anteriores con un SVM polinomial, con una exactitud de 98.5%; la más baja fue de 95.5%, con una SVM-RBF. La tabla 5.4 muestra los resultados. Sin

Tabla 5.2 Resultados de la clasificación con características de energía con perceptrones multicapa y SVM.

Clasificador	Parámetros	Exactitud	Desviación estándar
SVM	Polinomial	98 %	0.024
SVM	Lineal	97 %	0.033
SVM	RBF	96 %	0.044
Perceptrón 1	Una capa oculta con 10 unidades, backpropagation y 200 épocas.	95 %	0.042
Perceptrón 2	Tres capas ocultas con 10, 20 y 10 unidades, backpropagation y 200 épocas.	96.5 %	0.032
Perceptrón 3	Dos capas ocultas con 10 y 10 unidades, backpropagation y 200 épocas.	97 %	0.024

Tabla 5.3 Resultados de la clasificación con características de entropía con perceptrones multicapa y SVM.

Clasificador	Parámetros	Exactitud	Desviación estándar
SVM	Polinomial	97.5 %	0.040
SVM	Lineal	97 %	0.051
SVM	RBF	93 %	0.051
Perceptrón 4	Una capa oculta con 10 unidades, backpropagation y 200 épocas.	92 %	0.090
Perceptrón 5	Dos capas ocultas con 10 y 10 unidades, backpropagation y 200 épocas.	93.5 %	0.084
Perceptrón 6	Tres capas ocultas con 10, 50 y 50 unidades, backpropagation y 200 épocas.	94 %	0.070

Tabla 5.4 Resultados de la clasificación con características de entropía y energía con perceptrones multicapa y SVM.

Clasificador	Parámetros	Exactitud	Desviación estándar
SVM	Polinomial	98.5 %	0.032
SVM	Lineal	98 %	0.04
SVM	RBF	95.5 %	0.0437
Perceptrón 7	Una capa oculta con 10 unidades, backpropagation y 200 épocas.	95.5 %	0.052
Perceptrón 8	Tres capas ocultas con 10, 40 y 50 unidades, backpropagation y 200 épocas.	96.5 %	0.032
Perceptrón 9	Dos capas ocultas con 50 y 50 unidades, backpropagation y 200 épocas.	97 %	0.04

embargo, aunque estos son los mejores resultados con un SVM polinomial y lineal, su desviación estándar fue de 0.032 y 0.04 respectivamente, los cuales son valores mayores con respecto al uso de energía solamente (0.024).

Algunos experimentos sugieren extraer la media y la desviación estándar de la energía como características de textura usando filtros de Gabor [87, 88]. Sin embargo, utilizando la desviación estándar y la energía media, la exactitud inferior al 80 % con SVM, mientras que el perceptron con dos capas ocultas de 20 unidades alcanzó el 95 %. Por otro lado, empleando la desviación estándar y la media de la entropía, SVM y Perceptrón alcanzaron menos del 80 %. Finalmente, concatenando las cuatro características, la exactitud fue del 95 % para SVM Polinomial, 92,5 % para SVM lineal, 82,5 % para SVM RBF y 94,5 % para un Perceptrón.

La energía y la entropía se utilizan como características, sin embargo, probablemente, la energía es la característica predominante debido a que en la mayoría de los clasificadores fue mejor que la entropía. Por otro lado el mejor rendimiento se obtuvo utilizando SVM Polinomial para cualquier característica.

5.3.2. Clasificación de narcolepsia con EEFGabor

Para los experimentos realizados con el conjunto de estados de sueño, se usaron dos clases: stage1 y awake, debido a que el número de instancias por clase no está presente en todas las instancias. Se propusieron dos tipos de experimentos: el primero, utilizando todos los ejemplos de las instancias y evaluando con una validación cruzada; para el segundo, se usaron las señales de los sujetos de prueba como entrenamiento menos uno, las señales restantes correspondientes a una rata se utilizaron para la prueba. A

diferencia del conjunto de datos epilepsia, las señales oscilan a diferente frecuencia, por lo que los parámetros de la STFT aplicado son los mostrados en la tabla 5.1 del cuarto renglón, con una ventana de 2s (399 muestras) y un traslape de 300 muestras. Los parámetros de los clasificadores fueron como los del experimento de la subsección 5.3.3; aplicando un balanceo al conjunto.

Para el primer experimento con este conjunto de datos, los parámetros de la STFT fueron similares a los de epilepsia, se usó una ventana de tipo Hanning con una longitud de 399 muestras, el traslape de 75 % de la ventana fue de 300 muestras. La figura 5.4 representa el promedio de los resultados de la clasificación, los cuales fueron obtenidos usando todo el conjunto de datos con una validación cruzada de 10 pliegues, debido al número de instancias que se tiene por clase (stage1=425 y wake=146). Se utilizaron tres algoritmos con diferentes parámetros: para SVM-lineal se obtuvo una exactitud del 83 %; con un SVM-Polinomial el 93 %, obteniendo la exactitud más alta; con kNN con una $k=3$ y una distancia minkowsky el resultado fue de 90.10 %. Con el perceptrón multicapa con 1 capa oculta de 18 unidades, los resultados fueron los mas bajos (75.3 %).

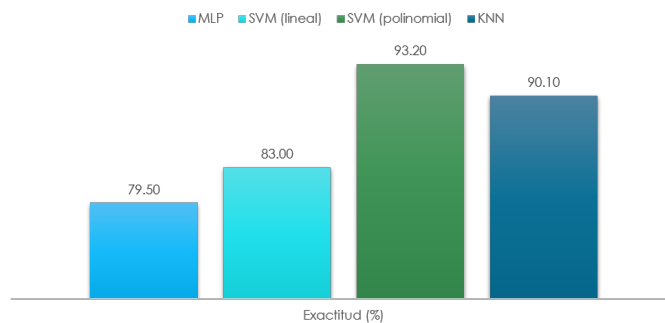


Figura 5.4 Resultados de la clasificación para todo el conjunto de estados de sueño.

El segundo experimento siguió el siguiente proceso: inicialmente, se entrenó con las señales de cuatro ratas de las cinco disponibles; posteriormente, la prueba se hizo con la que no se usó; este proceso se realizó para cada uno de las ratas. Los parámetros de la STFT utilizados para este experimento fueron los mismos que el experimentos anterior, los mostrados en la table 5.4; adicionalmente, se aplicó balanceo al subconjunto obtenido de las cuatro ratas. Cada uno de los subconjuntos estaba desbalanceado, con una proporción promedio de 1 a 3, por tal motivo, los cinco subconjuntos obtenidos de la combinación de las cinco ratas fueron balanceados. Torres [103] menciona que el balanceo ayuda al desempeño en la clasificación de señales EEG en emociones y Cui et al.[46] hizo una evaluación de técnicas de balanceo en un problema de señales EEG,

Tabla 5.5 Resultados de la clasificación del conjunto de estados de sueño de cada rata.

Clasificador	531(%)	693(%)	694(%)	698(%)	701(%)
MLP (30,30,12,12)	77.3	92.8	91.8	93.5	57.0
SVM (lineal)	81.8	90.3	91.3	87.3	69.6
SVM (polinomial)	78.8	94.2	90.8	95.7	69.6
kNN	81.8	87.1	87.2	89.1	53.2

en el cual menciona que la técnica SMOTE obtiene los mejores resultados para una clasificación de señales de este tipo; por tal motivo, se aplicó la técnica mencionada.

Los resultados son mostrados en la figura 5.5 y en la tabla 5.5, en la primera columna se presentan los clasificadores utilizados; las cinco columnas restantes de la derecha reflejan números que corresponden al número de rata con la que se hizo la prueba. Estos resultados denotan que la SVM-polinomial tiene los mejores resultados en la mayoría de los sujetos; para la rata 531 y 701 se tienen las menores exactitudes, ésto se relaciona con el número de instancias utilizadas para la prueba ya que para las dos ratas stage1 es casi el doble de wake, y para los demás conjuntos stage1 es casi 5 veces más.

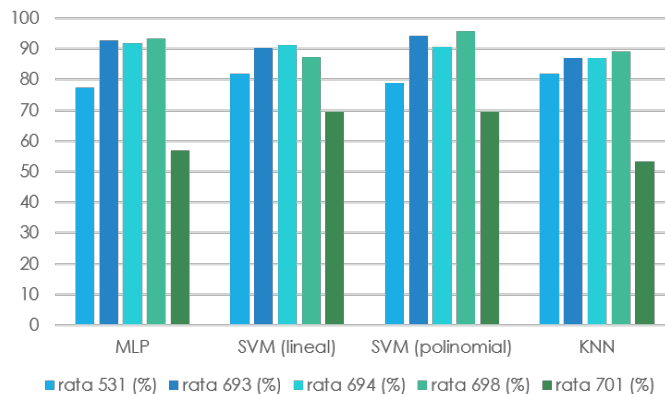


Figura 5.5 Resultados de la clasificación para todo el conjunto de estados de sueño por rata.

5.3.3. Experimentos complementarios de epilepsia con EEGabor

Los experimentos mostrados en la subsección 5.3.3 se realizaron con ventanas tipo Hanning de longitudes iguales para ambas clases, sin embargo, debido a que cualquier tipo de señal oscila en diferentes frecuencias, se probó con dos longitudes de ventanas,

Tabla 5.6 Nomenclatura del longitud de las ventanas Hanning.

Nomenclatura	Longitud de ventana (muestras)
v1	347
v2	174
v3	174 para epilepsia y 347 para sano
v4	347 para epilepsia y 174 para sano

Tabla 5.7 Resultados de la clasificación del conjunto de estados de epilepsia para diferentes tipos de ventana.

Clasificador	v1(%)	v2(%)	v3(%)	v4(%)
MLP (30,30,12,12)	62.0	55.5	55.5	61.0
SVM (lineal)	88.5	85.5	87.5	99.5
SVM (polinomial)	79.5	83.5	66.0	90.5
kNN	89.0	85.0	86.0	99.0

una para cada clase. La tabla 5.6 describe la nomenclatura usada para el tamaño de las ventanas, con v1 y v2 se usan las mismas longitudes para todas las clases; para v3 y v4 se usan diferentes longitudes por clase.

Estos experimentos fueron inicialmente para una clasificación binaria y posteriormente para una multiclase (3 y 5 clases). Para el primer, se usaron los conjuntos personas sanas (A) contra sujetos con ataques epilépticos presentes (E); en el segundo, se utilizaron tres clases, los empleados en el experimento binario, junto con el personas que sufren epilepsia sin ataque presente (F); para el experimento final fue con los cinco subconjuntos (A-E).

Los resultados de este experimento son mostrados en la figura 5.6 y la tabla 5.7, ambas tienen una nomenclatura descrita en la tabla 5.6 que corresponde a abreviaciones de las longitudes de las ventanas Hanning.

La tabla 5.7 y la figura 5.6 muestran los resultados con cada una de las propuesta de la tabla 5.6. Se observa que no mejoraron con respecto a experimentos mostrados en la subsección para las ventanas propuestas v1, v2 y v3, sin embargo, para las ventanas de v4 se superaron valores ya mostrados en otros experimentos. Se alcanzó un exactitud máxima de 99.5% con un SVM lineal y un 99% con kNN, tomando una ventana para cada clase, a diferencia de v1 y v2 que usan la misma ventana para ambas clases. Estos resultados muestran que los mejores desempeños con respecto a exactitud fueron para los clasificadores SVM-lineal y kNN.

A partir de los mejores parámetros del experimento para dos clases del conjunto de epilepsia, se usaron los mismos para tres y cinco clases, tomando una ventana de

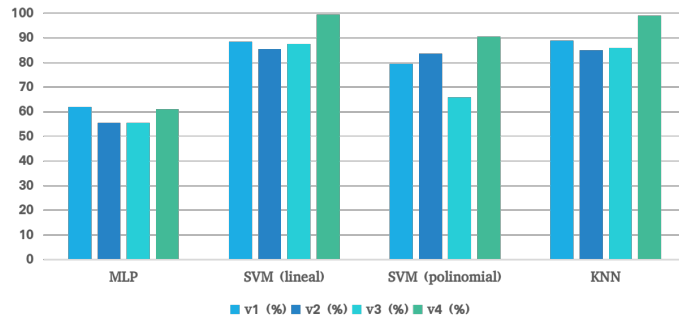


Figura 5.6 Resultados de la clasificación para dos clases del conjunto de epilepsia para diferentes ventanas.

Tabla 5.8 Resultados de la clasificación del conjunto de estados de epilepsia con los parámetros de V4 para tres y cinco clases.

Clasificador	A, E y C (%)	A, E y D (%)	A,B,C,D,E (%)
MLP	64	64	55
SVM (lineal)	94	94.7	73.4
SVM (Polinomial)	85	86	62.4
kNN	94	93.3	74.5

longitud 345 para la clase epilepsia y de 174 para las ventanas restantes. La figura 5.7 y la tabla 5.8 muestran los resultados de estos experimentos. En la primera columna se muestran los clasificadores; la segunda y la tercera corresponde a los resultados de la clasificación de tres clases; la última corresponde a la clasificación de los cinco tipos de señales.

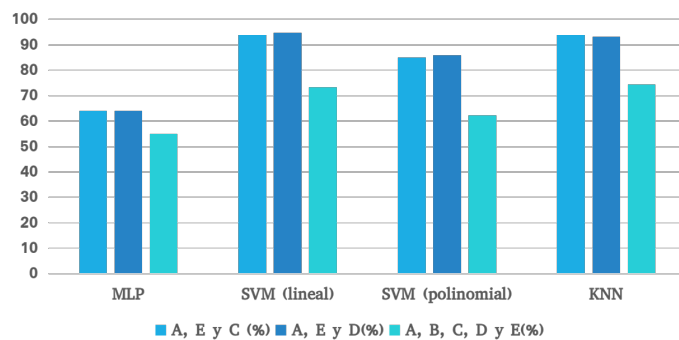


Figura 5.7 Resultados de la clasificación multiclase para todo el conjunto de epilepsia.

Con base en los resultados, se observa que los primeros dos experimentos (columna dos y tres de la tabla 5.8), resultan muy similares, las exactitudes más altas son de 94% y 94.7% respectivamente, se puede decir que no hay diferencia entre la zona

del cerebro de donde se toman las muestras para este caso. Con respecto a la última columna, se tienen exactitudes menores de 73.4% y de 74.5% para un SVM-lineal y kNN respectivamente, lo cual tiene coherencia al analizar más número de clases, y tomando en cuenta que C y D son conjuntos muy similares así como A y B que son de personas sanas con ojos cerrados y abiertos; los subconjuntos con mayor diferencia son A y E (sanos con ojos abiertos y pacientes con ataques epilépticos presentes), visualmente se nota la diferencia entre ellas. Estos experimentos, nos muestran que para cada tipo de señal se debe usar una longitud de ventana que se ajuste a la frecuencia en la que oscilan.

5.4. Evaluación del enfoque FOSP

El segundo enfoque está basado en tres técnicas descritas en la sección 4.3.2, el cual se evaluó con las dos bases de datos de epilepsia (Bonn y Boston). Sin embargo, con la de Bonn se hicieron evaluaciones estadísticas para los experimentos, debido a que en estos se obtuvieron los mejores resultados en clasificación. En esta primera subsección se muestran los resultados obtenidos del conjunto de Bonn con el enfoque FOSP y las diferentes técnicas; posteriormente, se presenta una evaluación del enfoque con el conjunto de Boston.

Como se señaló en la sección 4.3.2, se obtuvo el máximo para cada ventana por instancia. El número de máximos fue de 44 y de estos valores, se obtuvieron el mínimo, la media y el máximo para ser utilizados como características. La figura 5.8 muestra el comportamiento de las 200 instancias con sus 44 picos máximos. El eje horizontal representa el pico máximo para todas las ventanas de la 1 a la 44, mientras que el eje vertical representa la energía de estos picos. Los picos rojos corresponden a personas que sufren una crisis epiléptica abierta y los picos azules corresponden a personas sanas con los ojos abiertos. Se observa que con respecto a la energía de estos picos se pueden diferenciar las clases, ya que las personas que sufren una crisis epiléptica tienen una energía más alta que las de las personas sanas.

La figura 5.9 representa las características obtenidas para máximo, medio y mínimo, donde el eje x representa las 100 instancias y el eje y la energía. En esta figura, se representan los picos mínimo, medio y máximo de los máximos encontrados por la ventana de cada instancia. Las líneas continuas son de la clase de epilepsia y las líneas punteadas son de personas sanas. El color rojo representa el máximo pico por instancia, el verde el mínimo y el azul la media.

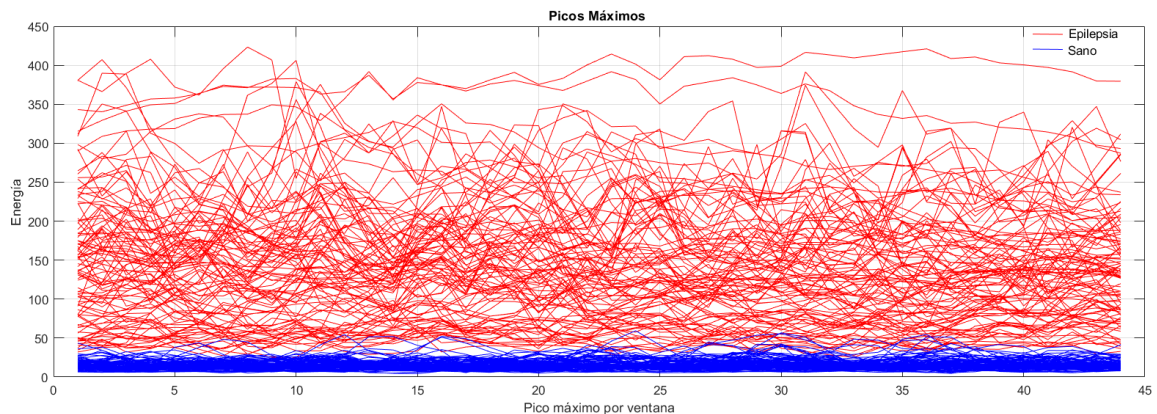


Figura 5.8 Representación de las 200 señales, 100 por cada clase usando sólo picos máximos por ventanas.

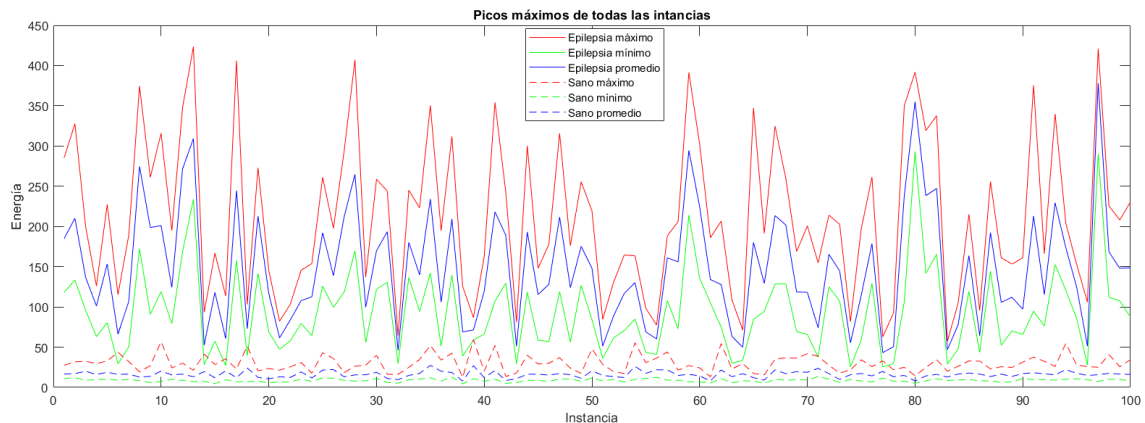


Figura 5.9 Picos máximos, mínimos y promedio de cada ejemplo.

El pico mínimo de todos los máximos encontrados para cada ventana de espectrograma se seleccionó como una característica en el último método, ya que mostró la desviación estándar más baja de 4.074 con respecto al máximo y la media.

5.4.1. Clasificación de epilepsia de Bonn con enfoque FOSP

Para la etapa de evaluación del conjunto de Bonn, se entrenaron cuatro clasificadores: MLP, SVM lineal, SVM Polynomial y kNN; seleccionados con base en comparación de literatura en la que son los clasificadores predominantes en el análisis EEG basado en espectrogramas [104]. Se utilizaron dos clases para este experimento: personas sanas y pacientes que sufren una crisis epiléptica, cada clase con 100 instancias, para un total de 200 instancias para evaluación. En este experimento, se utilizó una validación cruzada

de cinco folds, porque con un conjunto de 200 ejemplos un subconjunto estadístico representativo es una muestra de 160 ejemplos, con un intervalo de confianza del 99% y un margen de error del 5%. Por lo que, al dividir los 200 ejemplos en cinco pliegues, el resultado es de 40 elementos por cada uno; por lo tanto, al usar cuatro folds implica 160 elementos, que son un subconjunto representativo para entrenar el modelo, lo que resulta en una evaluación más precisa y justa. Los resultados se muestran en forma de tres experimentos diferentes: experimento 1, superficies, frecuencia y LBP; experimento 2: energía, k-means y LTP; experimento 3: picos espectrales y LTP; los cuales representan los tres diferentes enfoques que fueron comparados con el trabajo de Segur et al [51]. Como métrica relevante para la clasificación, se utilizó el coeficiente kappa para determinar el grado de concordancia entre dos observadores, para los tres experimentos. Los parámetros de clasificación se obtuvieron a través de la experimentación utilizando cotas superiores e inferiores, encontrando los mejores y modificando sus valores. Para MLP, se usaron dos capas ocultas con 100 unidades cada una y 200 épocas; para la SVM se usó un kernel lineal y polinomial de grado 3; finalmente con kNN se utilizó una distancia euclidiana $k = 1$.

Para el ajuste de hiperparámetros de los clasificadores, se utilizó la búsqueda de rejilla, que es el proceso de búsqueda de parámetros óptimos para un modelo dado, que crea una cuadrícula espaciada de puntos factibles, luego estos se evalúan sobre la función objetivo en cada punto, el espacio de puntos está determinado alcanzar la exactitud requerida [105]. La definición de los valores de hiperparámetros para el MLP se determinaron a partir de: 4, 10, 100 para una capa oculta y (100,100) para dos capas; una función de activación *tanh* y *ReLU* porque evitan el problema de desvanecimiento de gradiente. Para resolver la optimización de pesos del momentum se emplearon: 0.0001 y 0.05 y la tasa de aprendizaje usó: 0.001, 0.1 y 1. En el caso de las SVM's, los valores de margen entre vectores fueron 0.1, 1, 100 y 1000, y los núcleos seleccionados fueron lineales, polinomiales y RBF. Estos parámetros fueron reportados en la literatura [106, 104] para MLP y [107] para problemas similares de EEG con el uso de SVM.

A partir de la búsqueda de rejilla, los mejores resultados se obtuvieron con MLP con un *alfa* de 0.0001, una función de activación *ReLU*, con dos capas internas de 100 unidades, un optimizador de pesos *Adam* el cual se basa en el gradiente estocástico y una tasa de aprendizaje de 0.1; los SVM seleccionados fueron de núcleos lineales y polinomiales, el valor del margen entre vectores fue de 100. Para el kNN, se eligió $k = 1$ porque el conjunto de datos EEG se preprocesó y el ruido se atenuó. Se evaluaron las

métricas Manhattan, Euclidean y Chebyshev, obteniendo los mejores resultados con Euclidean.

La tabla 5.9 muestra los resultados de la comparación con otro autor, donde se puede ver que la exactitud es similar a la propuesta presentada con FOSP. El primer experimento consiste en el enfoque FOSP combinado con el primer método basado en las superficies obtenidas de volumen, área, picos de energía, número de picos y energía total, además de utilizar el primer experimento para la función de tiempo LBP, que codifica las frecuencias de pico en mayor que, igual a y menor que, formando una característica vectorial de 6 dimensiones. Los parámetros STFT fueron los propuestos por el autor [89]. También se usaron las ventanas propuestas basadas en las frecuencias de las señales de EEG; sin embargo, los resultados con respecto a la clasificación fueron bajos, mientras que el mejor resultado fue 99.5 % con un SVM lineal y sus coeficientes kappa promedio fueron 0.98, 0.99, 0.95, 0.87 para MLP, SVM lineal, SVM Polinómico y kNN, respectivamente; por lo tanto, su interpretación fue casi perfecta.

El segundo experimento aplicado: FOSP con LTP, energía y K-medias, mostró un buen rendimiento, muy cercano al 100 % de exactitud. La técnica LBP al adaptarlo a los picos, no contemplaba por completo las características de estos, ya que se hacían dos comparaciones de tres posibles, es de decir de tiempo, frecuencia y energía solo se comparaban con dos de estas; por lo que, debido a que LTP hacia estas comparaciones de una forma más justa se utilizó como alternativa. Se decidió probar con las características obtenidas en los experimentos iniciales para reforzar LTP, utilizando la energía total, el número de elementos por grupo y la energía representada por los centroides, extraídos utilizando K-means con una $K = 2$, así como dos códigos LTP, obteniendo un vector de 7 dimensiones, esta combinación de aprendizaje no supervisado y LTP para extracción de características reforzó la construcción de los modelos de clasificación. El resultado mejoró con respecto a los experimentos anteriores usando frecuencia y K-means para la extracción de características, alcanzando el 99.5 % para un SVM lineal y sus coeficientes kappa promedio fueron 0.96, 0.99, 0.9, 0.88 para MLP, SVM lineal, SVM polinomial y kNN, respectivamente; así, debido a la tabla de coeficientes kappa, la interpretación fue casi perfecta.

Los resultados obtenidos en el tercer experimento fueron del 100 % para la exactitud; en este método propuesto, se utilizaron los picos mínimos del espectrograma basados en los máximos de cada ventana o columna y características LTP del segundo experimento. Los resultados de la clasificación se compararon bajo las mismas condiciones metodológicas descritas por Sengür et al., utilizando técnicas de LBP, TFCM y GLCM. La comparación fue con el método que alcanzó los resultados más altos con respecto

Tabla 5.9 Resultados de la clasificación comparados con Sengür et al. [89].

Método	Clasificador	Exactitud(%)
Sengür	SVM	100
Sengür	Liblinear	100
Sengür	Homogeneous mapping + liblinear	99.5
FOSP		
Experimento 1	MLP	99
Experimento 1	SVM Linear	99.5
Experimento 1	SVM Polynomial	97.5
Experimento 1	kNN	93.5
FOSP		
Experimento 2	MLP	98
Experimento 2	SVM Linear	99.5
Experimento 2	SVM Polynomial	95
Experimento 2	kNN	94
FOSP		
Experimento 3	MLP	100
Experimento 3	SVM Linear	100
Experimento 3	SVM Polynomial	98.5
Experimento 3	kNN	99

a la precisión de la clasificación (usando LBP). Aunque existen otras metodologías, Sengür et al. [89] fue elegido para esta comparación debido a que tienen el rendimiento más relevante en términos de exactitud de clasificación. Finalmente, sus coeficientes kappa promedio fueron 1, 1, 0.97, 0.98 para MLP, SVM lineal, SVM Polinomial y kNN, respectivamente; por lo tanto, su interpretación fue casi perfecta.

Para validar los resultados de rendimiento, se realizó una prueba de significancia estadística de Friedman considerando $\alpha = 0.05$ como se sugirió en [45]. Este prueba es útil para varios clasificadores con una sola base de datos, la cual está basada en rankings; para su aplicación, se debe tener un conjunto de datos lo suficientemente grande, sin embargo, al no cumplirse el requerimiento, se aplicó 30 veces la validación cruzada, generando diferentes conjuntos de datos para los pliegues. Estos fueron usados en cada uno de las metodologías propuestas, donde se requiere que los pliegues generados sean los mismos para todos los enfoques y puedan compararse con cada uno de los clasificadores. El promedio obtenido de cada validación cruzada, ayuda a obtener lugares del ranking como se muestra en la tabla 5.10; el último renglón muestra las puntuaciones $\chi_F^2 = 86.3$, $\chi_F^2 = 81.9$ y $\chi_F^2 = 79.6$ para los tres experimentos, respectivamente, rechazando la hipótesis nula, por lo tanto, los resultados del clasificador muestran diferencias significativas y según la clasificación, el lineal SVM fue el mejor para los tres experimentos.

Tabla 5.10 Resultados de la prueba Friedman.

Clasificador	Ranking		
	Experimento 1	Experimento 2	Experimento 3
MLP	2	2	2
SVM-Lineal	1	1	1
SVM-Polinomial	3	4	3
kNN	4	3	4
χ_F^2	86.3	81.9	79.6

La complejidad computacional de los experimentos descritos es la siguiente: en la primera, la complejidad es $O(p \log p)$ debido al uso del algoritmo de cerco convexo rápido donde p son los picos que se encuentran y como máximo $p = 44$. Para el segundo experimento, la complejidad está relacionada con el proceso de agrupación por K-means, es decir, $O(imKt)$ donde $K =$ número de cluster que se encuentran i es el número de instancias en el conjunto de datos, m es el número de características que describen cada instancia y t es el número total de iteraciones durante el proceso de agrupación. El tercer experimento en nuestra propuesta (donde se aplica LTP) que tiene una complejidad de $O(wh)$, donde w , h son el ancho del espectro y la altura del espectro respectivamente; particularmente en nuestro enfoque $w = 44$ y $h = 512$. Entre nuestra experimentación, el mayor tiempo de ejecución se gasta en el segundo experimento, sin embargo, en nuestra propuesta $K = 2$, $m \ll i$ y, como es bien sabido, K-means converge en algunas iteraciones (en nuestro caso $t = 10$).

5.4.1.1. Experimentos complementarios

Para mostrar el rendimiento de la metodología con las características más relevantes (LTP y picos mínimos), se realizaron experimentos adicionales con los subconjuntos de la base de datos. El objetivo fue identificar cuándo un ataque epiléptico estaba presente, por lo tanto, a partir de la nomenclatura de la tabla 5.11 se muestran los cinco experimentos, donde se comparan algún subconjunto contra subconjunto E o la combinación de algunos subconjuntos contra el subconjunto E.

Estos experimentos fueron bajo las mismas condiciones del experimento FOSP con picos mínimos y LTP, mostrado en la Tabla 5.9, debido a que el enfoque empleado alcanzó una exactitud del 100 %.

La tabla 5.12 muestra los resultados de los experimentos con los subconjuntos propuestos, obteniendo exactitudes mayores o iguales a 90 %. Los resultados máximos de E2 se deben a la similitud de las señales de C y A, que se utilizan en los experimentos

Tabla 5.11 Nomenclatura de experimentos complementarios.

Nomenclatura	Experimentos
E1	B vs E
E2	C vs E
E3	D vs E
E4	C,D vs E
E5	A, B, C, D vs E

Tabla 5.12 Resultados de clasificación con nuevos subconjuntos para la identificación de epilepsia usando FOSP con picos mínimos y LTP.

Experimento	Clasificador	Exactitud(%)
E1	MLP	93.5
E1	SVM Lineal	92.0
E1	SVM Polinomial	94.0
E1	kNN	90.5
E2	MLP	99.5
E2	SVM Lineal	98.5
E2	SVM Polinomial	99.0
E2	kNN	97.5
E3	MLP	93.0
E3	SVM Lineal	92.0
E3	SVM Polinomial	91.0
E3	kNN	90.5
E4	MLP	95.0
E4	SVM Lineal	94.7
E4	SVM Polinomial	93.7
E4	kNN	95.0
E5	MLP	96.8
E5	SVM Lineal	96.2
E5	SVM Polinomial	96.2
E5	kNN	95.2

uno, dos y tres, mostrados en la tabla 5.9. En cuanto a los valores de exactitud más bajos, corresponden a señales obtenidas de otra área del cerebro; los resultados alcanzaron más del 90 % de exactitud. En los cinco experimentos se aplicó un análisis de coeficiente kappa; los clasificadores MLP, SVM-Lineal, SVM-Polynomial y kNN, obtuvieron los resultados promedio de los coeficientes kappa de la siguiente manera: 0.87, 0.84, 0.88, 0.81 para el primer experimento E1; en el segundo experimento (E2) se obtuvieron 0.86, 0.84, 0.82, 0.81; el tercero (E3), 0.86, 0.84, 0.82, 0.81; el cuarto (E5), 0.89, 0.88, 0.85, 0.87; finalmente, para E6, se obtuvieron 0.89, 0.86, 0.87, 0.84. El promedio de

Tabla 5.13 Resultados de la clasificación del conjunto de estados de epilepsia de Boston con características de volumen y tiempo.

Clasificador	S-NS (%)
ANN	76.2
SVM-Lineal	76.1
SVM-Polinomial	66.6
kNN	76.7

todos los coeficientes de kappa obtenidos en los cinco experimentos fue superior a 0.81, por lo que se encontró una interpretación casi perfecta.

5.4.2. Clasificación de epilepsia de Boston con enfoque FOSP

En la sección 5.1.3 se describió un conjunto de datos de epilepsia más extenso el de la Universidad de Boston. Usando el enfoque FOSP con las características de LTP y volumen, los del experimento 2 de la subsección 5.4.1, obteniendo los mejores resultados con ese enfoque para esta base de datos. Los resultados son mostrados en la tabla 5.13, donde la primera columna representa los clasificadores utilizados y la segunda los porcentajes de la exactitud de la clasificación de epilepsia contra no epilepsia.

Los resultados mostrados en la tabla 5.13 están por debajo de la exactitud reportada por Shoeb y Gutttag [93], sin embargo, su experimento se basó en una evaluación de una sola partición por sujeto; es decir, las grabaciones por persona se dividieron para entrenamiento y prueba. En comparación con nuestro experimento, se entrenó el modelo a partir de un segmento de todo el conjunto de datos: todas las señales de los canales disponibles de todos los sujetos y lo restante se utilizó para la evaluación, esto se repitió 10 veces, ya que se aplicó la técnica de validación cruzada con 5 pliegues.

5.5. Evaluación del enfoque basado en picos espectrales por medio de PEV

El último enfoque planteado se basó en DBSCAN como se describe en la subsección 4.3.3; se utilizaron los parámetros de la STFT propuestos en la sección 5.2. Sin embargo, aunque se siguen extrayendo picos espectrales, la etapa de extracción de características es distinta. Debido a que en ésta se contemplan puntos cercanos a partir del agrupamiento con DBSCAN y no solo por columna como se hacía anteriormente.

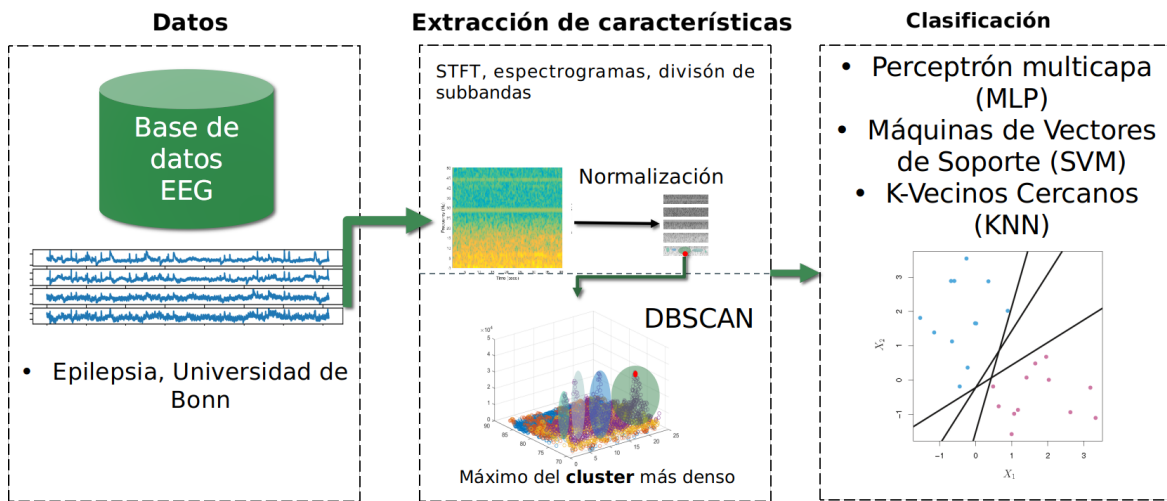


Figura 5.10 Enfoque de extracción de picos espectrales basado en PEV.

Como se muestra en la figura 5.10, a partir de la base de datos de EEG de la Universidad de Bonn, se aplicó la STFT con los parámetros ya previamente mencionados a la señales; posteriormente, se dividieron los espectrogramas en las cinco bandas ya conocidas con la finalidad de extraer características locales; después, una normalización min-max se aplicó para que la escala de las EEG no afectara a la extracción de características; un preprocesamiento de puntos máximos se aplicó a cada banda, donde a todos los puntos de energía se les agrupo por niveles a partir de cuartiles, descartando los puntos menores al valor de tercer cuartil; a los puntos máximos se les aplicó DBSCAN con métrica Manhattan y del cluster más denso se obtuvo el punto con mayor energía. Como esto se aplicó a las cinco bandas, el vector de características fue de dimensión cinco por cada instancia.

5.5.1. Clasificación de epilepsia de Bonn con enfoque PEV

Este último enfoque sólo fue complementario, sin embargo, se encontraron resultados aceptables en términos de clasificación para la identificación de epilepsia sobre una persona con ojos abiertos. La tabla 5.14 muestra los resultados del experimento, en donde un SVM-RBF alcanza la exactitud más alta con 98.5% de exactitud y los más bajos para el SVM-Polinomial con un 90.5%. Las características de evaluación fue la misma usada que en la descrita en la sección 5.4.

Tabla 5.14 Resultados de clasificación binaria(Epilepsia) usando DBSCAN con distancia Euclidiana.

Clasificador	E-A (%)
ANN	96.0
SVM-Lineal	96.0
SVM-Polinomial	90.5
SVM-RBF	98.5
kNN	92.5

5.6. Discusión

En las secciones anteriores se describieron los tres enfoques desarrollados, titulados EEFGabor, FOSP y PEV por medio de DBSCAN. La primer metodología (EEFGabor) es diferente a las otras dos, debido a que ésta no se basa en picos de energía, si no en una adaptación de extracción de características de textura sobre los resultantes de filtros Gabor, lo cual denota niveles de energía. Los resultados obtenidos con esta metodología con respecto a la exactitud fueron relevantes, debido a que con la base de datos de epilepsia de Bonn la clasificación alcanzó exactitudes cercanas al 100 % y para la de narcolepsia de la BUAP son variados, sin embargo, la mayoría de los resultados superan la exactitud del 90 %.

El segundo enfoque FOSP, hace uso de diferentes técnicas como LBP, LTP, máximos de energía, envolvente convexa, k-means; el cual dio los mejores resultados, específicamente, con el uso de LTP y puntos mínimos de los máxima energía de las ventanas de los espectrogramas (o columnas). Esta metodología obtiene los resultados más altos en comparación de los otros enfoques, sin embargo, los puntos máximos encontrados por ventana no contempla ventanas consecutivas o puntos máximos vecinos, lo que descarta el tiempo, que denota la ocurrencia de los eventos. El uso de LTP si contempla el uso de las tres características de los espectrogramas, tiempo, frecuencia y energía, lo que refuerza las características de los máximos. Se utilizó a la frecuencia de los puntos máximos como característica, sin embargo, al ser analizados por bandas no fue tan relevante debido a que las frecuencias siempre se encuentran en un rango marcado, delimitado por las frecuencias de las ondas; por otro lado, el tiempo no es relevante de una forma individual como en el caso de la frecuencia, debido a que es relativo con respecto al inicio o fin de eventos o de la toma de registros. Éste enfoque es el más importante debido a que la evaluación no solo estuvo basada en los resultados de clasificación, si no en coeficientes kappa y una evaluación de significancia estadística y la comparación con otro trabajo relevante que resuelve el mismo problema [89]. A

diferencia con Segür y otros enfoques analizados en el capítulo 3, el nuestro alcanza resultados de clasificación del 100 %, con la menor cantidad de características (3) que se han reportado en la literatura revisada.

A lo largo de la investigación, las técnicas fueron más justificadas, como el de la STFT, o LTP que contempla las características de los espectrogramas. El último PEV, no alcanzó los mejores resultados, la exactitud más alta es del 98.5 % como la de EEFGabor, sin embargo, este enfoque considera dos aspectos importantes; recordemos que al igual que FOSP, hace uso de puntos máximos de energía, a pesar de eso, la finalidad de usar DBSCAN es considerar puntos máximos vecinos, con similitud en frecuencia y tiempos cercanos; por otro lado, los puntos máximos, se aseguran que sean los de mayor energía, debido a un preprocesamiento basado en cuartiles, en donde solo los puntos mayores al valor de energía del tercer cuartil (según todos los puntos del espectrograma) hacia arriba son considerados, por lo que puntos de muy baja energía o posible ruido es descartado.

A pesar de que cada uno de los enfoques tiene consideraciones diferentes, los resultados fueron relevantes para todas las bases de datos, tanto para Bonn como para la de la BUAP se alcanzaron resultados superiores al 95 % en la clasificación; con respecto a la de Boston, fueron resultados menores, sin embargo, con respecto a lo que se ha reportado, la clasificación no fue muy diferente (no en más del 10 %) y la evaluación fue más justa en nuestro caso, haciendo uso de validación cruzada y no sólo de un conjunto de entrenamiento y otro de evaluación.

Capítulo 6

Conclusiones

El objetivo principal de este trabajo de investigación es proponer una metodología de extracción de características de señales EEG; esta fue desarrollada bajo el nombre de FOSP y evaluada estadísticamente sobre la exactitud que nos dieron diferentes clasificadores y comparado con lo que se ha reportado en la literatura, obteniendo resultados aceptables.

Se elaboró una breve revisión de los métodos y técnicas utilizadas para el análisis de señales EEG por medio de espectrogramas, logrando identificar diferentes características y número éstas, con diferentes resultados para diversas bases de datos, además de la notoriedad del poco uso de técnicas de aprendizaje profundo para el análisis de EEG pero con resultados aceptables ya reportados.

Se propusieron tres enfoques diferentes, con los que se permitió analizar tres diferentes conjuntos de datos, dos de ellos relacionados con la epilepsia y uno con narcolepsia en ratas, la finalidad fue ver el comportamiento de los enfoques ante señales con distinta naturaleza, a pesar de que se analizaron dos bases de datos de epilepsia el resultado fue distinto, debido a que una contenía datos preprocesados y con una previa selección de canales útiles (la de Bonn), a diferencia de la de Boston que mostraba datos crudos. La de narcolepsia mostró datos crudos también, sin embargo, los canales eran pocos y su recolección fue muy controlada, lo que ayudó a tener resultados aceptables.

El enfoque propuesto EEFGabor, empleó la extracción de energía y entropía como características de la respuesta de los filtros; los vectores de características se evaluaron por medio de MLPs y SVMs. Ambas características se usaron para la evaluación, sin embargo, la energía fue la característica predominante debido a que en la mayoría de los clasificadores fue mejor que la entropía.

En varios trabajos sobre análisis EEG, los valores de los diferentes parámetros de la STFT son fijos, sin embargo, no se justifican. A partir de un breve análisis

teórico y algunos experimentos, se notó que estos parámetros son importantes, ya que afectan al espectrograma. En particular, se debe tener en cuenta el ancho de la ventana y la superposición de la señal. Una buena opción de ventana podría ser Blackman, Hanning, Hanning o Gaussiana, debido a sus propiedades y comportamientos espectrales. La ventana mostró un buen rendimiento según lo reportado en [108] y los resultados mostrados, y los siguientes fueron Blackman y Hanning; la ventana rectangular fue la peor.

Debido a la energía que se muestra en los espectrogramas, probablemente se debería proponer una ventana para cada tipo de clase, ya que las ventanas requieren diferentes longitudes si la energía se muestra con el tiempo o la frecuencia. En el artículo [108], notamos que diferentes parámetros afectan el espectrograma y el rendimiento, por lo tanto, estos deben seleccionarse con un análisis previo de las señales EEG y de las ventanas, un análisis experimental podría no ser suficiente.

En el estudio de la STFT se analizaron los parámetros que afectan la resolución de los espectrogramas: tipo de ventana, longitud y superposición. Se encontró una ventana aceptable (Hanning) con respecto a las características de los lóbulos espectrales a lo largo del tiempo, ya que es similar tanto en tiempo como en frecuencia a su espectro, siendo una variante del coseno alzado; las longitudes encontradas fueron con respecto a la frecuencia de las señales, por lo que se propuso un método para seleccionarlas descrito en la sección 5.2.1; finalmente, se utilizó una superposición, como se ve en la literatura. Estos parámetros permiten aprovechar más la energía de los espectrogramas, la cual es una característica fundamental de la que se basan nuestros enfoques.

Las características de FOSP basadas en LTP, K-means y energía, mostraron un buen rendimiento (99.5 %) con menos características con una dimensión de 7. Por otra parte, se obtuvo el 100 % en Sengür et al. [89], así como con la metodología propuesta (picos mínimos y LTP); la diferencia es que el primero usa un vector de 1280 características por instancia, mientras que el último método usa solo tres características, un pico máximo y dos códigos de LTP. Se han propuesto otros enfoques en la literatura para extraer características de EEG de espectrograma, la mayoría de ellos consideran al menos cuatro características para alcanzar altas precisiones; nuestro enfoque usó solo tres, mostrando la relevancia de las características ya que lograron los mismos resultados.

La metodología FOSP presentada en este trabajo muestra una alternativa para la detección de ataques epilépticos, mostrando la relevancia de esto, debido a que los resultados extraen solo tres características (menos que otros enfoques reportados en la literatura), este enfoque es capaz de generalizar información relevante de un

pequeño subconjunto de características que es principalmente útil para reducir los costos computacionales. Finalmente, el enfoque FOSP puede funcionar como una aplicación clínica siendo utilizado como una herramienta auxiliar para la identificación de un ataque de epilepsia, ya que con la ayuda de expertos que interpretan las señales de EEG podría modelar o reforzar el análisis visual que se da comúnmente. La metodología se puede adaptar a los sistemas de EEG en una clínica para mejorar el rendimiento, solo definiendo el monitoreo del tiempo con la opinión de los médicos.

La última metodología PEV a pesar de no tener lo mejores resultados comparados con FOSP, tiene consideraciones relevantes, debido a que en el desarrollo de ésta se consideran también picos espectrales, a diferencia de FOSP, los picos encontrados en ese enfoque eran durante un instante de tiempo, es decir para todas las frecuencias de una sola columna del espectrograma. Por otro lado, el enfoque basado en DBSCAN considera vecinos con características en común del espectrograma, instantes cercanos, frecuencias parecidas y energías similares y a partir de esos se obtiene el pico, por lo que la característica o pico no es tan general.

6.1. Trabajo a futuro

Como trabajo a futuro, el enfoque de filtros Gabor puede ser analizado por medio de la extracción de otras características, en la literatura se mostró que basados en los filtros Gabor descriptores LBP pueden ser extraídos [109], otros descriptores de este tipo podrían ser aplicados como LTP o histogramas de gradientes orientados.

Por otra parte, para la propuesta de parámetros de la STFT fue una fase inicial, ya que podría considerarse analizar las ventanas con respecto al espectro, porcentaje de energía, lóbulos y otros factores. El porcentaje de energía del lóbulo principal con respecto a los lóbulos laterales podría ser un buen parámetro. Además, se propondrán longitudes de ventana basadas en las frecuencias no solo al doble.

Los tres diferentes enfoques mostraron diferentes resultados, sin embargo, estos pueden analizarse con conjuntos de datos diferentes como ECG, EMG, u otro tipos de bioseñales, debido a la naturaleza no estacionaria podrían ser de utilidad para otro tipo de aplicaciones.

6.2. Publicaciones

Los siguientes trabajos listados abajo, son productos alcanzados durante el desarrollo de este trabajo de investigación.

- Ricardo Ramos, José Arturo Olvera and Ivan Olmos, “Analysis of EEG Signal Processing Techniques based on Spectrograms”, presentado en el 5th International Symposium on Language and Knowledge Engineering y publicado en Research in Computing Science Journal, ISSN 1870-4069 (Indexed by DBLP, LatIndex, Periódica), 2017.
- Ricardo Ramos, José Arturo Olvera, Ivan Olmos and Manuel Martín “Time-Frequency Analysis of EEG Spectrograms using 2-D Gabor Filters for Epileptic Seizure Classification”, presentado en el student’s meeting del 11th Mexican Conference on Pattern Recognition (MCPR-2018) y publicado en Research in Computing Science Journal, ISSN 1870-4069 (Indexed by DBLP, LatIndex, Periódica), 2018.
- Ricardo Ramos, José Arturo Olvera e Ivan Olmos, “Análisis de ElectroEncefalogramas mediante Espectrogramas”, publicado como capítulo de libro en Tópicos actuales en la Ingeniería del lenguaje y del conocimiento, capítulo 6, Fcc-BUAP, Primera Edición, ISBN: 978-607-97282-7-4, 2018.
- Ricardo Ramos, José Arturo Olvera e Ivan Olmos, “Análisis de Tiempo-Frecuencia de espectrogramas EEG por medio de Aprendizaje no Supervisado para la Clasificación de Ataques Epilépticos”, capítulo 10, Avances en Tecnologías del Lenguaje y el Conocimiento, primera edición, ISBN BUAP:978-607-7512-97-2, 2019.
- Ricardo Ramos, José Arturo Olvera, Ivan Olmos y Susana Sánchez, “Parameter analysis on EEG Signal Identification using Short Time Fourier Transform”, student’s meeting del 12th Mexican Conference on Pattern Recognition (MCPR-2019), Research in Computing Science Journal, ISSN 1870-4069 (Indexed by DBLP, LatIndex, Periódica), 2019.
- Ricardo Ramos, José Arturo Olvera, Ivan Olmos y Susana Sánchez, "Feature extraction from EEG spectrograms for epileptic seizure detection", Pattern Recognition Letters 133 (2020) 202–209.
- Ricardo Ramos, José Arturo Olvera e Ivan Olmos, "Transformada de Fourier en tiempo corto para el análisis de espectrogramas", por publicar en Avances en Tecnologías del Lenguaje y el Conocimiento, 2020.

Bibliografía

- [1] M. Mustafa, Mohd Nasir Taib, Z. H. Murat, and Noor Hayatee Abdul Hamid. Gcm texture classification for EEG spectrogram image. In *2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pages 373–376, Nov 2010.
- [2] Rabie A. Ramadan and Athanasios V. Vasilakos. Brain computer interface: control signals review. *Neurocomputing*, 223:26 – 44, 2017.
- [3] J. Satheesh Kumar and P. Bhuvaneshwari. Analysis of electroencephalography (EEG) signals and its categorization: A study. *Procedia Engineering*, 38:2525 – 2536, 2012.
- [4] N. Alamdari, A. Haider, R. Arefin, A. K. Verma, K. Tavakolian, and R. Fazel-Rezai. A review of methods and applications of brain computer interface systems. In *2016 IEEE International Conference on Electro Information Technology (EIT)*, pages 0345–0350, May 2016.
- [5] Gregory A. Light, Lisa E. Williams, Falk Minow, Joyce Sprock, Anthony Rissling, Richard Sharp, Neal R. Swerdlow, and David L. Braff. Electroencephalography (EEG) and event-related potentials (ERPs) with human participants, 2010.
- [6] Marcus Siems, Anna-Antonia Pape, Joerg F. Hipp, and Markus Siegel. Measuring the cortical correlation structure of spontaneous oscillatory activity with EEG and MEG. *NeuroImage*, 129(Supplement C):345 – 355, 2016.
- [7] iMotions. *EEG POCKET GUIDE*. iMotions Biometric Research Platform, 2016.
- [8] Amal Feltane. *Time-Frequency Based Methods for Non-Stationary Signal Analysis with Application To EEG Signals*. PhD thesis, University of Rhode Island, 2016.
- [9] H. A. T. Nguyen, H. H. Tran, T. T. Vu, and T. T. Q. Bui. A combination of independent component analysis, relative wavelet energy, and support vector machine for mental state classification. In *2016 16th International Conference on Control, Automation and Systems (ICCAS)*, pages 733–738, Oct 2016.
- [10] A. Al-Ani, I. Koprinska, G. R. Naik, and R. N. Khushaba. A dynamic channel selection algorithm for the classification of EEG and EMG data. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 4076–4081, July 2016.

-
- [11] W. Ren, M. Han, J. Wang, D. Wang, and T. Li. Efficient feature extraction framework for EEG signals classification. In *2016 Seventh International Conference on Intelligent Control and Information Processing (ICICIP)*, pages 167–172, Dec 2016.
- [12] Y. Li, X. Wang, L. Luo, K. Li, X. Yang, and Q. Guo. Epileptic seizure classification of EEGs using time-frequency analysis based multiscale radial basis functions. *IEEE Journal of Biomedical and Health Informatics*, PP(99):1–1, 2017.
- [13] Sim Kuan Goh, Hussein A. Abbass, Kay Chen Tan, and Abdullah Al Mamun. *Artifact Removal from EEG Using a Multi-objective Independent Component Analysis Model*, pages 570–577. Springer International Publishing, Cham, 2014.
- [14] A. Khan Nabeel and Ali Sadiq. Classification of EEG signals using adaptive time-frequency distributions, 2016.
- [15] Varun Bajaj, Yanhui Guo, Abdulkadir Sengur, Siuly Siuly, and Omer F. Alcin. A hybrid method based on time–frequency images for classification of alcohol and control EEG signals. *Neural Computing and Applications*, pages 1–7, 2016.
- [16] Abdelhakim Ridouh, Daoud Boutana, and Messaoud Benidir. *Comparative Study of Time Frequency Analysis Application on Abnormal EEG Signals*, pages 355–368. Springer International Publishing, Cham, 2017.
- [17] H. H. Kha. Real-time brainwave-controlled interface using p300 component in EEG signal processing. In *2016 IEEE RIVF International Conference on Computing Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, pages 235–240, Nov 2016.
- [18] GFay Boudreaux-Bartels. Mixed time-frequency signal transformations. *Electrical Engineering Handbook*. CRC Press, January 2010.
- [19] W. Wu, S. Nagarajan, and Z. Chen. Bayesian machine learning: EEG/MEG signal processing measurements. *IEEE Signal Processing Magazine*, 33(1):14–36, Jan 2016.
- [20] Vaibhav Gandhi. Chapter 2 - interfacing brain and machine. In Vaibhav Gandhi, editor, *Brain-Computer Interfacing for Assistive Robotics*, pages 7 – 63. Academic Press, San Diego, 2015.
- [21] Desney Tan and Anton Nijholt. *Brain-Computer Interfaces and Human-Computer Interaction*, pages 3–19. Springer London, London, 2010.
- [22] Siuly Siuly, Yan Li, and Yanchun Zhang. *Electroencephalogram (EEG) and Its Background*, pages 3–21. Springer International Publishing, Cham, 2016.
- [23] Outi Väisänen and Jaakko Malmivuo. Improving the snr of eeg generated by deep sources with weighted multielectrode leads. *Journal of Physiology-Paris*, 103(6):306 – 314, 2009. Advanced Methods for the Estimation of Human Brain Activity and Connectivity.

-
- [24] A. Jalilifard and E. B. Pizzolato. An efficient k-nn approach for automatic drowsiness detection using single-channel EEG recording. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 820–824, Aug 2016.
- [25] Leonardo Duque-Muñoz, Jairo Jose Espinosa-Oviedo, and Cesar German Castellanos-Dominguez. Identification and monitoring of brain activity based on stochastic relevance analysis of short-time EEG rhythms. *BioMedical Engineering OnLine*, 13(1):123, 2014.
- [26] P. Kovacs, K. Samiee, and M. Gabbouj. On application of rational discrete short time fourier transform in epileptic seizure classification. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5839–5843, May 2014.
- [27] Christopher M. Sinton. *Narcolepsy-Cataplexy in the Rats*, pages 399–413. Springer International Publishing, Cham, 2016.
- [28] Boualem Boashash, Nabeel Ali Khan, and Taoufik Ben-Jabeur. Time-frequency features for pattern recognition using high-resolution tfds: A tutorial review. *Digital Signal Processing*, 40:1 – 30, 2015.
- [29] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.
- [30] M. Mohammed, M.B. Khan, and E.B.M. Bashier. *Machine Learning: Algorithms and Applications*. CRC Press, 2016.
- [31] Jiawei Han, Micheline Kamber, and Jian Pei. *Advanced Pattern Mining*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, Boston, third edition edition, 2012.
- [32] Sebastian Raschka. *Python Machine Learning*. Packt Publishing, 2015.
- [33] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, Jun 2002.
- [34] Esteban Vazquez-Fernandez, Angel Dacal-Nieto, Fernando Martin, and Soledad Torres-Guijarro. Entropy of Gabor filtering for image quality assessment. In Aurélio Campilho and Mohamed Kamel, editors, *Image Analysis and Recognition*, pages 52–61, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [35] Mohammad Haghghat, Saman Zonouz, and Mohamed Abdel-Mottaleb. Cloudid: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications*, 42(21):7905 – 7916, 2015.
- [36] *Computational Geometry*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

-
- [37] Xiaoyang Tan and Bill Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In S. Kevin Zhou, Wenyi Zhao, Xiaou Tang, and Shaogang Gong, editors, *Analysis and Modeling of Faces and Gestures*, pages 168–182, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [38] A. Satpathy, X. Jiang, and H. Eng. Lbp-based edge-texture features for object recognition. *IEEE Transactions on Image Processing*, 23(5):1953–1964, 2014.
- [39] Nathalie Japkowicz and Mohak Shah. *Performance Measures I*, page 74–110. Cambridge University Press, 2011.
- [40] Jiawei Han, Micheline Kamber, and Jian Pei. 6 - mining frequent patterns, associations, and correlations: Basic concepts and methods. In Jiawei Han, Micheline Kamber, and Jian Pei, editors, *Data Mining (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems, pages 243 – 278. Morgan Kaufmann, Boston, third edition edition, 2012.
- [41] Alejandro Antonio Torres García. *Clasificación de palabras no pronunciadas presentes en Electroencefalogramas (EEG)*. Ciencias computacionales, INAOE, oct 2011.
- [42] R. Benítez, G. Escudero, S. Kanaan, and D.M. Rodó. *Inteligencia artificial avanzada*. Editorial UOC, S.L., 2014.
- [43] JR Landis and GG Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, March 1977.
- [44] Ian H. Witten, Eibe Frank, and Mark A. Hall. Chapter 5 - credibility: Evaluating what’s been learned. In Ian H. Witten, Eibe Frank, and Mark A. Hall, editors, *Data Mining: Practical Machine Learning Tools and Techniques (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems, pages 147 – 187. Morgan Kaufmann, Boston, third edition edition, 2011.
- [45] Salvador García, Alberto Fernández, Julián Luengo, and Francisco Herrera. Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences*, 180(10):2044 – 2064, 2010. Special Issue on Intelligent Distributed Information Systems.
- [46] D.J. Sheskin. *Handbook of Parametric and Nonparametric Statistical Procedures: Third Edition*. CRC Press, 2003.
- [47] M. Z. Ilyas, P. Saad, and M. I. Ahmad. A survey of analysis and classification of EEG signals for brain-computer interfaces. In *2015 2nd International Conference on Biomedical Engineering (ICoBE)*, pages 1–6, March 2015.
- [48] M. Rajya Lakshmi and Chandra Prakash. Survey on EEG signal processing methods. 2014.
- [49] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis. Epileptic seizure detection in EEGs using time–frequency analysis. *IEEE Transactions on Information Technology in Biomedicine*, 13(5):703–710, Sept 2009.

-
- [50] Z. H. Murat M. Mustafa, M. N. Taib and N. Sulaiman. Comparison between KNN and ANN classification in brain balancing application via spectrogram image. *Journal of Computer Science & Computational Mathematics*, 2:17–22, 2012.
- [51] Abdulkadir Şengür, Yanhui Guo, and Yaman Akbulut. Time–frequency texture descriptors of EEG signals for efficient detection of epileptic seizure. *Brain Informatics*, 3(2):101–108, 2016.
- [52] O. Nieves and V. Manian. Automatic person authentication using fewer channel EEG motor imagery. In *2016 World Automation Congress (WAC)*, pages 1–6, July 2016.
- [53] J. Camacho and V. Manian. Real-time single channel EEG motor imagery based brain computer interface. In *2016 World Automation Congress (WAC)*, pages 1–6, July 2016.
- [54] Kai Fu, Jianfeng Qu, Yi Chai, and Yong Dong. Classification of seizure based on the time-frequency image of {EEG} signals using HHT and SVM. *Biomedical Signal Processing and Control*, 13:15 – 22, 2014.
- [55] Siyi Deng, Ramesh Srinivasan, Tom Lappas, and Michael D’Zmura. EEG classification of imagined syllable rhythm using hilbert spectrum methods. *Journal of Neural Engineering*, 7(4):046006, 2010.
- [56] A. Zabidi, W. Mansor, Y. K. Lee, and C. W. N. F. Che Wan Fadzal. Short-time fourier transform analysis of EEG signal generated during imagined writing. In *2012 International Conference on System Engineering and Technology (ICSET)*, pages 1–4, Sept 2012.
- [57] S. K. Bashar, A. R. Hassan, and M. I. H. Bhuiyan. Motor imagery movements classification using multivariate emd and short time fourier transform. In *2015 Annual IEEE India Conference (INDICON)*, pages 1–6, Dec 2015.
- [58] Yi Wang, Guanghua Xu, Sicong Zhang, Ailing Luo, Min Li, and Chengcheng Han. EEG signal co-channel interference suppression based on image dimensionality reduction and permutation entropy. *Signal Processing*, 134:113 – 122, 2017.
- [59] Siddharth Biswal, Joshua Kulas, Haoqi Sun, Balaji Goparaju, M. Brandon Westover, Matt T. Bianchi, and Jimeng Sun. SLEEPNET: automated sleep staging system via deep learning. *CoRR*, abs/1707.08262, 2017.
- [60] Y. Yuan, G. Xun, F. Ma, Q. Suo, H. Xue, K. Jia, and A. Zhang. A novel channel-aware attention framework for multi-channel EEG seizure detection via multi-view deep learning. In *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 206–209, 2018.
- [61] Giulio Ruffini, David Ibañez, Marta Castellano, Laura Dubreuil-Vall, Aureli Soria-Frisch, Ron Postuma, Jean-François Gagnon, and Jacques Montplaisir. Deep learning with EEG spectrograms in rapid eye movement behavior disorder. *Frontiers in Neurology*, 10:806, 2019.

- [62] Grega Vrbancic and Vili Podgorelec. Automatic classification of motor impairment neural disorders from EEG signals using deep convolutional neural networks. *Elektronika ir Elektrotechnika*, 24(4):3–7, Aug. 2018.
- [63] Longhao Yuan and Jianting Cao. Patients’ EEG data analysis via spectrogram image with a convolution neural network. In Ireneusz Czarnowski, Robert J. Howlett, and Lakhmi C. Jain, editors, *Intelligent Decision Technologies 2017*, pages 13–21, Cham, 2018. Springer International Publishing.
- [64] N. K. Ambulkar and S. N. Sharma. Detection of epileptic seizure in EEG signals using window width optimized s-transform and artificial neural networks. pages 1–6, Sep. 2015.
- [65] Ömer F. Alçin, Siuly Siuly, Varun Bajaj, Yanhui Guo, Abdulkadir Sengür, and Yanchun Zhang. Multi-category EEG signal classification developing time-frequency texture features based fisher vector encoding method. *Neurocomputing*, 218:251 – 258, 2016.
- [66] Shibli Nisar, Omar Usman Khan, and Muhammad Tariq. An efficient adaptive window size selection method for improving spectrogram visualization. *Computational Intelligence and Neuroscience*, 2016:6172453, Aug 2016.
- [67] W. A. W Azlan and Y. F. Low. Feature extraction of electroencephalogram (EEG) signal - a review. In *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, pages 801–806, Dec 2014.
- [68] Dean Cvetkovic, Elif Derya Übeyli, and Irena Cosic. Wavelet transform feature extraction from human ppg, ecg, and EEG signal responses to elf pemf exposures: A pilot study. *Digital Signal Processing*, 18(5):861 – 874, 2008.
- [69] Elif Derya Übeyli. Wavelet/mixture of experts network structure for EEG signals classification. *Expert Systems with Applications*, 34(3):1954 – 1962, 2008.
- [70] Grzegorz Rutkowski, Krzysztof Patan, and Paweł Leśniak. Comparison of time-frequency feature extraction methods for EEG signals classification. In Leszek Rutkowski, Marcin Korytkowski, Rafał Scherer, Ryszard Tadeusiewicz, Lotfi A. Zadeh, and Jacek M. Zurada, editors, *Artificial Intelligence and Soft Computing*, pages 320–329, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [71] R.W. Thatcher, D. North, and C. Biver. Eeg and intelligence: Relations between EEG coherence, EEG phase delay and power. *Clinical Neurophysiology*, 116(9):2129 – 2141, 2005.
- [72] Scott Sponheim, Brett Clementz, William Iacono, and Morton Beiser. Clinical and biological concomitants of resting state EEG power abnormalities in schizophrenia. *Biological psychiatry*, 48:1088–97, 01 2001.
- [73] Adrienne L. Tierney, Laurel Gabard-Durnam, Vanessa Vogel-Farley, Helen Tager-Flusberg, and Charles A. Nelson. Developmental trajectories of resting EEG power: An endophenotype of autism spectrum disorder. *PLOS ONE*, 7(6):1–10, 06 2012.

- [74] Yun Park, Lan Luo, Keshab Parhi, and Theoden Netoff. Seizure prediction with spectral power of eeg using cost-sensitive support vector machines. *Epilepsia*, 52:1761–70, 06 2011.
- [75] Martha Ann Bell. Power changes in infant EEG frequency bands during a spatial working memory task. *Psychophysiology*, 39(4):450–458, 2002.
- [76] Davide Moretti, Claudio Babiloni, Giuliano Binetti, Emanuele Cassetta, Gloria Dal Forno, Florinda Ferreri, Raffaele Ferri, Bartolo Lanuzza, Carlo Miniussi, Flavio Nobili, Guido Rodriguez, Serenella Salinari, and Paolo Rossini. Individual analysis of EEG frequency and band power in mild Alzheimer’s disease. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 115:299–308, 03 2004.
- [77] David P Allen and Colum D MacKinnon. Time-frequency analysis of movement-related spectral power in EEG during repetitive movements: a comparison of methods. *Journal of neuroscience methods*, 186(1):107–115, January 2010.
- [78] Thea Radüntz. Signal quality evaluation of emerging EEG devices. *Frontiers in Physiology*, 9:98, 2018.
- [79] Jason T. Moyer, Vadym Gnatkovsky, Tomonori Ono, Jakub Otáhal, Joost Wagenaar, William C. Stacey, Jeffrey Noebels, Akio Ikeda, Kevin Staley, Marco de Curtis, Brian Litt, and Aristeia S. Galanopoulou. Standards for data acquisition and software-based analysis of in vivo electroencephalography recordings from animals. a task1-wg5 report of the aes/ilae translational task force of the ilae. *Epilepsia*, 58(S4):53–67, 2017.
- [80] Boualem Boashash. *Heuristic Formulation of Time-Frequency Distributions*. Academic Press, Oxford, second edition edition, 2016.
- [81] Michael M. Goodwin. *The STFT, Sinusoidal Models, and Speech Modification*, pages 229–258. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [82] William Sethares. *Transforms*, pages 111–145. Springer London, London, 2007.
- [83] S J M Smith. Eeg in the diagnosis, classification, and management of patients with epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(suppl 2):ii2–ii7, 2005.
- [84] Edilberto Amorim, Craig Williamson, Lidia Moura, Mouhsin Shafi, Nicolas Gaspard, Eric Rosenthal, Mary Guanci, Venkatakrishna Rajajee, and M Brandon Westover. Performance of spectrogram-based seizure identification of adult EEGs by critical care nurses and neurophysiologists. *Journal of Clinical Neurophysiology*, 34:1, 12 2016.
- [85] Lingraj Dora, Sanjay Agrawal, Rutuparna Panda, and Ajith Abraham. An evolutionary single Gabor kernel based filter approach to face recognition. *Engineering Applications of Artificial Intelligence*, 62:286 – 301, 2017.

- [86] Fazal Malik and Baharum Baharudin. The statistical quantized histogram texture features analysis for image retrieval based on median and laplacian filters in the dct domain. *Int. Arab J. Inf. Technol.*, 10:616–624, 2013.
- [87] W. Li, K. Mao, H. Zhang, and T. Chai. Selection of Gabor filters for improved texture feature extraction. In *2010 IEEE International Conference on Image Processing*, pages 361–364, 2010.
- [88] Omar S. Al-Kadi. A Gabor filter texture analysis approach for histopathological brain tumor subtype discrimination. *Journal of Science and Technology*, 12(22):1–14, 2017.
- [89] Abdulkadir Şengür, Yanhui Guo, and Yaman Akbulut. Time–frequency texture descriptors of EEG signals for efficient detection of epileptic seizure. *Brain Informatics*, 3(2):101–108, Jun 2016.
- [90] Ayman Abaza and Thirimachos Bourlai. On ear-based human identification in the mid-wave infrared spectrum. *Image and Vision Computing*, 31(9):640 – 648, 2013.
- [91] Ralph G. Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David, and Christian E. Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6):061907, 2001.
- [92] Jose R. Eguibar and Ma. del Carmen Cortés. El mutante de melanina taiep como un modelo de crisis de ausencia. *Gaceta médica de México*, 146(1):11–18, Jan 2010.
- [93] Ali Shoeb and John Guttag. Application of machine learning to epileptic seizure detection. pages 975–982, 08 2010.
- [94] Nicolae Cristian Pampu. Study of the effects of the short time Fourier transform configuration on EEG spectral estimates, Nov 2011.
- [95] V. Gerla, E. Saifutdinova, M. Macas, A. Mladek, and L. Lhotska. P01-comparison of short-time fourier transform and continuous wavelet transform for frequency analysis of sleep EEG. *Clinical Neurophysiology*, 129(4):e14, 2018.
- [96] Julius O. Smith. *Spectral Audio Signal Processing*. <http://ccrma.stanford.edu/~jos/sasp/>, 2011. online book, 2011 edition.
- [97] Carlos Mateo and Juan Antonio Talavera. Short-time fourier transform with the window size fixed in the frequency domain. *Digital Signal Processing*, 77:13 – 21, 2018. Digital Signal Processing and SoftwareX - Joint Special Issue on Reproducible Research in Signal Processing.
- [98] K.M.M. Prabhu. *Window Functions and Their Applications in Signal Processing*. CRC Press, 2013.
- [99] Julie E. Greenberg and Natalie T. Smith. Spectral Analysis, Aug 2011. [Online; accessed 20. Nov. 2018].

-
- [100] CP Panayiotopoulos. *Idiopathic Generalised Epilepsies*. Bladon Medical Publishing, 2005.
- [101] Julius O. Smith III and Xavier Serra. PARSHL: An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation¹, Dec 2005. [Online; accessed 21. Nov. 2018].
- [102] P. Kovacs, K. Samiee, and M. Gabbouj. On application of rational Discrete Short Time Fourier transform in epileptic seizure classification. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5839–5843, May 2014.
- [103] E. T. Pereira and H. Martins Gomes. The role of data balancing for emotion classification using EEG signals. In *2016 IEEE International Conference on Digital Signal Processing (DSP)*, pages 555–559, 2016.
- [104] Ricardo Ramos, José Arturo Olvera-López, and Iván Olmos. Analysis of EEG signal processing techniques based on spectrograms. *Research in Computing Science*, 145:151–162, 2017.
- [105] Zeldá B. Zabinsky. *Introduction*, pages 1–23. Springer US, Boston, MA, 2003.
- [106] Iaroslav Omelianenko. Applying advanced machine learning models to classify electro-physiological activity of human brain for use in biometric identification. *CoRR*, 1708.01167, 2017.
- [107] Chih-Sheng Huang, Chun-Ling Lin, Li-Wei Ko, Shen-Yi Liu, Tung-Ping Su, and Chin-Teng Lin. Knowledge-based identification of sleep stages based on two forehead electroencephalogram channels. *Frontiers in Neuroscience*, 8:263, 2014.
- [108] Ricardo Ramos, José Arturo Olvera-López, Iván Olmos, and Susana Sánchez-Urrieta. Parameter experimentation for epileptic seizure detection in EEG signals using Short-Time Fourier transform. *Research in Computing Science*, 148:83–96, 2019.
- [109] T. Sunil Kumar, Vivek Kanhangad, and Ram Bilas Pachori. Classification of seizure and seizure-free EEG signals using local binary patterns. *Biomedical Signal Processing and Control*, 15:33 – 40, 2015.