



UNIVERSIDAD PONTIFICIA BOLIVARIANA
ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

RECONOCIMIENTO DE ACTIVIDADES HUMANAS USANDO REDES DE
GRAN MEMORIA DE CORTO PLAZO

Javier Fernando Villamizar Torres
Paula Andrea Lizarazo Sanabria

PROYECTO DE GRADO

Bucaramanga
2019



UNIVERSIDAD PONTIFICIA BOLIVARIANA
ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

**RECONOCIMIENTO DE ACTIVIDADES HUMANAS USANDO REDES DE
GRAN MEMORIA DE CORTO PLAZO**

Javier Fernando Villamizar Torres
Paula Andrea Lizarazo Sanabria

PROYECTO DE GRADO

Presentado ante la Universidad Pontificia Bolivariana
como requisito para optar al título de

INGENIERO ELECTRÓNICO

Realizado con la asesoría de:
Miguel Alfonso Altuve Paredes, PhD.

**Bucaramanga
2019**

DEDICATORIA

En primera instancia agradezco a mi familia por su gran apoyo en toda mi vida estudiantil, por darme ánimos de continuar con mis sueños sin importar los obstáculos y por inculcarme grandes valores y cualidades que definen la persona que soy ahora.

A Nelson Cala por enseñarme a ser mejor cada día, dándome fuerzas para continuar en mis proyectos y por ser una persona incondicional, haciéndome ver los momentos difíciles en posibles de solucionar con tal de verme triunfar.

Quisiera agradecer a mis profesores que estuvieron conmigo en mi formación académica y profesional, en especial al profesor Miguel Altuve quién siempre confió desde el principio en mis capacidades y me guio a ingresar al semillero BISEMIC, y por medio del semillero incentivarne a aprender cosas nuevas; también agradezco su dedicación y tiempo que tuvo con nosotros.

“Todos nuestros sueños se pueden hacer realidad si tenemos el coraje de perseguirlos.”

— Walt Disney

Paula Andrea Lizarazo Sanabria

AGRADECIMIENTOS

Al Dr. Miguel Altuve, por la confianza depositada en nosotros, por estar constantemente brindándonos su conocimiento, experiencia y consejos, los cuales nos han ayudado a aventurar científica y académicamente, obteniendo buenos resultados y momentos gratificantes, y finalmente por la guía, paciencia y seguimiento en todos los trabajos y actividades que hemos llevado a cabo.

A los Ingenieros Nelson Monroy, Luis Suárez y Jeyson Ardila, por su disposición en todo momento para aclararnos las dudas y por su agradable compañía en este tiempo.

A todos los profesores, desde la escuela hasta la universidad, quienes han aportado en gran medida a nuestra educación y conocimiento, haciéndonos mejores profesionales, y también por aquellos, unos más que otros, que nos han dejado huella y nos han hecho crecer más como personas.

Para finalizar, a nuestras familias quienes han sido el pilar de estos acontecimientos, gracias a su esfuerzo, dedicación, enseñanzas y apoyo constante, estando en todo momento.

ÍNDICE GENERAL

INTRODUCCIÓN	15
CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA	18
1.1 Definición del problema	18
1.2 Objetivos	19
1.2.1 Objetivo general	19
1.2.2 Objetivos específicos	19
1.3 Justificación e importancia	20
1.4 Impacto esperado	20
CAPÍTULO 2: MARCO TEÓRICO	22
2.1 Reconocimiento de actividades humanas	22
2.1.1 Señales usadas en el HAR	23
2.1.2 Sensores inerciales	24
2.2 Representación de señales en tiempo-frecuencia	25
2.2.1 Descomposición empírica en modos	26
2.2.2 Descomposición empírica en modos por conjunto completa con ruido adaptativo mejorada	27
2.3 Métricas para analizar las señales	29
2.3.1 Promedio	29
2.3.2 Varianza	30
2.3.3 Desviación estándar	30
2.3.4 Entropía de Shannon	30
2.4 Aprendizaje automático y aprendizaje profundo	31
2.4.1 Redes neuronales artificiales	32
2.4.2 Redes neuronales recurrentes	33
2.4.3 Medidas de desempeño	37
CAPÍTULO 3: ANTECEDENTES	39
3.1 Técnicas basadas en umbrales	39
3.2 Técnicas basadas en aprendizaje automático	40
CAPÍTULO 4: MARCO METODOLÓGICO	43
4.1 Materiales e instrumentos	43
4.1.1 Descripción del hardware y software	43
4.1.2 Conjunto de datos	44

4.2	Análisis de las señales inerciales	44
4.2.1	Descomposición de las señales inerciales	44
4.2.2	Estimación del rango representativo de frecuencias en cada IMF	47
4.2.3	Caracterización de las señales inerciales en cada IMF	48
4.3	Clasificación de las actividades humanas	48
4.3.1	Representación en tiempo-frecuencia	48
4.3.2	Estandarización	50
4.3.3	Segmentación	50
4.3.4	Clasificación	51
CAPÍTULO 5: RESULTADOS		53
5.1	Análisis de las señales inerciales	53
5.1.1	Análisis de las señales inerciales crudas	53
5.1.2	Análisis de las IMF de las señales inerciales	54
5.2	Clasificación de las actividades humanas	72
5.2.1	Representación en tiempo-frecuencia	72
5.2.2	Segmentación	76
5.2.3	Clasificación	77
CAPÍTULO 6: CONCLUSIONES Y RECOMENDACIONES		80
BIBLIOGRAFÍA		84
ANEXO A: ARTÍCULO PRESENTADO EN UN EVENTO INTERNACIONAL		89
ANEXO B: ARTÍCULO ENVIADO A UNA REVISTA		92

ÍNDICE DE FIGURAS

Figura 1.	Smartphone	23
Figura 2.	Acelerometro	24
Figura 3.	Giroscopio	25
Figura 4.	Procedimiento de sifting	27
Figura 5.	Redes neuronales artificiales	33
Figura 6.	Redes Neuronales Artificiales	34
Figura 7.	Redes Neuronales Artificiales	34
Figura 8.	Redes Neuronales Artificiales	35
Figura 9.	Redes Neuronales Artificiales	35
Figura 10.	Bloque de una red LSTM.	36
Figura 11.	Ejemplo de señales de los sensores inerciales (acelerómetro en la columna izquierda, giroscopio en la columna central y aceleración total en la columna derecha) en cada eje (x en la primera fila, y en la fila central, y z en la ultima fila), para las actividades dinámicas (WK en azul, WU en rojo y WU en amarillo).	45
Figura 12.	Ejemplo de señales de los sensores inerciales (acelerómetro en la columna izquierda, giroscopio en la columna central y aceleración total en la columna derecha) en cada eje (x en la primera fila, y en la fila central, y z en la ultima fila), para las actividades estáticas (ST en azul, SD en rojo y LD en amarillo)..	46
Figura 13.	Diagrama de bloques para la caracterización de las señales inerciales . .	47
Figura 14.	Procedimiento propuesto para reconocer las actividades humanas. . . .	49
Figura 15.	Arquitectura del modelo de clasificación propuesto..	51
Figura 16.	Barras de error de las varianzas de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del acelerómetro en el eje z para cada actividad.	64
Figura 17.	Barras de error de las varianzas de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del giroscopio en el eje z para cada actividad.	65
Figura 18.	Barras de error de las entropías de Shannon de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del acelerómetro en el eje x para cada actividad.	66

Figura 19.	Barras de error de las entropías de Shannon de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del giroscopio en el eje x para cada actividad.	67
Figura 20.	Barras de error de f_L del giroscopio en el eje z	72
Figura 21.	Barras de error de f_H del giroscopio en el eje z	73
Figura 22.	Descomposición en seis IMF, por medio de la técnica ICEEMDAN, de una señal del acelerómetro en el eje x	74
Figura 23.	Ejemplo de representación de una señal del acelerómetro en el eje x en cinco IMF. Las IMF están estandarizadas.	75
Figura 24.	Segmentación en ventanas de 64 muestras de la IMF 1 de una señal del acelerómetro en el eje x	76

ÍNDICE DE TABLAS

Tabla 1.	Matriz de confusión.	38
Tabla 2.	Promedio \pm desviación estándar de los vectores de entropía (\vec{H}) y varianza ($\vec{\sigma}^2$, expresadas en $\times 10^{-3}$) de las señales inerciales (acelerómetro A y giroscopio G) crudas (sin procesar) en cada eje (x , y y z), por cada clase (WK, WU, WD, ST, SD y LD).	54
Tabla 3.	Cantidad de registros que pueden ser descompuestos en un número máximo de IMF, tanto del conjunto de entrenamiento como del de prueba, usando las técnicas EMD e ICEEMDAN.	55
Tabla 4.	Varianza de las IMF del acelerómetro obtenidas con la técnica EMD. Los resultados están expresados en $\times 10^{-3}$	56
Tabla 5.	Varianza de las IMF del giroscopio obtenidas con la técnica EMD. Los resultados están expresados en $\times 10^{-3}$	57
Tabla 6.	Entropía de Shannon de las IMF del acelerómetro obtenidas con la técnica EMD.	58
Tabla 7.	Entropía de Shannon de las IMF del giroscopio obtenidas con la técnica EMD	59
Tabla 8.	Varianza de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN. Los resultados están expresados en $\times 10^{-3}$	60
Tabla 9.	Varianza de las IMF del giroscopio obtenidas con la técnica ICEEMDAN. Los resultados están expresados en $\times 10^{-3}$	61
Tabla 10.	Entropía de Shannon de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN.	62
Tabla 11.	Entropía de Shannon de las IMF del giroscopio obtenidas con la técnica ICEEMDAN.	63
Tabla 12.	Rango de frecuencia de las IMF del acelerómetro obtenidas con la técnica EMD.	68
Tabla 13.	Rango de frecuencia de las IMF del giroscopio obtenidas con la técnica EMD.	69
Tabla 14.	Rango de frecuencia de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN.	70
Tabla 15.	Rango de frecuencia de las IMF del giroscopio obtenidas con la técnica ICEEMDAN.	71
Tabla 16.	Cantidad de registros que pueden descomponerse hasta el máximo número de IMF indicado usando la técnica ICEEMDAN.	74
Tabla 17.	Matriz de confusión del primer modelo de clasificación.	77
Tabla 18.	Matriz de confusión del segundo modelo de clasificación.	78

Tabla 19. Matriz de confusión del modelo de clasificación de las seis actividades humanas
79

LISTA DE ACRÓNIMOS Y ABREVIACIONES

Acc	Exactitud, del inglés <i>Accuracy</i>
ANN	Redes Neuronales Artificiales, por sus siglas en inglés <i>Artificial Neural Networks</i>
CEEMDAN	Descomposición Empírica en Modos por Conjunto Completo con Ruido Adaptativo, por sus siglas en inglés <i>Complete Ensemble Empirical Mode Decomposition with Adaptive Noise</i>
EEMD	Descomposición Empírica en Modos por Conjuntos, por sus siglas en inglés <i>Ensemble Empirical Mode Decomposition</i>
EMD	Descomposición Empírica en Modos, por sus siglas en inglés <i>Empirical Mode Decomposition</i>
FFT	Transformada Rápida de Fourier, por sus siglas en inglés <i>Fast Fourier Transform</i>
HAR	Reconocimiento de la actividad Humana, por sus siglas en inglés <i>Human Activity Recognition</i>
HMM	Modelo Oculto de Markov, por sus siglas en inglés <i>Hidden Markov Model</i>
ICEEMDAN	Descomposición Empírica en Modos por Conjunto Completa con Ruido Adaptativo Mejorada, por sus siglas en inglés <i>Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise</i>
IMF	Función de Modo Intrínseco, por sus siglas en inglés <i>Intrinsic Mode Function</i>
kNN	k-vecinos más cercanos, por sus siglas en inglés <i>k-nearest Neighbors</i>
LD	Estar Acostado, del inglés <i>Laying Down</i>
LSTM	Redes de Gran Memoria de Corto Plazo, por sus siglas en inglés <i>Long Short-Term Memory</i>
MaxIter	Límite de iteraciones para el proceso de cribado o <i>sifting</i> de la técnica ICEEMDAN
MEMS	Sistema Microelectromecánico, por sus siglas en inglés <i>microelectromechanical systems</i>

Nstd	Amplitud del ruido blanco a ser utilizado por la técnica ICEEMDAN
Pre	Precisión, del inglés <i>Precision</i>
PSD	Densidad Espectral de Potencia, por sus siglas en inglés <i>Power Spectral Density</i>
Rec	Exhaustividad, del inglés <i>Recall</i> , también llamada sensibilidad
RNN	Redes Neuronales Recurrentes, por sus siglas en inglés <i>Recurrent Neural Networks</i>
SD	Estar Parado, del inglés <i>Standing</i>
ST	Estar Sentado, del inglés <i>Sitting</i>
SVM	Máquina de Soporte Vectorial, por sus siglas en inglés <i>Support Vector Machine</i>
WK	Caminar, del inglés <i>Walking</i>
WU	Bajar Escaleras, del inglés <i>Walking Down</i>
WU	Subir Escaleras, del inglés <i>Walking Up</i>

RESUMEN GENERAL DE TRABAJO DE GRADO

TITULO: RECONOCIMIENTO DE ACTIVIDADES HUMANAS USANDO REDES DE GRAN MEMORIA DE CORTO PLAZO

AUTOR(ES): Javier Fernando Villamizar Torres
Paula Andrea Lizarazo Sanabria

PROGRAMA: Facultad de Ingeniería Electrónica

DIRECTOR(A): Miguel Alfonso Altuve Paredes

RESUMEN

El reconocimiento de actividades humanas es un tema de mucha relevancia debido a su amplia gama de aplicaciones. Diferentes enfoques han sido propuestos para reconocer actividades humanas, desde la comparación de señales con umbrales hasta la aplicación de técnicas de aprendizaje automático y profundo. En este trabajo se realiza la clasificación de seis actividades humanas usando redes LSTM bidireccionales y, como representación de las señales inerciales a ser explotadas por estas redes, se propone la representación de estas señales en funciones de modo intrínseco (IMF) usando la descomposición empírica en modos por conjuntos completa con ruido adaptativo mejorada (ICEEMDAN). Las actividades clasificadas fueron: caminar, subir escaleras, bajar escaleras, estar sentado, estar parado y estar acostado. Los registros con las señales inerciales (acelerómetro y giroscopio) de 2.56 segundos (128 muestras) fueron recolectados por un teléfono inteligente a 30 voluntarios y se encuentran disponibles en el UCI Machine Learning Repository. En primer lugar, las señales inerciales fueron estandarizadas para llevarlas a una misma escala y fueron descompuestas en IMF usando la ICEEMDAN. Posteriormente, las IMF fueron segmentadas (divididas) en nueve segmentos de 64 muestras con solapamiento de 12.5% e introducidos a una primera red con cuatro salidas para identificar las actividades dinámicas y las estáticas como una única clase denominada "statics", arrojando una exactitud del 98.86%. Luego, las IMF sin segmentar de los registros asignados a la clase statics fueron introducidos a una segunda red para clasificar sus tres actividades, dando una exactitud de 88.46%. En total se obtuvo 92.91% de exactitud para clasificar las seis actividades humanas. Este desempeño se debe a que la ICEEMAN nos permitió extraer información que se encontraba implícita en la señal, además la segmentación de las señales permitió que la red pudiera discriminar entre actividades estáticas y dinámicas.

PALABRAS CLAVE:

Sensores inerciales, Descomposición Empírica en modos, LSTM, Reconocimiento de patrones

V° B° DIRECTOR DE TRABAJO DE GRADO

GENERAL SUMMARY OF WORK OF GRADE

TITLE: HUMAN ACTIVITY RECOGNITION USING LONG SHORT-TERM MEMORY NETWORKS

AUTHOR(S): Javier Fernando Villamizar Torres
Paula Andrea Lizarazo Sanabria

FACULTY: Facultad de Ingeniería Electrónica

DIRECTOR: Miguel Alfonso Altuve Paredes

ABSTRACT

The recognition of human activities is a topic of great relevance due to its wide range of applications. Different approaches have been proposed to recognize human activities, ranging from the comparison of the signals with thresholds to the application of deep and machine learning techniques. In this work, the classification of six human activities is performed using bidirectional LSTM networks and, as a representation of the inertial signals to be exploited by these networks, the representation of these signals in intrinsic mode functions (IMF) is proposed using improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN). The classified activities were: walking, walking downstairs, walking upstairs, standing, sitting, and lying down. Records with inertial signals (accelerometer and gyroscope) of 2.56 seconds (128 samples) were collected by a smartphone to 30 volunteers and are available in the UCI Machine Learning Repository. First, inertial signals were standardized to take them to the same scale and were decomposed into IMF using ICEEMDAN. IMF were then segmented (split) into nine segments of 64 samples with 12.5% overlap and introduced to a first network with four outputs to identify the dynamic activities and the statics as a single class called "statics", giving 98.86% accuracy. Then, the non-segmented IMF of the records assigned to the statics class were introduced to a second network to classify their three activities, giving an accuracy of 88.46%. In total, 92.91% accuracy was obtained to classify the six human activities. This performance is because ICEEMAN allowed the extraction of information that was embedded in the signal, and the segmentation of the signals allowed the network to discriminate between static and dynamic activities.

KEYWORDS:

Inertial sensors, Empirical Mode Decomposition, LSTM, Pattern Recognition

V° B° DIRECTOR OF GRADUATE WORK

INTRODUCCIÓN

La tecnología nos ha brindado muchos beneficios y facilidades que hacen que nuestro diario vivir sea más llevadero, ya sea por cuestiones de desplazamiento como puede ser un automóvil o en algo tan esencial para las personas como lo son las comunicaciones. No obstante, los avances tecnológicos han incluso ayudado en la solución de las dos anteriores necesidades con el dispositivo móvil. Gracias a este equipo se puede obtener un servicio o bien fácilmente, incluso sin salir del hogar, como la compra de alimentos y ropa, pagar algún servicio contratado o hasta pedir transporte público. Estos equipos contienen diversos sensores que pueden ser usados para medir, de una manera cuantitativa, nuestras acciones, lo cual abre muchas posibilidades para diferentes estudios, como alertar el sedentarismo (Sun, Zhang, Li, Guo, y Li, 2010), malas posturas (Guo y Miao, 2006), entre otros. Por estos motivos, el reconocimiento de las actividades humanas (HAR, por sus siglas en inglés *Human Activity Recognition*) se ha convertido en un área de mucho interés últimamente, ya que permite monitorear a las personas y presentar soluciones a diferentes necesidades.

Haciendo uso de estos dispositivos es posible recolectar información que describe nuestros comportamientos (Bulling, Blanke, y Schiele, 2014), desde adquisición de imágenes o vídeo hasta simplemente series de tiempo (señales), ya sean obtenidas por medio de cámaras (Lin, Sun, Poovendran, y Zhang, 2008) o teléfonos inteligentes. Además, dado que el teléfono celular se ha vuelto una herramienta indispensable que portamos constantemente, es capaz de adquirir información a partir de las señales de los sensores que contiene, sin que nos demos cuenta de ello, y, a partir de un procesamiento de estas señales, presentarle al usuario información a través de su pantalla.

El HAR consiste en reconocer acciones u objetivos de uno o más sujetos partiendo de una serie de observaciones realizadas a estos sujetos y a su entorno. Este campo de la investigación ha llamado la atención por su fuerte impacto en el apoyo de diferentes aplicaciones y su relación con muchos campos de estudio. En primer lugar, el HAR permite una gran variedad de aplicaciones, ya sea en domótica (encender las luces al hacer un movimiento de la mano) (Chahuara, Fleury, Portet, y Vacher, 2012), en deportes (encontrar el patrón correcto para efectuar el mejor lanzamiento de una pelota) (Piergiovanni y Ryoo, 2018), y en medicina (evaluar y hacer seguimiento de algún procedimiento quirúrgico) (Laverde, Rueda,

Amado, Rojas, y Altuve, 2018). En segundo lugar, se encuentran campos de estudio como la interacción humano-computador, ontología (Chen y Nugent, 2009), entre otros.

En este trabajo se propone el reconocimiento de actividades humanas a partir del procesamiento de señales inerciales provenientes de un teléfono celular inteligente (textitsmartphone), tomadas por dos sensores que se encuentran embebidos en éste: acelerómetro y giroscopio en tres ejes. Las señales fueron tomadas de una base de datos disponibles en la página web de UCI Machine Learning Repository ¹. Para esto se propone la representación de las señales inerciales en el dominio tiempo-frecuencia usando la descomposición empírica en modos (EMD, por sus siglas en inglés *Empirical Mode Decomposition*) para encontrar características relevantes que no son fácilmente observadas en la señales originales sino a través de su representación en otro dominio. La técnica EMD descompone una señal en funciones de modo intrínseco (IMF, por sus siglas en inglés, *Intrinsic Mode Functions*), componentes oscilatorios de duración finita, pero de frecuencia variables, ordenados de forma decreciente de acuerdo con su contenido frecuencial (las primeras IMF contienen las altas frecuencias mientras que las últimas IMF contienen las bajas frecuencias), donde la suma de todas las IMF permite reconstruir la señal original. Además, una técnica recientemente derivada de la EMD que presenta una descomposición más robusta denominada descomposición empírica en modos por conjunto completa con ruido adaptativo mejorada (ICEEMDAN, por sus siglas en inglés *Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise*), fue también implementada con el objetivo de obtener otra representación en tiempo-frecuencia. Luego, la representación de las señales en tiempo-frecuencia fueron usadas como entradas a una red de tipo profundo, denominada red de gran memoria de corto plazo (LSTM, por sus siglas en inglés *Long Short-Term Memory*), con el objetivo de clasificar automáticamente seis actividades humanas, tres dinámicas, caminar (WK, del inglés *walking*), subir escaleras (WU, del inglés *walking up*) y bajar escaleras (WD, del inglés *walking down*), y tres estáticas, estar sentado (ST, del inglés *sitting*), estar parado (SD, del inglés *standing*), y estar acostado (LD, del inglés *laying down*). Este trabajo se basa en la hipótesis de que la información contenida en las señales del acelerómetro y del giroscopio, adquirida en tres ejes (x, y, z) es resaltada cuando se descompone en IMF, y puede ser mejor explotada por una máquina de aprendizaje del tipo recurrente (especialmente concebida para procesar series de tiempo).

Dado el impacto de los resultados obtenidos en este trabajo, se ha enviado un artículo de investigación al evento LXIX Convención Anual de AsoVAC, a realizarse en Caracas, Venezuela, del 20 al 22 de noviembre de 2019, el cual ha sido aceptado por los organizadores del evento, y se ha enviado un artículo de investigación a una revista de alto impacto para

¹<https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

su evaluación por parte de expertos en el área y consideración para publicación.

El resto de este documento se encuentra estructurado de la siguiente manera:

- En el capítulo 1 se expone el planteamiento del problema. Primero se describe el problema, luego se muestran los objetivos generales y específicos del trabajo, y se dan a conocer la justificación del trabajo de investigación y el impacto esperado de los resultados obtenidos.
- El capítulo 2 se dedica a dar las bases teóricas para comprender el trabajo realizado. Se presentan los sensores y señales utilizadas en el HAR, la representación de señales en el dominio tiempo-frecuencia, las métricas empleadas para analizar señales, y, por último, las técnicas de clasificación de aprendizaje automático supervisado empleadas.
- Los referentes para el desarrollo de este trabajo son mostrados en el capítulo 3, en donde se detallan los métodos y estrategias empleadas por otros autores para realizar las representaciones de las señales y las técnicas de clasificación de actividades humanas.
- En el capítulo 4 se exhibe la metodología llevada a cabo en este trabajo de investigación. Se describe exhaustivamente los procedimientos realizados en las diferentes etapas del proyecto: descripción de los datos, análisis de las señales inerciales, y la clasificación automática de las actividades humanas. Con esto nos aseguramos de que nuestros experimentos puedan ser reproducidos por otros investigadores y puedan obtener resultados similares a los mostrados en este documento.
- Los resultados del proyecto se detallan en el capítulo 5. En este capítulo se presenta, además, un análisis de dichos resultados, resaltando no solamente la interpretación de los hallazgos, la comparación con investigaciones existentes y el impacto en el estado del arte, sino la explicación de la fenomenología, relacionando los resultados obtenidos con el comportamiento humano.
- Para finalizar son presentadas las conclusiones del trabajo realizado e investigaciones que se pudieran derivar a futuro por otros investigadores en el capítulo 6.

CAPÍTULO 1

PLANTEAMIENTO DEL PROBLEMA

1.1. Definición del problema

Se sabe que, para tener control de algo, conocer cómo funciona o para monitorear el estado en el que se encuentra y poder obtener beneficio de ello, es necesario medir ese proceso o, en el caso de este trabajo, medir las acciones que realiza el ser humano. Pero las mediciones que se realizan necesitan una respectiva interpretación para que esta toma de datos surta efecto: de nada sirve medir la temperatura de una habitación si no sabemos si está haciendo más calor o frío del que debe. Entonces, se presentan dos retos: la adquisición y la interpretación de estas mediciones.

En la actualidad se han desarrollado varios sensores, entre ellos se encuentran los acelerómetros y los giroscopios en todo tipo de tamaños, y que incluso se encuentran embebidos en otros dispositivos, lo que soluciona la parte de la adquisición, hasta tal punto que se encuentran grandes repositorios de base datos, dando aún más facilidades para la investigación y el desarrollo. Quedando la parte de la interpretación, teniendo en cuenta que el ser humano posee la capacidad de realizar diversas tareas, a diferencia de otros seres vivos. Estas tareas conllevan a acciones con mayor complejidad por lo que se hace necesario un alto grado de interpretación para describir adecuadamente las acciones que realizamos.

Lo anterior representa un reto, que en estos días está tomando mucha relevancia debido a las aplicaciones en las que puede servir y los resultados que se pueden lograr: proporcionar un ambiente inteligente (domótica) (Chahuara y cols., 2012), encontrar la manera de mejorar el lanzamiento de un balón (deporte) (Piergiovanni y Ryoo, 2018) y monitorear procedimientos quirúrgicos (medicina) (Laverde y cols., 2018). Aunque existen diversas alternativas para obtener información, desde cámaras de vídeo hasta objetos con identificación por radiofrecuencia, sensores portátiles, o sensores embebidos en otros dispositivos como los teléfonos inteligentes, el objetivo es reconocer patrones en los datos brindados por éstos.

En la literatura se encuentran diferentes métodos para lograr reconocer los patrones

de las actividades llevadas a cabo por los seres humanos, pero muchos de estos métodos utilizan los datos crudos (Altun y Barshan, 2010), es decir, sin procesar, otros usan ciertos filtros para reducir el ruido, sin embargo, son muy pocos los enfoques que realizan una representación de los datos en otros dominios (Z. Wang, Wu, Chen, Ghoneim, y Hossain, 2016) o la extracción de características relevantes de estos. Estas características relevantes o la representación en otros dominios ayudan a que la interpretación de los datos realizadas por los métodos de clasificación, como los encontrados en el campo del aprendizaje automático, obtengan un mejor desempeño. Por eso, la implementación de nuevos métodos que ayuden al reconocimiento de las actividades humanas es un reto que se puede abordar y representa un gran desafío.

Con base en lo descrito anteriormente, en este trabajo se plantea la siguiente hipótesis: ¿Si se representaran las señales inerciales en otros dominios y estas representaciones se usaran como entrada a una red LSTM, sería posible el reconocimiento de ciertas actividades humanas? Esta hipótesis se basa en el hecho de que la descomposición de una señal ya sea en espectrograma, en transformada de ondículas o en funciones de modo intrínseco, representarían mejor las características que se encuentran embebidas en las señales, debido a que estas representaciones permiten aumentar la observabilidad del sistema.

1.2. Objetivos

Para cumplir con la hipótesis planteada se proponen los siguientes objetivos:

1.2.1. Objetivo general

Desarrollar un clasificador basado en redes LSTM que utilice como entrada una representación de señales adquiridas de un teléfono inteligente con el fin de identificar seis actividades humanas.

1.2.2. Objetivos específicos

1. Describir los datos usando un análisis estadístico para obtener información de las señales adquiridas
2. Representar las señales adquiridas en otros dominios para incrementar la observabilidad del sistema

3. Emplear las representaciones de las señales en otros dominios como entradas a una red de clasificación de tipo profunda para identificar automáticamente las actividades humanas.

1.3. Justificación e importancia

El reconocimiento de patrones en las actividades realizadas por los seres humanos se ha vuelto un tema de mucha relevancia en la actualidad. Debido a los constantes desarrollos tecnológicos, el HAR puede facilitar, mejorar y complementar el desarrollo de herramientas y técnicas para el avance y evolución en muchos campos. Por ejemplo, en el campo de la medicina abarcaría tanto el monitoreo y seguimiento de actividades que requieran de mucha práctica hasta la creación de máquinas que obedecen al especialista mientras que él se encuentre en un lugar distante (telemonitoreo) (Satava, 2006). Para que esto sea llevado a cabo es necesario la creación de algoritmos que sean capaces de reconocer, según el enfoque, las diversas actividades que podemos realizar los seres humanos, haciendo uso de las observaciones recolectadas por los diversos dispositivos que se encuentran hoy en día.

A pesar de los diferentes campos en los que se puede indagar, el objetivo de este trabajo solo se enfoca en reconocer seis actividades humanas: estar sentado, parado, acostado, caminando, subiendo escaleras, y bajando escaleras, y pretende ser un referente en el campo de clasificación automática basada en la representación de los datos en un dominio tiempo-frecuencia.

Este trabajo sigue los esfuerzos llevados a cabo en una investigación que se realizó recientemente en el grupo de investigación BISEMIC de la UPB, en las líneas de aprendizaje automático y procesamiento de señales. En esta ocasión se enfoca en la implementación de una representación tiempo-frecuencia de las señales inerciales provenientes de un teléfono celular inteligente que es explotada por una red neuronal de aprendizaje profundo del tipo LSTM bidireccional.

1.4. Impacto esperado

En este trabajo, por medio de señales inerciales tomadas de un dispositivo de uso común sin que el usuario lo note, se reconocen de manera automática seis actividades que podemos realizar los seres humanos. Reconocer automáticamente actividades humanas podría ayudar en la interacción entre seres humanos y en sus relaciones interpersonales ya que proporciona información sobre la identidad de una persona, su personalidad y su estado

psicológico (Vrigkas, Nikou, y Kakadiaris, 2015). La capacidad que tienen los seres humanos de reconocer las actividades que realiza otra persona es extensamente estudiada en el campo de la visión por computador y en el aprendizaje automático. El reconocimiento de actividades humanas podría contribuir a mejorar nuestra calidad de vida, por ejemplo, en domótica para brindar servicio, comodidad y ayuda de manera automática, o en un ambiente clínico para monitorear pacientes y poderles brindar soluciones ajustadas a sus necesidades. Además, el reconocimiento de actividades humanas podría ser utilizado en sistemas de videovigilancia, en la interacción humano-máquina y en la robótica para la caracterización del comportamiento humano.

Aunque se trabajará con señales adquiridas en condiciones adecuadas dentro de un laboratorio, aceptadas por la comunidad científica para el desarrollo de algoritmos para el aprendizaje automático, en un futuro se podría validar el enfoque propuesto en un ambiente real.

CAPÍTULO 2

MARCO TEÓRICO

A continuación se presenta la fundamentación teórica de este documento de tesis. Se hace particular interés en la representación tiempo-frecuencia de señales y en la clasificación automática usando algoritmos de aprendizaje profundo. Dado que este trabajo está enmarcado en una línea de investigación consolidada del grupo de investigación y del semillero BISEMIC, gran parte de los fundamentos teóricos concernientes a la representación tiempo-frecuencia de señales y a la clasificación automática usando algoritmos de aprendizaje profundo ha sido tomada de la tesis de pregrado (Suárez Mantilla y Ardila Navas, 2019).

2.1. Reconocimiento de actividades humanas

Reconocer actividades humanas consiste en interpretar los gestos o movimientos del ser humano por medio de sensores con el fin de determinar la acción o actividad humana (Ann y Theng, 2014). Por ejemplo, un sistema HAR puede reportar las actividades realizadas por pacientes fuera de las instalaciones hospitalarias, lo que lo convierte en una herramienta útil para la evaluación de intervenciones sanitarias y progresos de terapias, y en la toma de decisiones clínicas (Qi y cols., 2019). El HAR puede ser supervisado o no supervisado. El sistema HAR supervisado requiere un entrenamiento previamente con un conjunto de datos etiquetados, por el contrario, el no supervisado no requiere entrenamiento, sino que tiene configurado un conjunto de reglas durante el desarrollo. En este trabajo particular nos enfocamos en un sistema HAR del tipo supervisado para reconocer las siguientes seis actividades humanas: caminar (WK), subir escaleras (WU), bajar escaleras (WD), parado (ST), estar acostado (LD) y estar sentado (SD). Denominamos, de manera particular, a las actividades WK, WU y WD como actividades dinámicas dado que involucran un movimiento voluntario que provoca un desplazamiento y se refleja claramente en los sensores inerciales, y a las actividades ST, LD y SD las denominamos actividades estáticas dado no involucran movimientos voluntarios del sujeto y no existe un desplazamiento de la persona.

2.1.1. Señales usadas en el HAR

En los sistemas HAR es común utilizar señales e imágenes que provienen de sensores que pueden ser ubicados en un espacio físico determinado, como en una habitación, o que pueden ser colocados o portados por las personas, como los que encontramos en los teléfonos celulares inteligentes o en los relojes inteligentes. Los teléfonos inteligentes son teléfonos móviles, que son capaces de realizar tareas similares a las de una computadora, como la capacidad de almacenar y procesar datos y poder navegar en Internet (Palencia-Vizcarra y Palencia-Díaz, 2013). Además, en comparación con los computadores personales, los teléfonos inteligentes son ampliamente aceptados debido a su reducido tamaño, bajo peso, un aparato más personal y, en especial, una gran conectividad que le permite acceder en todo momento y lugar a sitios de información y a redes sociales (Organista-Sandoval, McAnally-Salas, y Lavigne, 2013). Otras aplicaciones que suelen estar presentes son las cámaras integradas, la administración de contactos, un software de multimedia capaz de reproducción de música y poder visualizar fotos y vídeos, y el uso de programas de navegación, y, además, tener la habilidad de visualizar documentos de negocios de diferentes formatos como PDF y Microsoft Office (Alonso, Artime, Rodríguez, y Baniello, 2011).

Actualmente, en la mayoría de teléfonos inteligentes vienen instalados diferentes sensores, como sensores de posicionamiento, sensores de proximidad, sensores de temperatura, acelerómetro, giroscopio, magnetómetro, micrófono, etc., tal como se muestra en la figura 1. Esta alta gama de sensores abre las puertas para ofrecer servicios para monitorear los aspectos físicos y fisiológicos de los seres humanos. Esto es un reto en la actualidad llevado a cabo por diferentes comunidades científicas, en particular de los campos de visión por computador, procesamiento de señales y aprendizaje automático. Los sensores suelen ser operados por un microcontrolador o microprocesador, el cual cumple la función de un computador.



Figura 1: Sensores encontrados en un teléfono celular inteligente. Imagen tomada de ¹

2.1.2. Sensores inerciales

Los sensores inerciales son sensores basados en el principio de inercia, la tendencia de un cuerpo a conservar su velocidad (en ausencia de una influencia externa, un cuerpo se mantiene en un movimiento rectilíneo uniforme). Hay diferentes tipos de sensores para medir las señales que pueden ser usados por los sistemas HAR. Dos de los más usados son el acelerómetro y el giroscopio. El acelerómetro mide la aceleración (en metros por segundo cuadrado, m/s^2) basada en las diferentes variaciones que realiza una capacitancia en el interior del sensor. Esta capacitancia es un sistema microelectromecánico (MEMS por sus siglas en inglés *microelectromechanical systems*) que consiste en la suspensión de partículas de silicio que se encuentran situadas en un punto fijo y son movidas libremente en el eje donde se mide. Al ocurrir una aceleración, las partículas se desplazan y rompen con el equilibrio en la capacitancia; esto es medido para brindar la información que ocurre en un determinado eje. Para medir la aceleración tridimensionalmente se deben incluir tres sensores en un teléfono inteligente. Estos sensores deben colocarse ortogonalmente entre sí, es decir, formando un ángulo recto (condición de perpendicularidad), para obtener las aceleraciones A_x , A_y y A_z de cada dirección espacial (ejes x , y y eje z) independientemente (Vogt y Kuhn, 2012).



Figura 2: Ejes de medición de un acelerómetro en un teléfono inteligente. Imagen tomada de ²

¹<https://www.mdpi.com/1424-8220/19/9/2164>

²<https://bcninfoservices.files.wordpress.com/2013/08/acelerometro-sensor-smartphone.png>

El giroscopio es otro sensor que mide la velocidad angular y que está destinado a complementar la información sobre la orientación del móvil que ofrece el acelerómetro ³. El giroscopio funciona transformando la fuerza de un movimiento angular que se efectúa sobre éste en una señal eléctrica proporcional. Básicamente el giroscopio mide la velocidad de giro y su orientación (Ferrer Mínguez, 2009). Esta velocidad angular se mide en radianes por segundo (rad/s). A través de un ángulo de giro con respecto a un eje, en este caso tres ejes (cabeceo *pitch*, balanceo *roll* y guiñada *yaw*) se adquieren valores de rotación, generando así el sistema de coordenadas (Ferrer Cid, 2017).

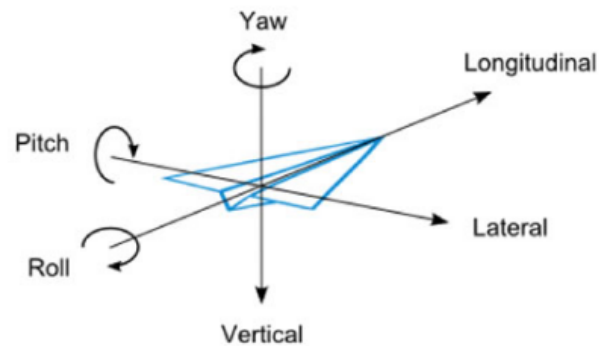


Figura 3: Ejes de medición de un giroscopio en un avión. Imagen tomada de ⁴

Usando el giroscopio y acelerómetro juntos se obtienen información sobre el plano del espacio en el que el teléfono inteligente se encuentra, su aceleración con respecto al punto de partida y si el dispositivo está girando sobre el plano o inclinándose, es decir, que pueden detectar el movimiento del dispositivo en seis ejes (Marcos Plaza, 2017).

2.2. Representación de señales en tiempo-frecuencia

Una serie de tiempo, también llamada serie cronológica, se basa en una secuencia de observaciones, valores o datos, tomados en momentos específicos y ordenados de manera cronológica (Villavicencio, 2010). Al momento de medir los datos, estos pueden estar separados en intervalos iguales, como en un electrocardiograma en el que existe una frecuencia de muestreo, o en intervalos desiguales, como en la medición del peso de una persona que se toma en diferentes ocasiones. Por lo anterior, se puede decir que las señales inerciales son series de tiempo.

³<https://www.xataka.com/basics/12-sensores-que-encontraras-tu-movil-sirven>

⁴<https://bcninfoservices.wordpress.com/2013/08/23/sensores-de-smartphones-conoce-cuales-son-y-como-funcionan/>

El análisis de series de tiempo se puede realizar en varios dominios, tales como el dominio del tiempo y el dominio de la frecuencia. El análisis en el dominio del tiempo generalmente involucra el cálculo de medidas descriptivas, como media, varianza, etc., para obtener información acerca de la señal. Con respecto al análisis en el dominio de la frecuencia se usa una función espectral para estudiar cómo la variación de señal puede ser representada como una mezcla de senos y cosenos (Wei, 2006). Dadas las limitaciones de realizar análisis en forma independiente en el dominio del tiempo y en el dominio de la frecuencia, por ejemplo, en la estimación de una frecuencia instantánea en un tiempo determinado, se han propuesto técnicas para realizar conjuntamente la representación de señales en tiempo-frecuencia, tales como el espectrograma, la transformada de Fourier de tiempo corto, la transformada de ondícula, la transformación de Wigner-Ville, y la descomposición empírica en modos.

2.2.1. Descomposición empírica en modos

La descomposición de una señal en varios componentes es una poderosa técnica de procesamiento digital de señales que ha sido aplicada con éxito en diferentes contextos, particularmente para eliminar componentes indeseables (filtrado, reducción de ruido), para identificar componentes deseables (filtrado, recuperación de información) y para reducir la cantidad de datos necesarios para el almacenamiento y la transmisión (compresión de datos). Se han propuesto diferentes métodos para descomponer una señal, tanto en tiempo discreto como en tiempo continuo, como la serie de Fourier, la transformada de Fourier, las transformadas de ondículas, las transformadas seno y coseno, la transformada de Karhunen-Loève y la descomposición empírica en modos. La EMD es una técnica de análisis de tiempo-frecuencia, adaptativa y basada en los datos, que descompone una señal en un número finito de funciones de modo intrínseco (IMF) que contiene oscilaciones rápidas y lentas, de manera que la suma de IMF proporciona la señal original (Huang y cols., 1998).

La EMD se basa en las siguientes suposiciones: (i) la señal posee al menos dos extremos, un máximo y un mínimo, (ii) las características de la escala de tiempo están dadas por el lapso de tiempo entre los extremos y si los datos están exentos de extremos, pero tienen solo puntos de inflexión, éstos pueden ser diferenciadas una o más veces para revelar el extremo.

Para el proceso de EMD se calculan los extremos locales (máximos y mínimos). Estos extremos son interpolados con *splines* cúbicos para obtener las envolventes: los puntos máximos para la envolvente superior y los mínimos para la inferior. Luego se halla el promedio de estas dos envolventes para obtener una tercera envolvente, pero promedio. Esto se puede apreciar en la figura 4. Se procede a obtener el residuo entre la señal original y la envolvente promedio, y el resultado sería un posible candidato a IMF. Para establecerla como una

IMF real se debe tener en cuenta las siguientes condiciones: (i) el número total de extremos locales y el número de cruces por cero deben ser iguales o distintos por valor máximo de uno, y (ii) que en cualquier lugar de la IMF el valor medio de la envolvente definido por los máximos locales debe ser cero. Si estas condiciones no se cumplen se repite el procedimiento mencionado anteriormente, pero esta vez la señal a procesar es con el residuo encontrado. Este procedimiento es denominado cribado (del inglés *sifting*) y es repetido hasta que la señal no contenga más de dos extremos locales. La señal original quedará expresada como la suma de estas IMF (Huang y cols., 1998).

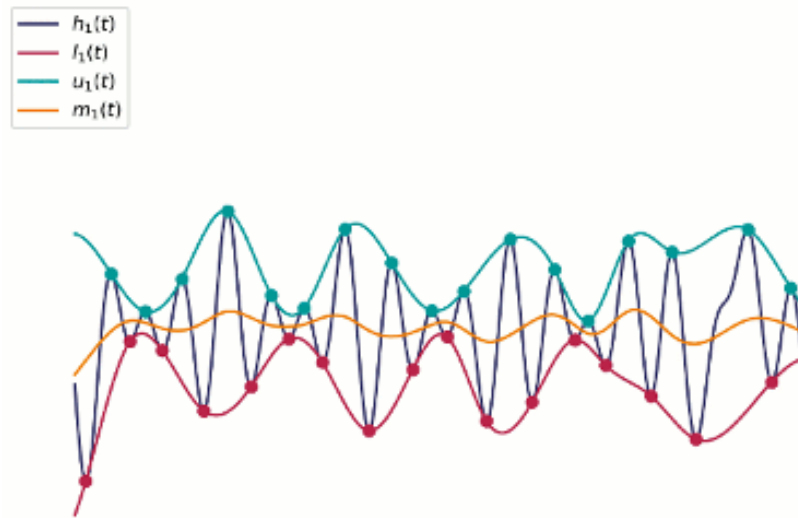


Figura 4: Procedimiento de Sifting. Imagen tomada de ⁵

2.2.2. Descomposición empírica en modos por conjunto completa con ruido adaptativo mejorada

Dada la versatilidad de EMD para descomponer una señal en el dominio del tiempo en una base completa y casi ortogonal (es decir, las IMF), todas con la longitud de la señal original, que reside en el dominio del tiempo, pero con frecuencias variables, versiones mejoradas (familia de EMD) han sido propuestas para enfrentar algunas limitaciones y problemas de la EMD (Huang y Wu, 2008). Por ejemplo, cuando la señal a descomponer contiene señales intermitentes, es decir, señales que son más cortas que la señal completa, se produce el

⁵https://en.wikipedia.org/wiki/File:Emd.example_lowres.gif

problema de mezcla de modos, las IMF pierden su significado físico y la distribución en tiempo-frecuencia se superpone. La mezcla de modos implica que las oscilaciones con diferentes escalas de tiempo residen en la misma IMF o que las oscilaciones con una escala de tiempo similar residen en diferentes IMF (Huang y cols., 1998; Wu y Huang, 2009).

La adición de ruido a la señal a descomponer se ha propuesto como una solución al problema de la mezcla de modos, y ha dado lugar a variantes de la EMD como la EMD por conjuntos (EEMD, por sus siglas en inglés, Ensemble Empirical Mode Decomposition) (Wu y Huang, 2009), la EEMD completa con ruido adaptativo (CEEMDAN, por sus siglas en inglés, Complementary Ensemble Empirical Mode Decomposition with Assisted Noise) (Torres, Colominas, Schlotthauer, y Flandrin, 2011) y la CEEMDAN mejorada (ICEEMDAN) (Colominas, Schlotthauer, y Torres, 2014). Esta última técnica no solo aborda el problema de la mezcla de modos, sino que también reduce el ruido residual en las IMF, logrando así una mejor descomposición y reconstrucción de una señal (error de reconstrucción insignificante) (Colominas y cols., 2014).

La EMD y sus variantes asistidas por ruido pueden descomponer señales no lineales y no estacionarias (como las generadas por los sistemas biológicos) sin definir una función base (una función wavelet madre debe definirse explícitamente antes de aplicar la transformada wavelet) y sin establecer a priori el nivel de descomposición (Labate y cols., 2013; Motin, Karmakar, y Palaniswami, 2019). Además, la inserción de ruido y la cantidad de realizaciones de ruido requeridas para descomponer una señal han hecho que las técnicas basadas en EMD asistidas por ruido superen a otras técnicas en varios dominios (Ren, Suganthan, y Srikanth, 2014; Kærgaard, Jensen, y Puthusserypady, 2016; Zhan y Li, 2016), como para analizar el electroencefalograma con ataques epilépticos (Hassan y Subasi, 2016), para mejorar la precisión de predicción del flujo mensual (L. Wang, Li, Ma, y Bai, 2019), y para extraer la frecuencia respiratoria del fotopletoxiograma (Motin y cols., 2019). Estas propiedades hacen que las técnicas basadas en EMD sean particularmente adecuadas para descomponer señales de los sensores inerciales acelerómetro y giroscopio en modos que se pueden analizar por separado para obtener información reveladora.

La ICEEMDAN es el producto de la constante mejora y solución de problemas que presenta el algoritmo de la descomposición empírica en modos. Esta técnica estima las medias locales, en vez de estimar los modos como se hacía en técnicas anteriores. La ICEEMDAN se rige por la siguiente ecuación:

$$\tilde{h}_1 = x - \frac{1}{I} \sum_{j=1}^I m_1^{(j)}, \quad (2.1)$$

donde \tilde{h}_1 es la IMF₁ obtenida por ICEEMDAN, I el número de intentos o realizaciones de ruido blanco, $m_1^{(j)}$ es la media local de la IMF₁ obtenida por EMD al descomponer la señal x más la realización j de ruido blanco $R^{(j)}$ ($j = 1, \dots, I$).

El algoritmo planteado es el siguiente:

1. $n = 0, r_0 = x$.
2. Generar las señales más ruido $x^{(j)} = r_n + R^{(j)}$.
3. $n = n + 1$. Calcular la media local para cada $x^{(j)}$ utilizando EMD, obteniendo $m_n^{(j)}$.
4. Calcular $r_n = \frac{1}{I} \sum_{j=1}^I m_n^{(j)}$.
5. Calcular $\tilde{h}_n = r_{n-1} - r_n$ y volver al paso 2.
6. Continuar hasta que r_n no tenga más de dos extremos.

De esta manera, se alivian las alteraciones ocasionadas por el ruido residual en las IMF reduciéndolo significativamente. Sin embargo, esta técnica exige una mayor carga computacional respecto a sus antecesoras.

2.3. Métricas para analizar las señales

A lo largo de este trabajo se han empleado varias métricas para analizar las señales derivadas de los sensores inerciales con el fin de encontrar información valiosa que me permita diferenciar las actividades realizadas por las personas. Estas métricas se describen a continuación.

2.3.1. Promedio

El promedio o también llamada media aritmética de un grupo de datos o serie de tiempo es quien la caracteriza cuantitativamente, es obtenido sumando todos sus valores y dividiendo esto por entre el número de sumandos:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (2.2)$$

2.3.2. Varianza

La varianza corresponde a una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media:

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.3)$$

La varianza puede aportar mucha información de una variable en concreto. Los valores atípicos influyen significativamente sobre la varianza, por esto no se recomienda usarla cuando las distribuciones de las variables aleatorias tienen colas pesadas.

2.3.3. Desviación estándar

La desviación estándar, representada por σ , es una medida que se implementa para cuantificar la dispersión de un conjunto de datos o variación. La desviación estándar es igual a la raíz cuadrada de la varianza. Tener una desviación baja nos dice que los datos están agrupados cerca de su promedio, que es lo que se espera, mientras que un valor alto quiere decir que los datos se encuentran en valores más amplios.

2.3.4. Entropía de Shannon

La entropía de Shannon mide la incertidumbre de una fuente de información. La entropía también se considera como la cantidad de información promedio que contienen los valores de una señal. La entropía de Shannon de una señal o serie de tiempo x viene definida por:

$$H_x = - \sum_{i=1}^n P(x_i) \log P(x_i), \quad (2.4)$$

donde $P(x_i)$ corresponde a la probabilidad de x_i y cumple con la condición $\sum_{i=0}^n P(x_i) = 1$, y \log corresponde, en este documento, al logaritmo natural (\ln o \log_e) y tiene unidades en nats (unidades naturales). Los valores con menor probabilidad son los que aportan más información. Cuando todos los símbolos son igualmente probables (distribución de probabilidad plana), todos aportan información relevante y la entropía es máxima. Básicamente, una entropía más alta indica sistemas más complejos o caóticos, por lo tanto, menos previsibilidad.

2.4. Aprendizaje automático y aprendizaje profundo

El aprendizaje automático es una rama de las ciencias de la computación y de la inteligencia artificial que se centra en la creación de técnicas para que las máquinas aprendan y que tengan la capacidad de cambiar cuando son expuestos a nuevos datos. Estos algoritmos se basan en la detección de patrones y obtención de predicciones a través de estos datos. Actualmente, los términos aprendizaje automático (del inglés *machine learning*) e inteligencia artificial se han vuelto muy popular debido a sus innumerables usos y aplicaciones, y a su vez, a los buenos resultados que se obtienen con estos, siendo usados en asistentes virtuales, recomendación de productos según búsqueda y gustos, entre otros.

El aprendizaje automático se puede clasificar en supervisado y no supervisado. El supervisado se basa de información previamente etiquetada o clasificada para para realizar sus propias acciones, y el no supervisado está capacitado para extraer e inferir información relevante de los datos sin una previa etiqueta.

Al momento de realizar clasificaciones, predicciones, entre otras tareas, lo ideal es que el modelo tenga la capacidad de generalizar su conocimiento aprendido, si esta tarea no se realiza de manera correcta el modelo puede caer en dos grandes fallos llamados sobreajuste (*overfitting*) y sub-ajuste (*underfitting*). El sobreajuste ocurre cuando el modelo solo aprende con casos particulares y al ingresar un dato nuevo, al no ser idéntico a éstos, no es capaz de identificarlo, mientras que el sub-ajuste sucede cuando la información de entrenamiento es muy mínima e identifica cualquier cosa como correcta.

El ingreso del tipo de datos para el aprendizaje a estos modelos también es de suma importancia debido a que si los datos pueden resaltar y destacar entre sí información que permita distinguir las clases, el modelo aprenderá y podrá clasificar o reconocer patrones de una mejor manera. Para esto se realizan diferentes tipos de procedimientos, tales como operaciones de transformación a los datos (potenciación, logaritmo, espectro de frecuencias y descomposiciones).

El aprendizaje profundo (del inglés *deep learning*) es un método del aprendizaje automático que se caracteriza por que sus modelos deben ser entrenados por medio de grandes cantidades de datos previamente etiquetados. El aprendizaje profundo está resaltando últimamente debido a que está consiguiendo resultados que anteriormente no eran posibles de alcanzar, incluso para los humanos. Es utilizado en vehículos autónomos (sin conductor) para reconocer señales de tránsito, obstáculos en la vía y a las personas, y es una clave fundamental para su funcionamiento. Al necesitar grandes cantidades de datos y poseer extensos números de capas ocultas, este método exige y requiere una potencia de cálculo significativa, es decir, demanda una carga computacional grande.

La clasificación usando técnicas de aprendizaje automático o aprendizaje profundo puede ser binaria, en la que se debe distinguir entre dos clases únicas, como tomar una decisión de verdadero o falso, y multiclase, que, como su nombre lo indica, involucra distinguir entre más de dos clases. En este trabajo se utiliza el aprendizaje del tipo supervisado para clasificación multiclase (seis clases), es decir tener la capacidad de identificar la etiqueta de una clase, con una lista de seis posibilidades previamente definidas. Varios métodos de aprendizaje automático supervisado pueden ser utilizados en la clasificación multiclase, tales como regresión logística multinomial, árbol de decisión, k vecinos más próximos (kNN, por sus siglas en inglés *k-nearest neighbors*), redes neuronales artificiales y redes neuronales profundas. En este trabajo se utiliza las redes del tipo profunda.

2.4.1. Redes neuronales artificiales

Las redes neuronales artificiales (ANN, por sus siglas en inglés *Artificial Neural Networks*) son un modelo de aprendizaje automático compuesta por un conjunto de estructuras llamadas neuronas, que emulan el funcionamiento del cerebro humano, y que se encuentran conectadas entre sí para transmitir información. La información fluye a través de éstas para obtener unos valores de salida. Estas conexiones de neuronas se realizan por medio de enlaces (emula una sinapsis, estructura del sistema nervioso que permite que una neurona transmita una señal a otra neurona o a una célula) en los cuales la salida se rige por un valor de peso. Los pesos pueden inhibir o aumentar el estado permitiendo la activación de las siguientes neuronas.

Las neuronas se agrupan en capas, en las cuales se realizan diferentes transformaciones a sus entradas. Las neuronas poseen señales de entrada ponderadas produciendo una señal de salida por medio de una función limitadora que modifica el resultado o pone un límite que hay que superar antes de pasar a la siguiente neurona, conocida como función de activación. En la figura 5 se aprecia la estructura de una red neuronal artificial y su función de activación. Esta función de activación se encarga de comprobar el valor de salida de una neurona, para determinar si las conexiones siguientes tomen la neurona como activada o no. Las funciones de activación más comunes son la tangente hiperbólica (\tanh), ReLu (por sus siglas en inglés *Rectified Linear Unit*), Softmax y la función logística o sigmoide.

Las redes neuronales artificiales están organizadas en capas, como se muestran en la figura 6. La topología de una red neuronal está definida por las capas y la naturaleza de las conexiones entre las neuronas. Una capa es un conjunto de neuronas con características similares. Existen capas de entrada, capas ocultas y capas de salida. Las conexiones entre las neuronas pueden ser unidireccionales (hacia adelante, del inglés *feedforward*) o recurrentes

⁶<http://samuelabad1991.blogspot.com/2014/02/analisis-con-redes-neuronales-neural.html>

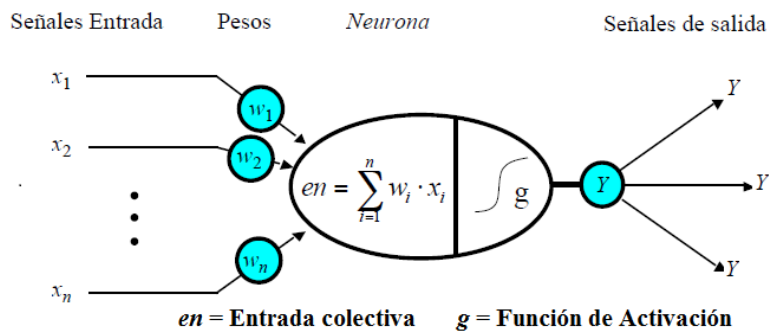


Figura 5: Arquitectura de una red neuronal artificial. Imagen tomada de ⁶

(con al menos un lazo de realimentación, del inglés *feedback*). De esta forma se definen tres topologías fundamentales (Miranda Calle y Gamboa Entralgo, 2018):

- Las redes monocapa unidireccionales en donde los datos fluyen en un único sentido y sólo se tienen dos capas (la capa de entrada y la capa de salida). Se le denomina monocapa porque sólo se consideran las capas que contienen neuronas; la capa de entrada es un conjunto de fuentes de información.
- Las redes multicapa unidireccionales en donde los datos fluyen en un único sentido, pero a diferencia de las redes monocapa unidireccionales, existen una o más capas ocultas. La función de las capas ocultas es realizar un procesamiento parcial tomando los datos de la capa de entrada o de otra capa oculta y generar una salida que sirva como entrada para la capa de salida u otra capa oculta.
- Las redes recurrentes en donde existe al menos un lazo de realimentación que comunique las salidas de la red neuronal con sus entradas.

2.4.2. Redes neuronales recurrentes

Un problema que está siempre presente en las redes neuronales tradicionales es que la información no persiste en cada uno de los ciclos si no que empieza de nuevo desde cero, es decir, que se pierde a través del tiempo. Para ello se desarrollaron las redes neuronales recurrentes (RNN, por sus siglas en inglés *Recurrent Neural Networks*) en las cuales se establece una retroalimentación donde valores posteriores dependen directamente de los anteriores.

⁷<https://la.mathworks.com/discovery/deep-learning.html#howitworks>

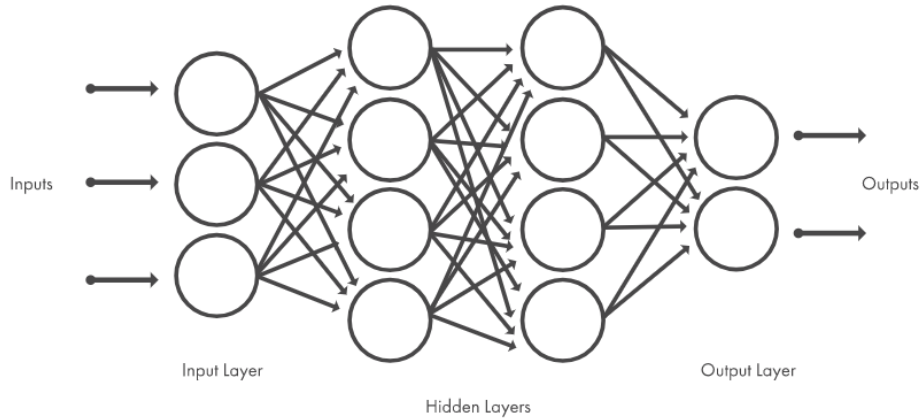


Figura 6: Redes neuronales organizadas en capas. Imagen tomada de ⁷

Estas redes no poseen una estructura definida de capas, sino que tienen conexiones arbitrarias entre sí, creando ciclos y dando paso a que la red obtenga memoria. Son muy usadas para series de tiempo y análisis de secuencias (vídeos, sonido, textos). La estructura de una red neuronal recurrente se muestra en la figura 7. A pesar de todos estos beneficios, este tipo de redes presentan un problema en cuanto se habla de un largo plazo de tiempo.

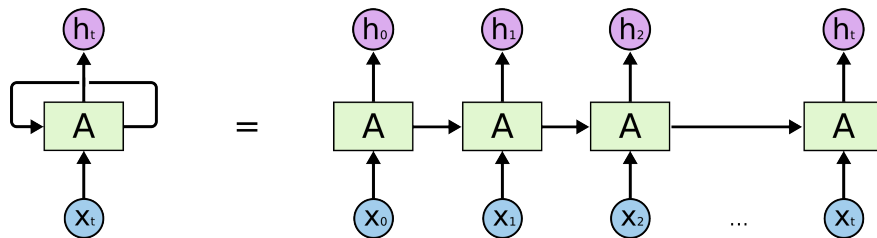


Figura 7: Estructura red neuronal recurrente. Imagen tomada de ⁸

Red LSTM

Las RNN por lo general presentan dificultades en su entrenamiento a causa de que los gradientes retro propagados no dependen únicamente de los errores presentes, sino que también de errores pasados, por esto son propensos a crecer o desvanecerse. La constante aglomeración de estos errores ocasiona impedimento para memorizar dependencias a largo plazo. En las redes LSTM introducidas por Hochreiter y Schmidhuber en 1997 (Hochreiter y Schmidhuber, 1997) se establece una secuencia de pasos para saber qué información se almacenará

⁸<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

o será eliminada. Están predeterminadas para recordar información por largos periodos de tiempo. Todas las RNN se forman por una cadena de módulos repetitivos de red neuronal, pero en una RNN tradicional, el módulo repetitivo tiene una estructura muy simple, ver figura 8.

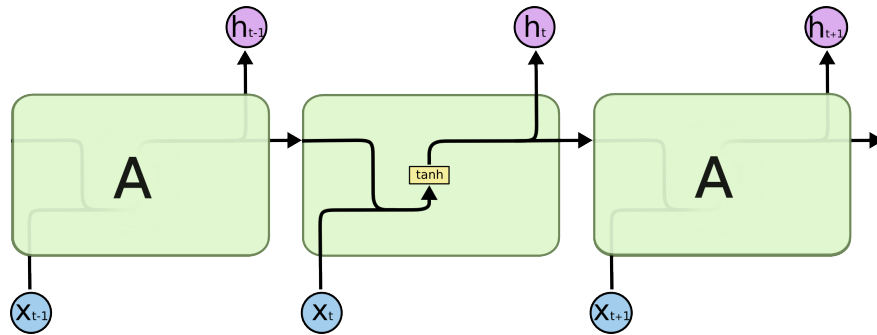


Figura 8: Módulo de repetición de una RNN tradicional. Imagen tomada de ⁹

Las redes LSTM también se forman por una cadena, pero el módulo de repetición varía con respecto a la anterior, en vez de poseer una sola capa de red neuronal, tiene cuatro, las cuales cumplen una función específica (ver figura 9).

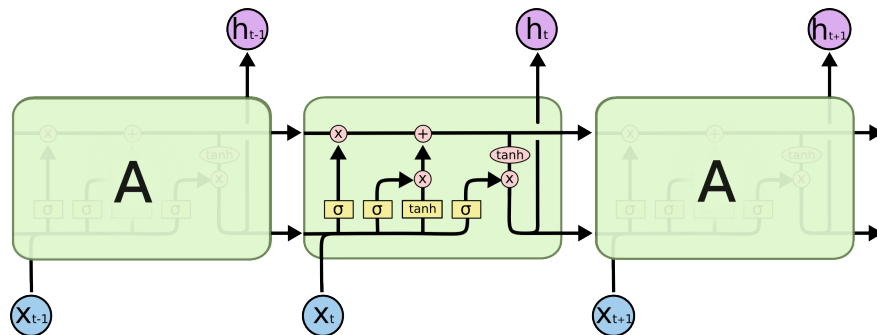


Figura 9: Módulo de repetición de una red LSTM. Imagen tomada de ¹⁰

Lo más fundamental para las redes LSTM es el estado de la celda; la línea horizontal que pasa por la parte superior de la figura 10. El estado de la celda actúa como una especie de cinta transportadora. Fluye directamente por toda la cadena, con solo algunas interacciones lineales menores. Es muy fácil que la información fluya sin cambios. Estas redes tienen la capacidad de eliminar o agregar información al estado de la celda, cuidadosamente regulado por estructuras llamadas puertas. Las puertas son una forma opcional de dejar pasar

⁹<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

¹⁰<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

la información. Se componen de una capa de red neuronal sigmoidea y una operación de multiplicación puntual.

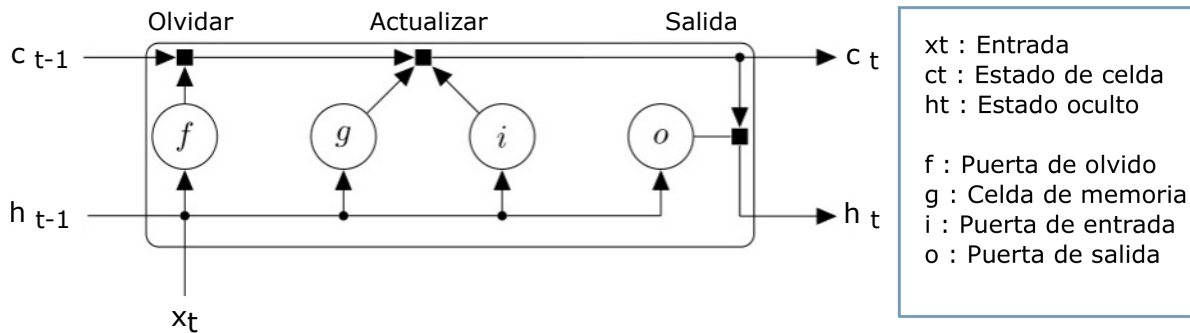


Figura 10: Bloque de una red LSTM. Imagen tomada de de (Suárez Mantilla y Ardila Navas, 2019)

A continuación se muestra la función de cada una de las puertas (Suárez Mantilla y Ardila Navas, 2019):

- Puerta de olvido: Es el primer paso en la red LSTM. Se encarga de decidir qué información debe ser olvidada del estado de celda. La decisión se toma examinando h_{t-1} y x_t por cada celda de estado C_{t-1} . La salida es un número entre 0 y 1, donde un 0 representa olvidar algo completamente y un 1 mantener la información.
- Puerta de entrada: El siguiente paso es decidir qué nueva información vamos a almacenar en el estado de celda. La puerta de entrada consiste en una función logística y su salida h_{i_t} representa el valor que va a actualizar la celda LSTM.
- Celda de memoria: Ahora es el momento de actualizar el antiguo estado de celda, C_{t-1} , al nuevo estado de celda C_t . Los pasos anteriores ya han ayudado a decidir qué hacer, sólo hace falta la ejecución. La celda de memoria crea un vector con diferentes valores que son combinados con h_{i_t} y hace una actualización del estado. La actualización de la celda de estado consiste en olvidar lo que debe ser olvidado y agregar lo que debe ser agregado.
- Puerta de salida: Finalmente, la salida de la celda LSTM combina la información actual junto con la pasada. La puerta de salida representa la información de la celda de estado combinada con la información pasada.

Las redes LSTM son redes neuronales profundas, las cuales aprenden directamente de los datos, sin necesidad de extraer manualmente características. Estas redes pueden tener hasta 150 capas, de ahí el término profundo (MathWorks, 2018).

Las opciones de entrenamiento de una red LSTM corresponden a los parámetros de configuración del modelo de clasificación que actúan internamente. Estos valores pueden ser estimados según los datos que se procesen. Si estos parámetros no son especificados, se utilizarán valores por defecto brindados por el programa, los cuales fueron determinados debido a que funcionan bien para grandes cantidades de casos o tareas. A continuación se mencionan y describen los parámetros (MathWorks, 2019):

- Tasa de aprendizaje, del inglés *Learning Rate*: Se especifica por un escalar positivo, corresponde al valor inicial utilizado para el entrenamiento. Los valores comunes por defecto se encuentran entre 0.01 y 0.001. Si la tasa de aprendizaje es demasiado baja, el entrenamiento toma mucho tiempo. Si en cambio, es demasiado alta, entonces el entrenamiento puede alcanzar un resultado deficiente.
- Tamaño del grupo, del inglés *Batch Size*: Se especifica por un número entero positivo, corresponde al tamaño de los grupos que deben utilizarse para cada iteración de entrenamiento. Un *batch* o grupo, es un subconjunto del entrenamiento que se utiliza para evaluar el gradiente de la función de pérdida y actualizar los pesos de la red.
- Periodo de cambio, del inglés *Epochs*: Se especifica por un número entero positivo y corresponde al número de períodos que deben utilizarse para el entrenamiento. Una iteración es un paso que se da en el algoritmo de la gradiente de descenso para minimizar la función de pérdida haciendo uso de un *batch*. Un *epoch* o periodo de cambio, es el recorrido completo del algoritmo de entrenamiento a lo largo de todo el conjunto de datos.

2.4.3. Medidas de desempeño

El desempeño de un clasificador suele observarse en una matriz de confusión, en la cual se visualizan las discrepancias entre las clases predichas y las clases verdaderas: las clases resultantes dadas por un clasificador contrastan con las clases verdaderas en las mismas instancias (Fawcett, 2006; Powers, 2011).

Una matriz de confusión general para n clases ($Clase_1, \dots, Clase_n$) se muestra en la tabla 1: las columnas corresponden a las clases predichas y las filas a las clases verdaderas. La diagonal principal (A_{11}, \dots, A_{nn}) corresponde a la clasificación correcta realizada por el clasificador.

Tabla 1: Matriz de confusión.

Clase verdadera	Predicción				
	$Clase_1$	$Clase_2$	$Clase_3$	\dots	$Clase_n$
$Clase_1$	A_{11}	A_{12}	A_{13}	\dots	A_{1n}
$Clase_2$	A_{21}	A_{22}	A_{23}	\dots	A_{2n}
$Clase_3$	A_{31}	A_{32}	A_{33}	\dots	A_{3n}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Clase_n$	A_{n1}	A_{n2}	A_{n3}	\dots	A_{nn}

Para un clasificador multiclase, el desempeño suele evaluarse usando las métricas exactitud, precisión, y exhaustividad.

Exactitud: del inglés *Accuracy*, representa la proporción de instancias clasificadas correctamente del total de instancias:

$$Acc = \frac{\sum_{i=1}^n A_{ii}}{\sum_{i=1}^n \sum_{j=1}^n A_{ij}} \quad (2.5)$$

Precisión: del inglés *precision*. La precisión Pre_i de la clase i representa la proporción de instancias clasificadas correctamente de la clase i (verdaderos positivos) del total de clasificaciones que se asignaron a la clase i :

$$Pre_i = \frac{A_{ii}}{\sum_{j=1}^n A_{ji}} \quad (2.6)$$

Exhaustividad: del inglés *recall*, también llamada sensibilidad. La exhaustividad Rec_i de la clase i representa la proporción de instancias clasificadas correctamente de la clase i (verdaderos positivos) del total de instancias que tiene la clase i como etiqueta verdadera:

$$Rec_i = \frac{A_{ii}}{\sum_{j=1}^n A_{ij}} \quad (2.7)$$

CAPÍTULO 3

ANTECEDENTES

En esta sección se presenta una revisión del estado del arte de las técnicas y de los enfoques concebidos para reconocer las acciones que realizamos como seres humanos.

3.1. Técnicas basadas en umbrales

En el 2002, Mathie and Lovell (Mathie, Lovell, Coster, y Celler, 2002) reconocieron siete actividades de la vida diaria en un ambiente controlado como estar de pie y sentarse durante 30 s cada uno, estar de pie durante 10 s, caminar, subir y bajar escaleras, etc. usando datos de un acelerómetro triaxial colocado en la cintura de 26 personas. El reconocimiento se basó en la comparación de la energía consumida por cada actividad con umbrales. Con este enfoque se lograron identificar correctamente las actividades dinámicas con una sensibilidad del 99 % y una especificidad del 94 %. No obstante, es importante resaltar que este alto desempeño se debe a que las actividades fueron colectadas en un ambiente controlado poco realista.

En el 2006, Heinz y colaboradores (Heinz, Kunze, Gruber, Bannach, y Lukowicz, 2006) usaron sensores inerciales (acelerómetro triaxial) colocados en las muñecas, pies, rodillas, cuello y en la parte trasera de la cintura, para reconocer en tiempo real movimientos típicos de Wing Tsun, un tipo de arte marcial, para desarrollar videojuegos. La técnica consistió en comparar las señales inerciales en el dominio del tiempo y en el dominio de la frecuencia con umbrales fijos. Los resultados obtenidos permitieron diferenciar entre los movimientos realizados por un aficionado y por un experto en este arte marcial.

En el 2007, Bidargaddi y colaboradores (Bidargaddi, Sarela, Klingbeil, y Karunanithi, 2007) identificaron la actividad de caminar de otras cinco actividades diferentes como montar bicicleta y remar, con el fin de monitorear pacientes en rehabilitación cardiaca. El enfoque se basó en la comparación con umbrales de la descomposición en ondículas de las señales de un acelerómetro triaxial colocado en la cintura. Los resultados obtenidos mostraron una exactitud superior al 95 %.

3.2. Técnicas basadas en aprendizaje automático

Uno de los primeros enfoques en realizar reconocimiento de actividades humanas fue realizado en 1985 hecho por Recker y McNally (Recker, McNally, y Root, 1985) en donde analizaron modelos de viaje y actividades usando un proceso de clasificación en el que el conjunto de medidas que definen el movimiento humano estuvo representado por un vector patrón. Se aplicaron técnicas de transformación a los vectores de patrón para desarrollar una taxonomía para el espacio de patrones. La inversión de los patrones transformados produjo patrones de actividad representativos que se relacionaron con las características socioeconómicas y urbanas.

En el 2004, Bao and Stephen (Bao y Intille, 2004) reconocieron diecinueve actividades de la vida diaria aplicando arboles de decisión a las señales de acelerómetros triaxial colocados en la muñeca, codo, tobillo, pierna y cintura de 20 sujetos. Se obtuvo un 84% de exactitud para clasificar las diecinueve actividades, no obstante, las actividades de estirarse y subir al ascensor tuvieron un desempeño inferior al 50%.

En el 2006, Dean y colaboradores (Karantonis, Narayanan, Mathie, Lovell, y Celler, 2006) diseñaron un sistema de clasificación basado árboles de decisión implementado en un microcontrolador para reconocer actividades ambulatorias (caminar, caerse, sentarse, etc.) en tiempo real a partir de señales de acelerómetros. Se obtuvo una exactitud del 90.8% para doce tareas, siendo la actividad caerse la más fácil de detectar con un desempeño superior al 95%.

En el 2006, Pirttikangas, Fujinami y Nakajima (Pirttikangas, Fujinami, y Nakajima, 2006) realizaron el reconocimiento de actividades humanas mediante el uso de pequeños sensores ubicados en cuatro partes del cuerpo: muslo izquierdo, muñecas izquierda y derecha, y un collar. Se reconocieron diecisiete actividades de la vida cotidiana de hombres y mujeres. Los datos fueron adquiridos por medio de un acelerómetro triaxial. Además, se adquirió la frecuencia cardiaca de las personas. Se escogieron las mejores características con el algoritmo de búsqueda secuencial hacia adelante y hacia atrás, y se evaluaron dos clasificadores, donde el mejor resultado fue 90.61% con un clasificador kNN.

En el 2010, Altun y Barshan (Altun y Barshan, 2010) reconocieron diecinueve actividades usando de diferentes métodos de clasificación, como máquinas de soporte vectorial, redes neuronales artificiales, toma de decisiones Bayesiana, etc. Para la extracción de características usaron métricas de estadística descriptiva de los datos tales como valores máximos, mínimos, media, varianza, etc. Esto proporcionó una gran cantidad de características, sin embargo, no todas ellas aportaban información de importancia para la clasificación. Para solucionar este problema usaron el análisis de componente principales, la cual es una transformación

que encuentra la combinación lineal óptima de las características, reduciendo de 1170 a 30 características. Gracias a esta reducción de características obtuvieron un desempeño de clasificación mayor a 95 % usando toma de decisiones Bayesiana.

En el 2010, Khan y colaboradores (Khan, Lee, Lee, y Kim, 2010) propusieron un método de validación de cinco actividades físicas diarias que fueron recolectadas de cinco partes diferentes del cuerpo humano usando un teléfono inteligente que tenía incorporado un acelerómetro triaxial. De cada señal fueron calculadas características como los coeficientes autorregresivos y el área de la magnitud de la señal y, usando un análisis discriminatorio de núcleos (*kernels*), se extrajeron las características discriminatorias no lineales. Usando redes neuronales artificiales obtuvieron una precisión promedio del 96 %.

En el 2011, Kwapisz, Weiss y Moore (Kwapisz, Weiss, y Moore, 2011) implementaron un sistema de reconocimiento donde recolectaron datos a veintinueve personas que realizaban actividades diarias como caminar, trotar, escalar, entre otras, usando el acelerómetro para adquirir los datos. Usaron la data de entrenamiento resultante con el fin de predecir el modelo de reconocimiento de actividades y poder crear una base de datos que detecte las actividades del usuario y así crear un perfil que determine si el usuario está teniendo una vida saludable.

En el 2011, Lee y Cho (Lee y Cho, 2011) realizaron un sistema de reconocimiento de actividades adquiriendo señales de acelerómetro por medio de un teléfono inteligente. Las señales fueron procesadas usando modelos jerárquicos ocultos de Markov. Desarrollaron una aplicación de reconocimiento de actividades en tiempo real con un teléfono inteligente por medio de una plataforma para Android de Google.

En el 2013, Anguita y colaboradores (Anguita, Ghio, Oneto, Parra, y Reyes-Ortiz, 2013) tomaron señales de acelerómetro y giroscopio de 30 sujetos que realizaban actividades de la vida diaria, como caminar o estar parado, mientras usaban un teléfono inteligente ubicado en la cadera con sensores inerciales integrados. Tuvieron una exactitud de 96 % usando máquinas de soporte vectorial.

En el 2016, Wang y colaboradores (Z. Wang y cols., 2016), recolectaron información de los tres ejes del acelerómetro para reconocer nueve actividades innovando al momento de resaltar características usando descomposición de modo empírico por conjunto (EEMD). Para la clasificación tuvieron dos métodos: SVM y kNN, y obtuvieron una mejora notoria con respecto a usar solo las señales crudas, aumentando el desempeño de la clasificación.

En el 2017, Tsunoda y colaboradores (Tsunoda, Komori, Matsugu, y Harada, 2017) usaron una red LSTM de jerarquía para reconocer jugadas de fútbol sala realizadas por los jugadores durante los partidos, a partir de una secuencia de imágenes. Usando como entrada a la red dos partes en las que se integraban múltiples características centradas en la persona y su ubicación en tres dimensiones, dedujeron que la utilización de múltiples características

centradas en una persona tenía más exactitud que usar una sola característica centrada en una persona.

En el 2018, Yu y Qin (Yu y Qin, 2018) clasificaron seis actividades a partir de señales recolectadas por sensores inerciales, usando una arquitectura de red bidireccional LSTM, trabajando con los datos crudos solo con el mínimo preprocesamiento brindado por la base de datos obteniendo una exactitud de 93.79%. En su trabajo compararon diferentes métodos de clasificación de otros autores con el obtenido por ellos, siendo el enfoque Bi-LSTM el que proporcionó el mejor resultado a pesar de solo utilizar los datos crudos.

Finalmente, en la Universidad Pontificia Bolivariana seccional Bucaramanga, en un trabajo reciente llevado a cabo en el grupo de investigación BISEMIC, Hernández y colaboradores (Hernández, Suárez, Villamizar, y Altuve, 2019) clasificaron seis actividades frecuentes del ser humano haciendo uso de redes neuronales recurrentes LSTM. Los parámetros de la red neuronal se escogieron mediante un enfoque de búsqueda exhaustiva, evaluando diferente cantidad de capas y neuronas, y tomando como función de costo la exactitud de la clasificación. En ese trabajo se obtuvo como resultado que los mejores parámetros eran tres capas 175 neuronas. Con esta arquitectura se obtuvo una exactitud de 92.67% usando sólo las señales crudas (sin procesar).

Como se puede observar, los trabajos de reconocimientos son muy variados y con diferentes aplicaciones. No solamente se han reconocidos actividades diarias como caminar sino también acciones deportivas. Además, se pudo observar que las señales para realizar este reconocimiento son muy variadas, pueden provenir de imágenes o de señales de sensores inerciales. Finalmente, se resalta que la gran mayoría de los trabajos se enfocan en un proceso de extracción de características, lo cual es una tarea que demanda una gran cantidad de tiempo y experticia, y muy pocos se han enfocados en redes del tipo profundo, como nuestro enfoque actual, en donde la extracción de característica es llevada a cabo por la propia máquina de aprendizaje, gracias a una gran cantidad de ejemplos (instancias) de entrada que son usados en el proceso de entrenamiento del modelo.

CAPÍTULO 4

MARCO METODOLÓGICO

En este capítulo se describe la metodología llevada a cabo en este trabajo. Este capítulo se divide en tres secciones. En la primera se describe los materiales e instrumentos utilizados para la realización del trabajo, en la segunda parte se detalla el procedimiento llevado a cabo para el análisis de las señales inerciales, y en la tercera se presenta el desarrollo del modelo de clasificación de las actividades humanas usando redes LSTM bidireccionales.

4.1. Materiales e instrumentos

Los materiales empleados para la realización del trabajo constan de herramientas de software en las que se desarrollaron los algoritmos, y de hardware para el procesamiento y almacenamiento de las diferentes etapas del proyecto.

4.1.1. Descripción del hardware y software

Para la elaboración de este trabajo se utilizó MATLAB[®] R2018a, en un sistema operativo basado en Linux (Ubuntu 18.04.1 LTS) instalado en un PC DELL equipado con una tarjeta gráfica Nvidia Quadro M4000 de 8 GB, 32 GB de RAM, una ROM de 2 TB y una CPU Intel[®] Xeon[®] Quad Core (2800MHz). El código para el algoritmo de ICEEMDAN ¹ fue producido por Colominas y colaboradores y se encuentra disponible de manera gratuita para su descarga y uso en MATLAB[®]. El documento fue escrito en LaTeX y las imágenes fueron editadas en Inkscape.

¹http://bioingenieria.edu.ar/grupos/ldnllys/meteorres/re_inter.htm#Codigos

4.1.2. Conjunto de datos

Las señales procesadas en este trabajo fueron obtenidas de la base de datos UCI Machine Learning repository ² (Anguita y cols., 2013). El conjunto de datos está compuesto por 30 sujetos, el cual está dividido en dos subconjuntos: 70 % para entrenamiento, correspondiente a 7352 registros, y 30 % para prueba, correspondientes a 2947 registros. Estos registros fueron divididos aleatoriamente por los creadores de la base de datos con el fin de asegurar la reproducibilidad de los resultados.

Cada registro de la base de datos consta de nueve señales, correspondientes a tres ejes de acelerómetro, tres de giroscopio y otras tres señales derivadas del acelerómetro. En las figuras 11 y 12 se muestran ejemplos de estas señales (acelerómetro A , giroscopio G , y aceleración total TA) para cada eje (x, y , y z) para las actividades dinámicas (WK, WU, y WD) y estáticas (ST, SD, y LD), respectivamente. En estas figuras se puede observar que la alta variabilidad y amplitud de las señales dinámicas en comparación con las actividades estáticas. Además, se puede observar cierto patrón embebido en las señales de las actividades dinámicas. Las actividades estáticas no hay un patrón definido que sea observable en las señales. Esto es debido a que las actividades dinámicas que se representan en las señales, como caminar, subir escaleras y bajar escaleras, tienden a tener a involucrar movimientos repetitivos para realizar dichas actividades. Por ejemplo, caminar puede verse como un movimiento repetitivo con una duración aproximadamente de un paso cada 0.5 segundos (un ciclo de marcha dura alrededor de 1 segundo).

4.2. Análisis de las señales inerciales

A continuación, se presenta el procedimiento realizado para el análisis de las señales inerciales de acelerómetro y giroscopio en los tres ejes, a través de las diferentes IMF, obtenidas por medio las técnicas EMD e ICEEMDAN, haciendo uso de la varianza y entropía de Shannon para describirlas.

4.2.1. Descomposición de las señales inerciales

Se unieron los dos subconjuntos brindados por la base de datos, es decir, tanto entrenamiento como prueba, teniendo un total de 10299 registros. Luego se descompusieron todas las señales inerciales $x(t)$ en N diferentes IMF: $IMF_1(t), IMF_2(t), \dots, IMF_N(t)$ por medio

²<https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

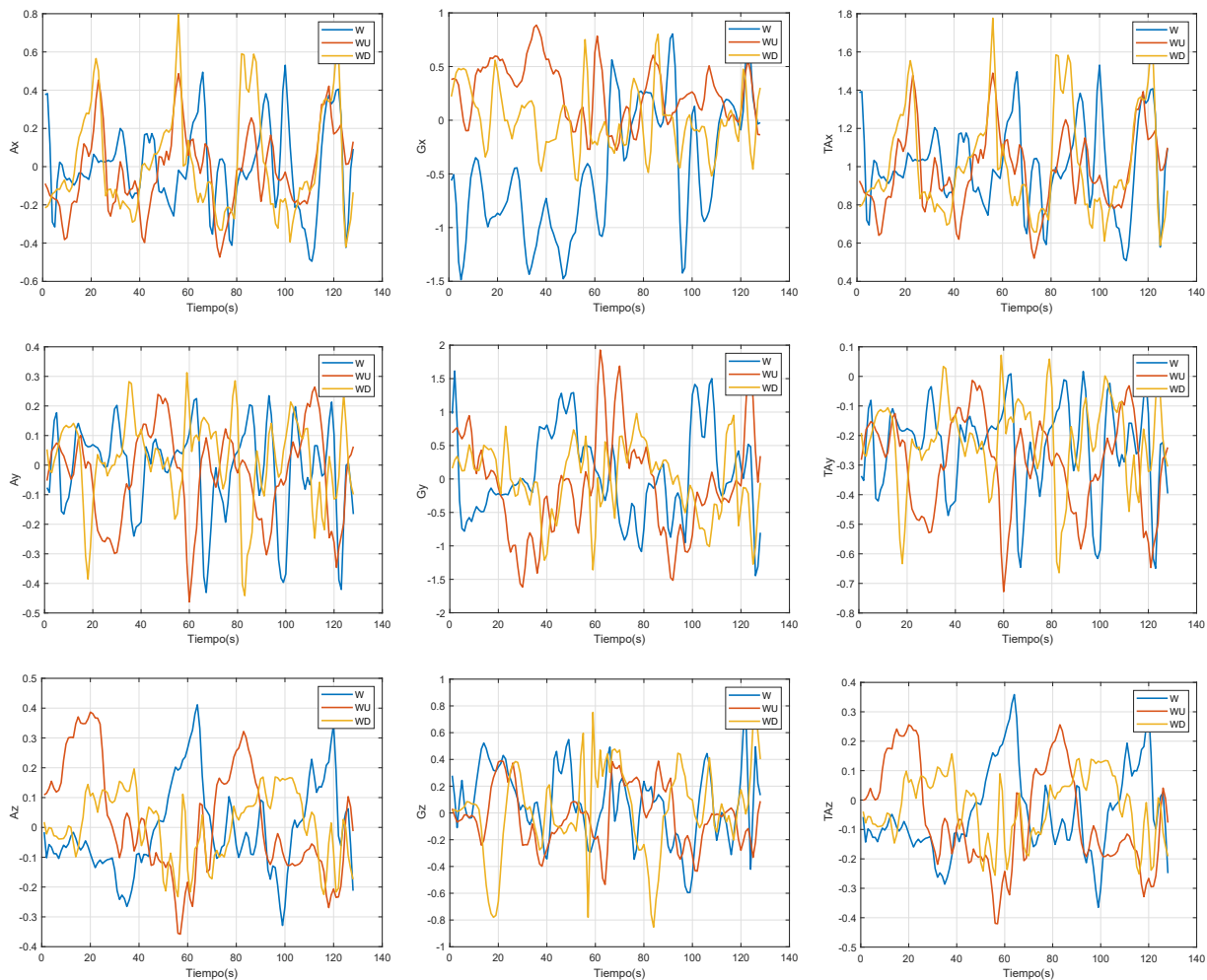


Figura 11: Ejemplo de señales de los sensores inerciales (acelerómetro en la columna izquierda, giroscopio en la columna central y aceleración total en la columna derecha) en cada eje (x en la primera fila, y en la fila central, y z en la ultima fila), para las actividades dinámicas (WK en azul, WU en rojo y WU en amarillo).

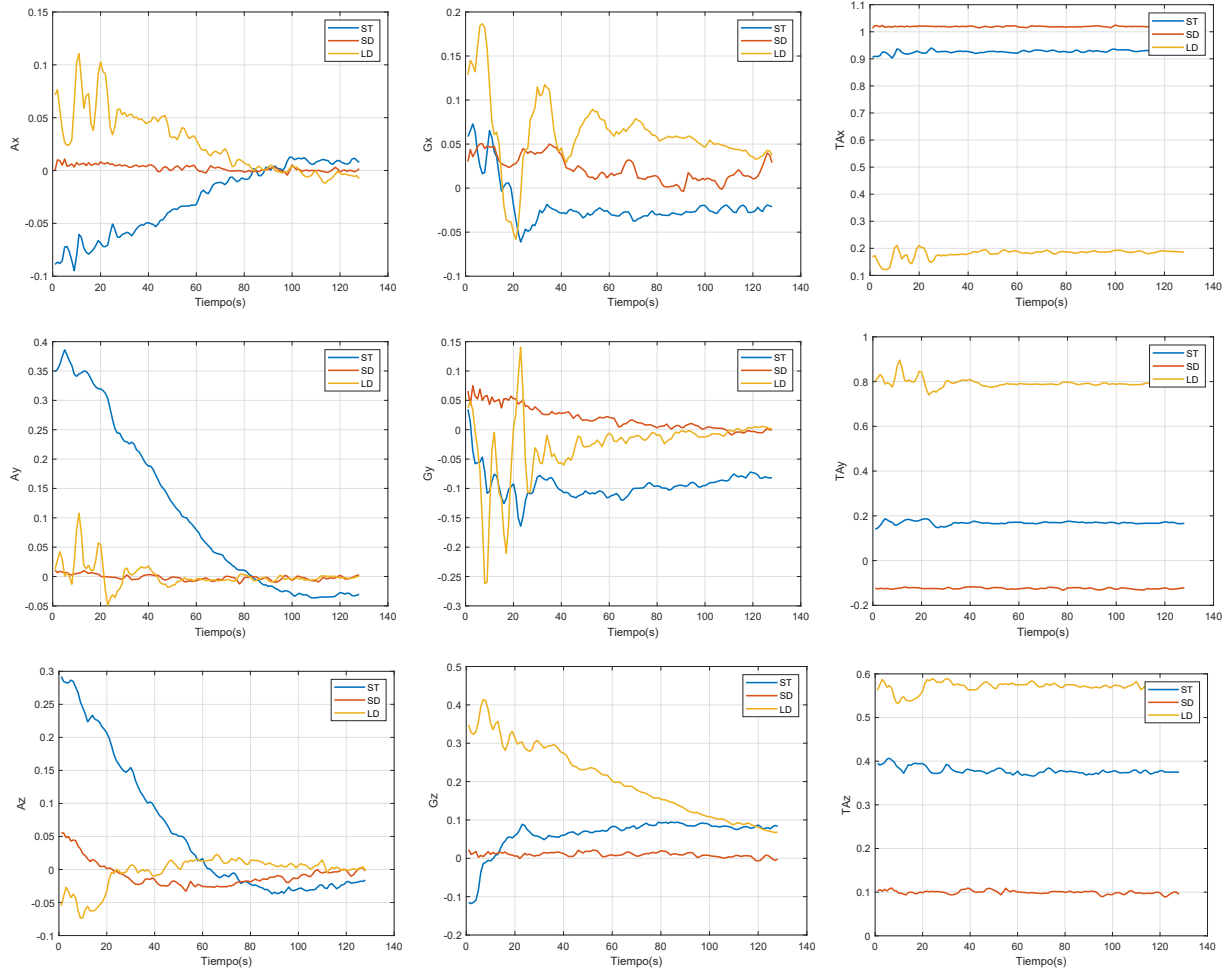


Figura 12: Ejemplo de señales de los sensores inerciales (acelerómetro en la columna izquierda, giroscopio en la columna central y aceleración total en la columna derecha) en cada eje (x en la primera fila, y en la fila central, y z en la última fila), para las actividades estáticas (ST en azul, SD en rojo y LD en amarillo).

de las técnicas EMD e ICEEMDAN, donde N dependen de la longitud de cada señal. Se decidió trabajar únicamente con las señales inerciales provenientes del acelerómetro y giroscopio. No se procesaron las señales correspondientes a la aceleración del cuerpo estimada ya que dichas señales son derivadas del acelerómetro. Se cálculo la cantidad de sujetos por cada IMF por cada clase.

En el diagrama de bloques de la figura 13 se muestra el procedimiento llevado a cabo en el análisis de las señales inerciales. Como medidas descriptivas para caracterizar las señales se determinó la varianza, la entropía y los rangos de frecuencia de las IMF de las señales inerciales de cada clase.

Se obtuvieron la cantidad registros que podían ser descompuestos en diferentes números de IMF para ambas técnicas.

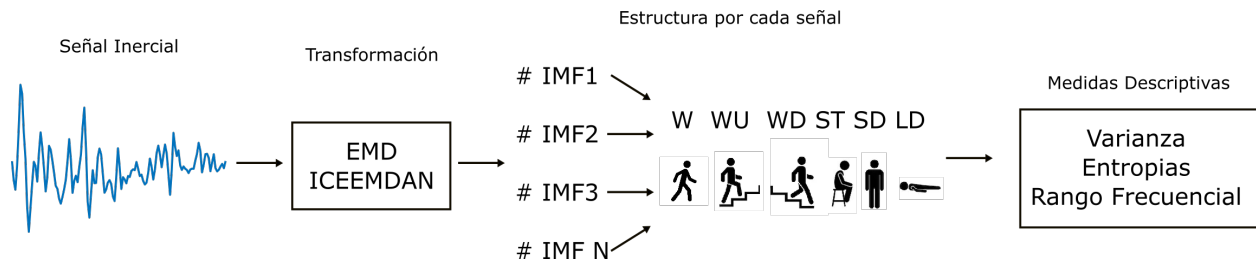


Figura 13: Diagrama de bloques para la caracterización de las señales inerciales.

Los datos fueron representados gráficamente por medio de barras de error con intervalos de confianza del 95 %.

4.2.2. Estimación del rango representativo de frecuencias en cada IMF

Para hallar el rango representativo de frecuencias en cada IMF, se estimó el espectro mediante la transformada discreta de Fourier, por medio del algoritmo de la transformada rápida de Fourier (FFT, por sus siglas en inglés *fast Fourier transform*), con una cantidad de 128 puntos, debido a la longitud de las señales. Luego, estas señales fueron suavizadas con un filtro de media móvil. En esta señal resultante, la componente de frecuencia más baja (f_L) y la más alta (f_H) para cada IMF se establecieron como aquellas frecuencias en las que la magnitud de su espectro de potencia cae por debajo del 20 % del valor más alto.

Para cada IMF y para cada clase, se calculó la media y la desviación estándar de los vectores de frecuencias $\vec{f}_L = [f_{L_1}, \dots, f_{L_M}]$ y $\vec{f}_H = [f_{H_1}, \dots, f_{H_M}]$, donde $M = 3752$ es el número de registros en el conjunto de datos utilizado.

4.2.3. Caracterización de las señales inerciales en cada IMF

Para caracterizar las señales inerciales en cada IMF, descompuestas por EMD e ICEEM-DAN, se utilizó la varianza y la entropía de Shannon calculadas en la totalidad cada señal, y para cada una de las seis clases.

Por cada IMF y por cada clase, se calculó la media y la desviación estándar de los vectores de varianza y entropía de Shannon, en concreto $\vec{\sigma}_{WK}^2 = \sigma_{WK_1}^2, \dots, \sigma_{WK_P}^2$, $\vec{\sigma}_{WU}^2 = \sigma_{WU_1}^2, \dots, \sigma_{WU_P}^2, \dots, \vec{\sigma}_{LD}^2 = \sigma_{LD_1}^2, \dots, \sigma_{LD_P}^2$, $\vec{H}_{WK} = H_{WK_1}, \dots, H_{WK_P}$, y $\vec{H}_{WU} = H_{WU_1}, \dots, H_{WU_P}, \dots, \vec{H}_{LD} = H_{LD_1}, \dots, H_{LD_P}$, donde $P = 10299$, que es total de registros de la base de datos (uniendo entrenamiento y prueba).

La varianza se utilizó como medida de la potencia de la señal y la entropía de Shannon se utilizó para medir la complejidad, según si su entropía es más alta representa sistemas complejos lo que lleva a una baja tasa de predictibilidad, y máxima cuando todos los valores de la serie están distribuidos uniformemente dentro de un rango fijo (Chang Francis, Ping-Yen, Hsuan-Hao, Long, y Chi, 2019).

4.3. Clasificación de las actividades humanas

En esta sección se presenta la metodología utilizada para reconocer las actividades caminar (WK), subir escaleras (WU), bajar escaleras (WD), estar parado (ST), estar acostado (LD) y estar sentado (SD).

A continuación, se describen las estrategias utilizadas para las etapas de transformación, preprocesamiento, clasificación y evaluación del modelo de clasificación de las señales inerciales. Estas etapas se mencionan en la figura 14

4.3.1. Representación en tiempo-frecuencia

Queriendo ampliar la observabilidad de las señales inerciales, se optó por implementar la EMD. A pesar de existen diferentes variantes de la EMD (familia EMD), se seleccionó la ICEEDMAN, debido a que sumando cada una de sus IMF, se logra un error de reconstrucción casi nulo de la señal. También se evita el mayor error de la EMD, la mezcla de modos, que es la aparición de señales más cortas a la original o señales intermitentes, en las cuales se puede perder información importante para el previo entrenamiento de la red.

De este modo, se descomponen los segmentos extraídos $x(t)$ en K número de IMF utilizando la ICEEMDAN: $IMF_1(t), IMF_2(t), \dots, IMF_K(t)$, donde K depende de cada segmento.

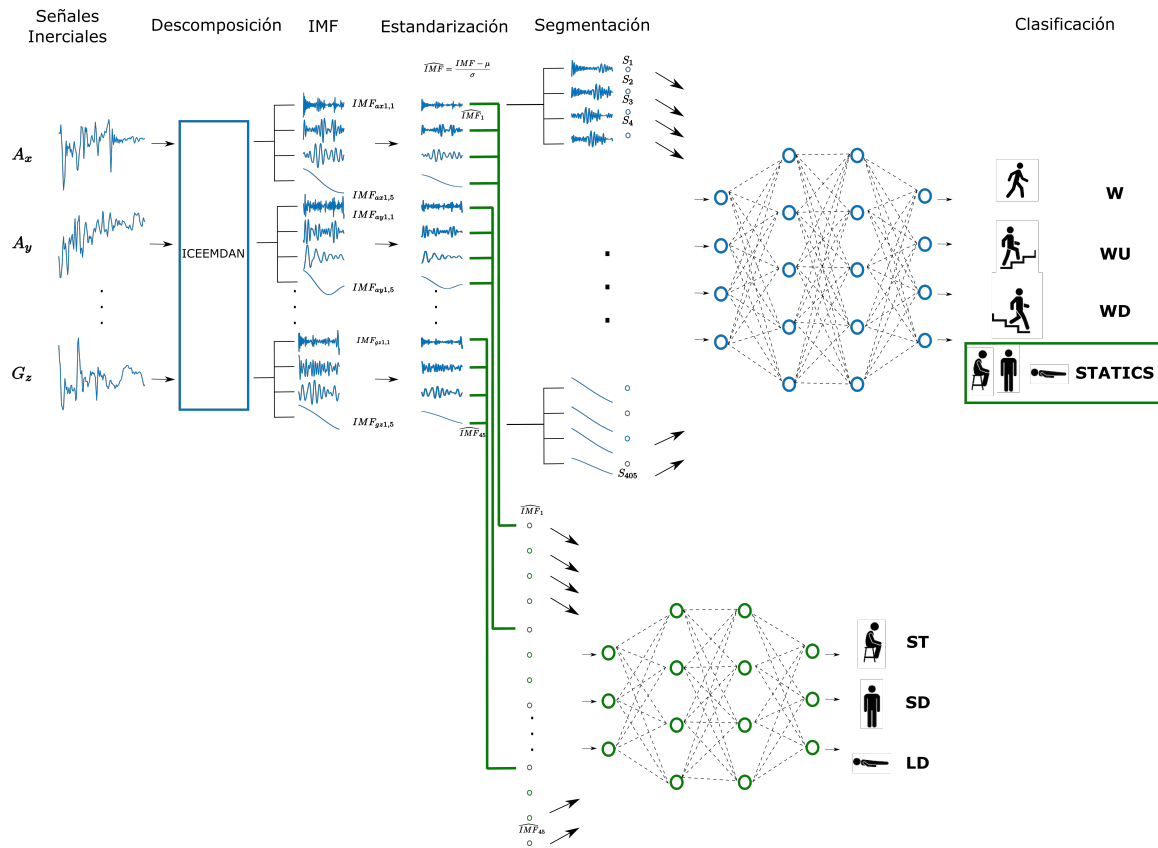


Figura 14: Procedimiento propuesto para reconocer las actividades humanas.

En cuanto a lo que concierne con los parámetros de entrada del algoritmo se escogió para la amplitud del ruido blanco $N_{std} = 0.2$, el número de realizaciones $NR = 5000$, y como máximo número de iteraciones para el proceso de shifting o cribado un valor de $MaxIter = 500$. En las primeras IMF se encuentran las oscilaciones de mayor frecuencia y conforme se extraen las últimas componentes, aparecen las oscilaciones de menor frecuencia y su residuo.

Debido al número diferente de IMF que se obtuvo para cada señal, había un desequilibrio en cuanto a observabilidad, ya que unas señales tenían más IMF que otras. Según lo anterior se optó por eliminar las señales que poseían más de cinco IMF, permitiendo así que todas tengan el mismo número de IMF.

4.3.2. Estandarización

Para hacer que un modelo de clasificación converja en una menor cantidad de iteraciones se realiza un escalamiento de los datos de entrada, en este caso las IMF, siendo la estandarización esta etapa de escalamiento. Se aplicó como estandarización los puntajes Z (del inglés z-scores), definido como $z = \frac{x-\mu}{\sigma}$, donde μ es la media de la población y σ es la desviación estándar de la población. En vista de que se desea que esta técnica de procesamiento sea aplicada en tiempo real, los valores de varianza y media para la mencionada estandarización extraídos son los mismos tanto para el conjunto de entrenamiento como el de prueba. Estos valores fueron determinados exclusivamente del conjunto de entrenamiento. Esto asegura la generalización del modelo clasificador.

4.3.3. Segmentación

Un modelo de aprendizaje profundo requiere de una gran cantidad de datos para obtener buenos resultados al momento de la clasificación. Sumado a esto la segmentación de señales es una etapa crucial en el proceso de reconocimiento de actividades humanas (Banos, Galvez, Damas, Pomares, y Rojas, 2014). En este trabajo se propone segmentar las IMF en ventanas de 1.28 s (64 muestras) con solapamiento del 50 %, este tiempo es suficiente para acoger un ciclo completo de las actividades que se desean clasificar, quedando así cada señal dividida en nueve partes. No obstante, la segmentación de la señal solo fue realizada para reconocer las actividades dinámicas (WK, WU y WD) dado que presentan un patrón repetitivo. Esto no ocurre en las señales estáticas (ST, SD y LD) dado que no hay un patrón repetitivo a extraer de la señal. La segmentación logra entonces que la señal de 2.56 s (128 muestras) sea observada en nueve segmentos consecutivos.

4.3.4. Clasificación

Ya que nuestras señales son series de tiempo, se propone un enfoque de redes LSTM, dado que estas al poseer una celda de memoria, almacenan información a lo largo del tiempo. Específicamente se utilizaron redes LSTM bidireccionales o redes BiLSTM, las cuales consideran datos anteriores y futuros. Se planteó realizar dos redes en cascada: una para clasificar las actividades dinámicas y la otra para clasificar las actividades estáticas. Como se aprecia en la figura 14 vista en la sección 4.3, nuestro trabajo conlleva dos partes de clasificación, la primera en la que se identifican las actividades dinámicas una por una, junto a las actividades estáticas todas en una sola clase, obteniendo así una red con una salida de cuatro clases: WK, WU, WD, estáticas (ST, SD, LD), y la segunda parte donde son identificadas cada una de las actividades estáticas que fueron clasificadas como un solo grupo en la primera parte, dando como resultado una segunda red con una salida de tres clases: ST, SD, LD. La arquitectura de la red para ambos modelos propuestos es la misma y se aprecia en la figura 15, sin embargo se diferencian en sus datos de entrada y parámetros.

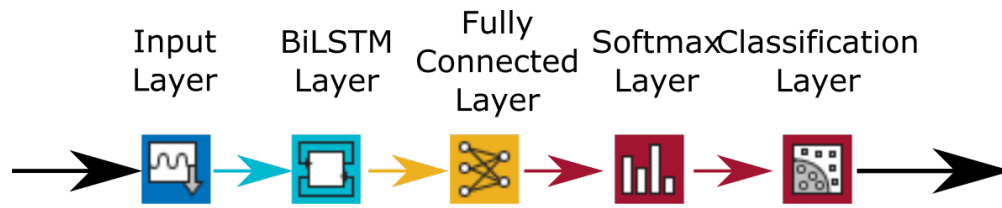


Figura 15: Arquitectura del modelo de clasificación propuesto.

A continuación se explica cada una de las cinco capas por la que está compuesto ambos modelos de clasificación:

- **Input layer:** Capa de entrada en la cual se ingresa la secuencia de datos a la red. Para nuestro caso recibe arreglos $9 \text{ señales} \times 5 \text{ IMF} \times 9 \text{ segmentos}$, en total serían 405 entradas para el modelo que reconoce las actividades dinámicas y 45 entradas para el modelo que reconoce las actividades estáticas.
- **BiLSTM layer:** Capa que aprende de forma bidireccional a largo plazo. Es útil cuando se quiere aprender de la serie completa de datos en cada paso de tiempo. Está compuesta por 175 neuronas.
- **Fully Connected Layer:** Una capa totalmente conectada que multiplica la entrada por una matriz de peso y luego agrega un vector de polarización. Posee seis neuronas debido a la cantidad de clases que se quiere clasificar.

- **Softmax Layer:** Normaliza los valores provenientes de la capa anterior en el un rango de 0 a 1, lo cual representa la probabilidad de cada una de las seis clases.
- **Classification Layer:** Compara la probabilidad con la etiqueta real para realimentar el proceso de aprendizaje del modelo.

Los parámetros utilizados para el diseño de la red son los siguientes:

- **Algoritmo de optimización:** Adam.
- **Learning rate:** un valor inicial de 0.001 con una tasa de caída de 0.1 cada 50 *epochs*
- **Batch size:** 1098 para las actividades dinámicas y 128 para las actividades estáticas.
- **Número de *epochs*:** 40.
- **Función de activación de las neuronas:** sigmoide y tanh para la celda de memoria de las redes LSTM bidireccionales.
- **Función de costo:** Entropía cruzada.

Se determinó la exactitud, precisión, exhaustividad de cada clasificador en el conjunto de prueba.

CAPÍTULO 5

RESULTADOS

En este capítulo se presentan los resultados obtenidos del presente proyecto. En primer lugar, en la sección 5.1 se presentan los resultados del análisis de las señales inerciales para reconocer seis actividades humanas, y, en segundo lugar, en la sección 5.2, se presentan los resultados de la clasificación de seis actividades humanas usando el enfoque basado en redes LSTM.

5.1. Análisis de las señales inerciales

El análisis de las señales inerciales se llevó a cabo tanto en las señales crudas (sin procesar) como en las IMF obtenidas por EMD y por ICEEMDAN.

5.1.1. Análisis de las señales inerciales crudas

En la tabla 2 se presentan los valores de entropía y varianza de las señales inerciales del giroscopio y acelerómetro, es decir, sin descomponerse en IMF, en cada uno de sus ejes y para cada clase de actividad humana. De esta tabla se puede observar que los mayores valores de varianza y entropía de las actividades dinámicas se encuentran en el eje x en ambos sensores inerciales, exceptuando la entropía en el giroscopio de la clase WD, cuyo valor más grande se encuentra en el eje z . Además, la magnitud de las actividades estáticas es menor que la de las dinámicas debido a que las actividades estáticas no involucran ni rotaciones ni movimientos significativos que los sensores puedan recolectar. Por otro lado, los mayores valores de entropía de las actividades estáticas se encuentran en los ejes y y z en ambos sensores, mientras que la varianza en el giroscopio es más grande en el eje Y . Finalmente, el valor de estas variables depende, en gran medida, de la posición en la que se encontraba el teléfono inteligente al momento de coleccionar las señales.

Tabla 2: Promedio \pm desviación estándar de los vectores de entropía (\vec{H}) y varianza ($\vec{\sigma}^2$, expresadas en $\times 10^{-3}$) de las señales inerciales (acelerómetro A y giroscopio G) crudas (sin procesar) en cada eje (x , y y z), por cada clase (WK, WU, WD, ST, SD y LD).

Parám	Clase					
	WK	WU	WD	ST	SD	LD
\vec{H}_{A_x}	0.91 \pm 0.16	0.80 \pm 0.16	0.76 \pm 0.16	1.21 \pm 0.39	1.26 \pm 0.36	0.95 \pm 0.44
\vec{H}_{A_y}	1.00 \pm 0.17	0.89 \pm 0.16	0.97 \pm 0.18	1.04 \pm 0.42	0.95 \pm 0.35	1.10 \pm 0.44
\vec{H}_{A_z}	0.96 \pm 0.21	0.76 \pm 0.20	1.02 \pm 0.21	1.11 \pm 0.41	1.070 \pm 0.40	1.22 \pm 0.40
\vec{H}_{G_x}	0.95 \pm 0.19	0.81 \pm 0.23	0.85 \pm 0.23	1.00 \pm 0.43	0.71 \pm 0.34	0.87 \pm 0.43
\vec{H}_{G_y}	1.04 \pm 0.22	0.89 \pm 0.26	1.17 \pm 0.22	0.87 \pm 0.36	0.92 \pm 0.35	0.91 \pm 0.40
\vec{H}_{G_z}	0.92 \pm 0.19	0.77 \pm 0.20	1.01 \pm 0.21	0.88 \pm 0.44	0.89 \pm 0.37	1.03 \pm 0.50
$\vec{\sigma}_{A_x}^2$	52.37 \pm 21.82	64.25 \pm 24.69	134.23 \pm 53.64	0.16 \pm 0.83	0.10 \pm 0.35	0.83 \pm 5.08
$\vec{\sigma}_{A_y}^2$	29.57 \pm 12.96	30.01 \pm 12.20	35.52 \pm 18.31	0.66 \pm 3.10	0.38 \pm 1.32	0.63 \pm 4.06
$\vec{\sigma}_{A_z}^2$	19.29 \pm 11.22	24.73 \pm 15.85	24.14 \pm 16.17	0.66 \pm 3.46	0.45 \pm 1.99	0.62 \pm 3.49
$\vec{\sigma}_{G_x}^2$	233.98 \pm 159.54	238.35 \pm 154.21	377.46 \pm 271.82	1.33 \pm 10.91	6.16 \pm 19.35	3.85 \pm 27.00
$\vec{\sigma}_{G_y}^2$	289.06 \pm 234.82	298.35 \pm 273.83	295.14 \pm 255.83	2.14 \pm 10.59	2.86 \pm 12.22	2.01 \pm 6.68
$\vec{\sigma}_{G_z}^2$	110.47 \pm 54.41	157.76 \pm 132.44	136.12 \pm 89.53	1.40 \pm 5.65	1.20 \pm 3.72	1.27 \pm 5.22

5.1.2. Análisis de las IMF de las señales inerciales

Número máximo de IMF por señal

Dado que el número máximo de IMF en el que una señal se puede descomponer es diferente para cada señal, en esta sección se analiza la distribución de los registros en cuanto al número máximo de IMF en el que las señales pueden ser descompuestas. La tabla 3 muestran la cantidad de registros, tanto del conjunto de entrenamiento como del conjunto de prueba, que pueden ser descompuestos en un máximo número de IMF, usando los métodos ICEEM-DAN y EMD. De esta tabla se puede observar que todos los registros pueden descomponerse en una y dos IMF. El menor valor del máximo número de IMF en la que los registros se pueden descomponer es tres IMF, tanto usando EMD como ICEEMDAN, no obstante, son muy pocos los registros que pueden ser descompuestos hasta tres IMF. De manera similar, muy pocos registros pueden descomponerse en ocho IMF. En cinco y seis IMF se encuentran la mayor cantidad de registros que pueden ser descompuestos hasta esos valores de IMF. Esta información es muy útil ya que nos permite conocer el máximo número de IMF en las que los registros pueden ser representados en IMF, para ser usados en el proceso de clasificación usando redes LSTM.

Tabla 3: Cantidad de registros que pueden ser descompuestos en un número máximo de IMF, tanto del conjunto de entrenamiento como del de prueba, usando las técnicas EMD e ICEEMDAN.

IMF máxima	EMD		ICEEMDAN	
	Entrenamiento	Prueba	Entrenamiento	Prueba
1	0	0	0	0
2	0	0	0	0
3	122	22	6	2
4	3010	1204	311	103
5	7300	2934	6372	2572
6	6601	2706	7305	2936
7	845	331	1901	722
8	14	2	42	15

Varianza y entropía de Shannon de las IMF por cada clase

En las tablas 4 a 7, se muestran los resultados de varianza y entropía de Shannon de las señales acelerómetro y giroscopio en cada IMF descompuestas con EMD. De manera similar, en las tablas 8 a 11, se muestran los resultados de varianza y entropía de Shannon de las señales acelerómetro y giroscopio en cada IMF descompuestas con ICEEMDAN. De estas tablas se desprende que las varianzas de las IMF del acelerómetro y del giroscopio usando EMD e ICEEMDAN son muy similares: las mayores varianzas de las actividades dinámicas se concentran en las IMF 2 y 3 con EMD, y en las IMF 3 y 4 con ICEEMDAN; las varianzas de las actividades estáticas son muy pequeñas. La entropía de Shannon de las IMF usando EMD e ICEEMDAN de las actividades estáticas también son muy similares, tienden a aumentar a medida que aumenta el índice de la IMF. En general, las menores entropías se encuentran en las IMF 1 y 2, mientras que las mayores entropías se encuentran en las IMF 4 y 5.

En las figuras 16 y 17 se muestran barras de error de la varianza de las IMF descompuestas por EMD e ICEEMDAN del acelerómetro y del giroscopio en el eje z , respectivamente, por cada actividad humana. Estas figuras concuerdan con lo mostrado en las tablas anteriores, en particular que las varianzas de las IMF del acelerómetro y del giroscopio se concentran en las primeras tres IMF con EMD, y en las IMF 2 a 5 con ICEEMDAN, y que las varianzas de las actividades estáticas son más pequeñas que las de las actividades dinámicas.

En las figuras 16 y 17 se muestran barras de error de las entropías de Shannon de las IMF descompuestas por EMD e ICEEMDAN del acelerómetro y del giroscopio en el eje x , respectivamente, por cada actividad humana. Estas figuras también concuerdan con lo detallado en las tablas anteriores, en particular que la entropía de Shannon de las IMF

Tabla 4: Varianza de las IMF del acelerómetro obtenidas con la técnica EMD. Los resultados están expresados en $\times 10^{-3}$.

Parám	WK	WU	WD	ST	SD	LD
$\bar{\sigma}_{A_{xIMF1}}^2$	10.91±8.87	7.78±7.11	13.19±10.98	0.01±0.04	0.02±0.11	0.02±0.11
$\bar{\sigma}_{A_{xIMF2}}^2$	16.40±13.17	21.70±21.07	54.77±52.57	0.01±0.11	0.02±0.07	0.02±0.08
$\bar{\sigma}_{A_{xIMF3}}^2$	24.26±14.67	34.99±18.80	73.67±42.05	0.01±0.28	0.01±0.04	0.04±0.42
$\bar{\sigma}_{A_{xIMF4}}^2$	3.60±4.65	5.91±8.22	11.43±14.87	0.05±0.52	0.01±0.16	0.52±4.54
$\bar{\sigma}_{A_{xIMF5}}^2$	0.68±1.00	1.12±1.54	2.00±3.13	0.04±0.44	0.02±0.19	0.20±1.75
$\bar{\sigma}_{A_{xIMF6}}^2$	0.36±0.82	0.55±0.98	0.96±1.78	0.00±0.04	0.00±0.04	0.02±0.21
$\bar{\sigma}_{A_{xIMF7}}^2$	0.24±0.34	0.39±0.66	0.68±1.04	0.00±0.00	0.00±0.02	0.11±0.38
$\bar{\sigma}_{A_{yIMF1}}^2$	9.03±5.93	4.72±4.07	9.57±8.20	0.01±0.06	0.02±0.09	0.02±0.14
$\bar{\sigma}_{A_{yIMF2}}^2$	8.60±6.72	8.11±7.26	11.66±8.49	0.02±0.09	0.06±0.18	0.04±0.45
$\bar{\sigma}_{A_{yIMF3}}^2$	10.89±7.44	13.47±8.87	11.22±9.98	0.11±1.14	0.05±0.19	0.11±1.83
$\bar{\sigma}_{A_{yIMF4}}^2$	3.27±3.55	5.84±6.27	4.75±4.66	0.31±2.25	0.12±0.89	0.24±2.31
$\bar{\sigma}_{A_{yIMF5}}^2$	0.71±1.20	1.47±2.32	1.69±2.18	0.13±1.00	0.10±0.73	0.23±3.00
$\bar{\sigma}_{A_{yIMF6}}^2$	0.31±0.47	0.65±1.13	0.73±1.15	0.03±0.29	0.03±0.11	0.05±0.58
$\bar{\sigma}_{A_{yIMF7}}^2$	0.18±0.31	0.13±0.22	0.48±0.54	0.00±0.01	0.03±0.07	0.15±0.74
$\bar{\sigma}_{A_{zIMF1}}^2$	4.66±4.39	1.94±2.07	6.08±5.93	0.01±0.04	0.02±0.05	0.02±0.08
$\bar{\sigma}_{A_{zIMF2}}^2$	3.94±3.69	4.02±5.20	6.02±5.99	0.02±0.07	0.02±0.05	0.03±0.36
$\bar{\sigma}_{A_{zIMF3}}^2$	7.38±6.67	9.90±10.84	7.26±9.47	0.09±1.58	0.05±4.62	0.04±0.51
$\bar{\sigma}_{A_{zIMF4}}^2$	3.77±3.92	9.16±9.11	5.21±4.58	0.31±2.25	0.14±0.71	0.26±2.53
$\bar{\sigma}_{A_{zIMF5}}^2$	1.08±1.91	2.29±4.01	1.64±2.19	0.15±0.98	0.17±1.29	0.20±1.64
$\bar{\sigma}_{A_{zIMF6}}^2$	0.40±0.78	0.82±1.32	0.79±1.50	0.03±0.28	0.04±0.18	0.06±0.73
$\bar{\sigma}_{A_{zIMF7}}^2$	0.26±0.46	0.84±1.22	0.71±1.51	0.05±0.18	0.05±0.12	0.04±0.15

Tabla 5: Varianza de las IMF del giroscopio obtenidas con la técnica EMD. Los resultados están expresados en $\times 10^{-3}$.

Parám	WK	WU	WD	ST	SD	LD
$\vec{\sigma}_{G^{xIMF_1}}^2$	49.33±41.64	26.54±29.47	43.28±47.73	0.07±0.42	0.47±5.48	0.48±6.30
$\vec{\sigma}_{G^{xIMF_2}}^2$	50.15±51.46	31.48±26.82	68.84±59.51	0.27±3.48	0.98±4.38	0.64±5.94
$\vec{\sigma}_{G^{xIMF_3}}^2$	72.27±87.50	53.34±63.49	75.24±108.56	0.44±6.50	1.87±8.94	1.05±9.27
$\vec{\sigma}_{G^{xIMF_4}}^2$	64.54±73.20	72.8±86.40	102.66±135.29	0.47±5.78	1.81±8.41	1.37±12.31
$\vec{\sigma}_{G^{xIMF_5}}^2$	15.86±48.67	65.06±91.18	100.32±133.61	0.15±0.69	0.82±3.40	0.83±7.48
$\vec{\sigma}_{G^{xIMF_6}}^2$	7.00±12.29	47.66±75.53	56.84±99.90	0.05±0.43	0.76±4.09	0.29±2.25
$\vec{\sigma}_{G^{xIMF_7}}^2$	2.95±4.32	1.99±2.82	53.84±65.94	0.06±0.15	0.66±1.35	0.37±1.57
$\vec{\sigma}_{G^{yIMF_1}}^2$	99.32±160.52	45.25±60.28	113.88±152.46	0.15±0.57	0.45±2.53	0.41±1.97
$\vec{\sigma}_{G^{yIMF_2}}^2$	100.4±102.34	42.11±52.03	90.68±101.92	0.41±4.19	0.60±2.79	0.35±2.06
$\vec{\sigma}_{G^{yIMF_3}}^2$	66.47±68.26	112.26±158.02	61.67±63.22	0.52±4.49	0.69±3.76	0.32±1.39
$\vec{\sigma}_{G^{yIMF_4}}^2$	49.55±68.87	101.94±151.95	36.83±41.26	0.49±3.08	0.63±4.28	0.43±1.37
$\vec{\sigma}_{G^{yIMF_5}}^2$	7.96±18.22	19.24±45.31	14.34±19.72	0.57±4.14	0.37±1.71	0.40±1.34
$\vec{\sigma}_{G^{yIMF_6}}^2$	3.26±10.18	7.05±14.08	7.97±15.65	0.3±1.34	0.37±1.89	0.27±1.27
$\vec{\sigma}_{G^{yIMF_7}}^2$	1.45±1.96	1.89±1.75	1.42±1.25	0.21±0.80	0.93±2.53	0.33±1.08
$\vec{\sigma}_{G^{zIMF_1}}^2$	20.57±19.33	15.19±16.16	36.41±39.79	0.06±0.45	0.11±0.42	0.06±0.49
$\vec{\sigma}_{G^{zIMF_2}}^2$	55.36±40.00	25.14±36.27	54.92±50.22	0.14±1.27	0.26±1.14	0.12±1.26
$\vec{\sigma}_{G^{zIMF_3}}^2$	31.59±25.09	72.99±97.65	37.02±38.97	0.34±2.37	0.26±0.88	0.55±10.50
$\vec{\sigma}_{G^{zIMF_4}}^2$	13.91±16.49	45.89±64.02	16.18±20.36	0.41±1.95	0.29±1.10	0.75±13.82
$\vec{\sigma}_{G^{zIMF_5}}^2$	3.11±5.18	11.94±25.33	7.01±11.21	3.68±1.73	0.20±0.73	0.30±1.47
$\vec{\sigma}_{G^{zIMF_6}}^2$	1.22±3.13	3.67±6.07	4.22±12.2	0.16±1.10	0.20±1.26	0.12±0.70
$\vec{\sigma}_{G^{zIMF_7}}^2$	0.85±0.88	1.44±0.84	2.97±6.73	0.10±0.37	0.48±1.71	0.52±2.55

Tabla 6: Entropía de Shannon de las IMF del acelerómetro obtenidas con la técnica EMD.

Parám	WK	WU	WD	ST	SD	LD
$\vec{H}_{A_x IMF_1}$	1.88±0.15	1.79±0.17	1.86±0.16	1.81±0.23	1.79±0.3	1.81±0.23
$\vec{H}_{A_x IMF_2}$	2.06±0.12	2.09±0.13	2.08±0.13	2.007±0.19	1.98±0.20	2.01±0.19
$\vec{H}_{A_x IMF_3}$	2.21±0.07	2.21±0.07	2.21±0.08	2.11±0.15	2.11±0.15	2.11±0.16
$\vec{H}_{A_x IMF_4}$	2.18±0.08	2.17±0.09	2.17±0.09	2.16±0.11	2.17±0.10	2.14±0.12
$\vec{H}_{A_x IMF_5}$	2.1±0.12	2.1±0.12	2.11±0.12	2.11±0.12	2.11±0.12	2.11±0.12
$\vec{H}_{A_x IMF_6}$	2.15±0.11	2.15±0.12	2.15±0.12	2.15±0.12	2.13±0.12	2.16±0.12
$\vec{H}_{A_x IMF_7}$	2.16±0.13	2.19±0.13	2.16±0.12	2.17±0.13	2.19±0.11	2.17±0.1
$\vec{H}_{A_y IMF_1}$	1.88±0.14	1.88±0.17	1.82±0.17	1.81±0.24	1.8±0.24	1.8±0.23
$\vec{H}_{A_y IMF_2}$	2.1±0.11	2.1±0.11	2.08±0.12	1.99±0.2	1.98±0.2	2±0.19
$\vec{H}_{A_y IMF_3}$	2.21±0.07	2.21±0.06	2.18±0.09	2.1±0.15	2.11±0.15	2.11±0.15
$\vec{H}_{A_y IMF_4}$	2.19±0.07	2.18±0.08	2.19±0.08	2.14±0.12	2.16±0.1	2.16±0.11
$\vec{H}_{A_y IMF_5}$	2.1±0.12	2.1±0.12	2.11±0.12	2.1±0.13	2.1±0.12	2.11±0.12
$\vec{H}_{A_y IMF_6}$	2.13±0.11	2.16±0.11	2.14±0.11	2.16±0.13	2.17±0.12	2.16±0.12
$\vec{H}_{A_y IMF_7}$	2.12±0.13	2.2±0.12	2.17±0.13	2.18±0.11	2.18±0.08	2.2±0.12
$\vec{H}_{A_z IMF_1}$	1.85±0.17	1.84±0.17	1.81±0.17	1.85±0.2	1.84±0.21	1.84±0.21
$\vec{H}_{A_z IMF_2}$	2.08±0.13	2.05±0.13	2.05±0.13	2.02±0.17	2.02±0.17	2.03±0.17
$\vec{H}_{A_z IMF_3}$	2.2±0.08	2.18±0.09	2.17±0.1	2.12±0.14	2.11±0.14	2.12±0.14
$\vec{H}_{A_z IMF_4}$	2.2±0.06	2.19±0.08	2.19±0.07	2.16±0.11	2.15±0.1	2.16±0.1
$\vec{H}_{A_z IMF_5}$	2.1±0.11	2.1±0.11	2.11±0.12	2.11±0.12	2.1±0.12	2.1±0.13
$\vec{H}_{A_z IMF_6}$	2.13±0.12	2.15±0.11	2.13±0.12	2.16±0.11	2.17±0.11	2.16±0.12
$\vec{H}_{A_z IMF_7}$	2.2±0.1	2.16±0.12	2.14±0.14	2.13±0.15	2.14±0.13	2.19±0.1

Tabla 7: Entropía de Shannon de las IMF del giroscopio obtenidas con la técnica EMD

Parám	WK	WU	WD	ST	SD	LD
$\vec{H}_{G_x IMF_1}$	1.94±0.16	1.91±0.15	1.87±0.16	1.8±0.3	1.74±0.28	1.76±0.27
$\vec{H}_{G_x IMF_2}$	2.12±0.1	2.1±0.12	2.1±0.12	1.95±0.22	1.94±0.22	1.96±0.22
$\vec{H}_{G_x IMF_3}$	2.18±0.09	2.17±0.1	2.17±0.1	2.1±0.17	2.11±0.14	2.09±0.16
$\vec{H}_{G_x IMF_4}$	2.17±0.1	2.15±0.1	2.15±0.11	2.16±0.1	2.15±0.1	2.15±0.11
$\vec{H}_{G_x IMF_5}$	2.09±0.12	2.1±0.12	2.09±0.13	2.1±0.12	2.11±0.12	2.11±0.12
$\vec{H}_{G_x IMF_6}$	2.18±0.11	2.15±0.12	2.17±0.11	2.15±0.11	2.15±0.13	2.16±0.12
$\vec{H}_{G_x IMF_7}$	2.18±0.19	2.28±0.004	2.08±0.13	2.14±0.13	2.13±0.16	2.14±0.17
$\vec{H}_{G_y IMF_1}$	1.89±0.15	1.88±0.16	1.84±0.18	1.77±0.3	1.76±0.26	1.76±0.26
$\vec{H}_{G_y IMF_2}$	2.13±0.11	2.07±0.12	2.07±0.13	1.96±0.22	1.94±0.22	1.96±0.23
$\vec{H}_{G_y IMF_3}$	2.2±0.07	2.18±0.09	2.17±0.1	2.09±0.16	2.09±0.16	2.09±0.16
$\vec{H}_{G_y IMF_4}$	2.19±0.08	2.17±0.1	2.19±0.08	2.16±0.11	2.15±0.1	2.14±0.11
$\vec{H}_{G_y IMF_5}$	2.09±0.12	2.09±0.11	2.1±0.12	2.1±0.12	2.1±0.12	2.11±0.12
$\vec{H}_{G_y IMF_6}$	2.16±0.12	2.15±0.11	2.15±0.11	2.15±0.12	2.15±0.11	2.15±0.12
$\vec{H}_{G_y IMF_7}$	2.15±0.12	2.21±0.08	2.12±0.12	2.17±0.11	2.15±0.13	2.14±0.12
$\vec{H}_{G_z IMF_1}$	1.91±0.15	1.92±0.16	1.87±0.17	1.8±0.3	1.78±0.26	1.81±0.3
$\vec{H}_{G_z IMF_2}$	2.16±0.11	2.09±0.13	2.11±0.12	1.96±0.22	1.96±0.21	1.99±0.2
$\vec{H}_{G_z IMF_3}$	2.19±0.08	2.19±0.088	2.19±0.09	2.1±0.17	2.1±0.15	2.11±0.16
$\vec{H}_{G_z IMF_4}$	2.16±0.11	2.15±0.1	2.18±0.08	2.16±0.1	2.15±0.11	2.16±0.1
$\vec{H}_{G_z IMF_5}$	2.11±0.12	2.09±0.12	2.09±0.12	2.1±0.12	2.1±0.12	2.11±0.11
$\vec{H}_{G_z IMF_6}$	2.15±0.12	2.15±0.11	2.16±0.12	2.16±0.12	2.16±0.11	2.15±0.12
$\vec{H}_{G_z IMF_7}$	2.09±0.14	2.1±0.17	2.21±0.08	2.16±0.12	2.15±0.16	2.19±0.1

Tabla 8: Varianza de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN. Los resultados están expresados en $\times 10^{-3}$.

Parám	WK	WU	WD	ST	SD	LD
$\bar{\sigma}_{A_{xIMF_1}}^2$	3.82±3.25	1.62±1.57	3.70±.33	0±0.02	0.01±0.05	0.01±0.05
$\bar{\sigma}_{A_{xIMF_2}}^2$	7.66±5.21	7.14±5.37	13.38±9.86	0.01±0.03	0.01±0.06	0.01±0.05
$\bar{\sigma}_{A_{xIMF_3}}^2$	24.76±15.04	35.91±21.57	71.50±47.12	0±0.05	0.01±0.03	0.01±0.10
$\bar{\sigma}_{A_{xIMF_4}}^2$	4.53±4.31	7.57±8.59	18.88±18.80	0.02±0.32	0±0.02	0.20±3.07
$\bar{\sigma}_{A_{xIMF_5}}^2$	0.59±0.73	1.11±1.30	1.59±2.00	0.06±0.53	0.02±0.22	0.35±3.02
$\bar{\sigma}_{A_{xIMF_6}}^2$	0.23±0.31	0.47±.70	0.76±1.05	0.01±0.15	0±0.05	0.20±2.19
$\bar{\sigma}_{A_{xIMF_7}}^2$	0.14±0.33	0.20±0.36	0.36±0.59	0±0.01	0.01±0.10	0±0.01
$\bar{\sigma}_{A_{yIMF_1}}^2$	3.54±2.52	1.11±1.20	3.16±2.94	0±0.03	0±0.03	0.01±0.07
$\bar{\sigma}_{A_{yIMF_2}}^2$	5.99±4.43	4.16±3.14	7.60±5.51	0.01±0.04	0.03±0.11	0.03±0.30
$\bar{\sigma}_{A_{yIMF_3}}^2$	8.51±5.75	9.37±7.41	9.43±7.84	0.02±0.08	0.03±0.11	0.03±0.24
$\bar{\sigma}_{A_{yIMF_4}}^2$	3.96±3.63	7.00±5.65	5.79±4.66	0.08±1.00	0.05±0.56	0.10±1.80
$\bar{\sigma}_{A_{yIMF_5}}^2$	1.13±1.41	2.81±3.33	2.06±2.41	0.33±2.13	0.12±0.91	0.30±2.82
$\bar{\sigma}_{A_{yIMF_6}}^2$	0.28±0.42	0.74±1.19	0.77±1.09	0.12±0.97	0.06±0.27	0.10±1.18
$\bar{\sigma}_{A_{yIMF_7}}^2$	0.09±0.18	0.19±0.45	0.28±0.52	0.01±0.09	0.02±0.07	0.03±0.19
$\bar{\sigma}_{A_{zIMF_1}}^2$	2.40±2.66	0.63±0.82	3.07±3.33	0±0.03	0±0.02	0.01±0.03
$\bar{\sigma}_{A_{zIMF_2}}^2$	2.12±1.52	1.73±1.80	3.79±3.26	0.01±0.02	0.01±0.03	0.02±0.24
$\bar{\sigma}_{A_{zIMF_3}}^2$	5.64±5.97	4.89±5.86	5.13±7.16	0.01±0.07	0.01±0.06	0.01±0.08
$\bar{\sigma}_{A_{zIMF_4}}^2$	3.23±2.97	8.46±8.02	5.00±4.20	0.03±0.28	0.05±0.43	0.04±0.54
$\bar{\sigma}_{A_{zIMF_5}}^2$	1.74±2.43	4.17±5.73	2.01±2.29	0.34±2.37	0.17±0.97	0.34±2.75
$\bar{\sigma}_{A_{zIMF_6}}^2$	0.32±0.57	0.79±1.32	0.81±1.34	0.13±0.87	0.13±1.13	0.11±1.03
$\bar{\sigma}_{A_{zIMF_7}}^2$	0.07±0.10	0.14±0.36	0.35±0.66	0.06±0.37	0.02±0.08	0.02±0.16

Tabla 9: Varianza de las IMF del giroscopio obtenidas con la técnica ICEEMDAN. Los resultados están expresados en $\times 10^{-3}$.

Parám	WK	WU	WD	ST	SD	LD
$\bar{\sigma}_{G^x IMF_1}^2$	13.30±11.79	5.56±8.86	10.62±15.72	0.03±0.24	0.06±0.30	0.08±0.70
$\bar{\sigma}_{G^x IMF_2}^2$	39.13±33.84	19.63±18.33	35.66±33.26	0.07±0.49	0.33±1.12	0.29±3.40
$\bar{\sigma}_{G^x IMF_3}^2$	26.68±26.80	22.07±18.01	43.33±31.68	0.19±2.36	0.75±3.14	0.53±3.95
$\bar{\sigma}_{G^x IMF_4}^2$	86.13±83.83	52.01±60.05	75.44±118.04	0.3±2.46	1.41±5.50	0.64±5.57
$\bar{\sigma}_{G^x IMF_5}^2$	16.49±32.94	61.82±81.28	94.90±126.15	0.33±2.19	1.31±5.58	1.01±8.10
$\bar{\sigma}_{G^x IMF_6}^2$	9.52±37.00	56.51±82.41	75.52±113.03	0.16±1.70	0.95±5.28	0.63±3.76
$\bar{\sigma}_{G^x IMF_7}^2$	1.66±4.42	25.00±44.64	47.66±98.90	0.04±0.22	0.77±4.87	0.79±5.04
$\bar{\sigma}_{G^y IMF_1}^2$	62.35±125.88	16.24±22.95	65.02±96.85	0.06±0.30	0.20±1.28	0.21±1.15
$\bar{\sigma}_{G^y IMF_2}^2$	41.31±46.12	25.56±37.52	58.45±78.93	0.11±0.41	0.39±2.31	0.22±1.19
$\bar{\sigma}_{G^y IMF_3}^2$	70.87±70.09	33.24±34.03	53.78±47.62	0.24±1.34	0.51±2.89	0.27±1.88
$\bar{\sigma}_{G^y IMF_4}^2$	58.78±69.09	133.00±164.98	35.17±34.86	0.36±2.82	0.53±3.74	0.24±1.01
$\bar{\sigma}_{G^y IMF_5}^2$	7.60±15.05	33.82±66.44	15.83±21.42	0.70±5.08	0.42±1.89	0.41±1.23
$\bar{\sigma}_{G^y IMF_6}^2$	2.86±6.89	10.80±21.35	8.06±14.09	0.34±1.79	0.30±1.31	0.35±1.41
$\bar{\sigma}_{G^y IMF_7}^2$	0.82±1.92	3.58±7.83	2.70±7.81	0.13±0.38	0.18±0.75	0.10±0.29
$\bar{\sigma}_{G^z IMF_1}^2$	7.53±11.16	3.54±6.37	12.92±17.74	0.02±0.15	0.03±0.17	0.02±0.11
$\bar{\sigma}_{G^z IMF_2}^2$	19.10±16.71	10.70±11.06	29.28±34.26	0.04±0.17	0.12±0.48	0.05±0.38
$\bar{\sigma}_{G^z IMF_3}^2$	35.48±21.95	23.52±31.43	35.99±34.52	0.11±1.14	0.17±0.65	0.10±0.58
$\bar{\sigma}_{G^z IMF_4}^2$	22.98±22.75	70.45±92.14	18.96±20.71	0.28±1.62	0.23±1.04	0.29±1.36
$\bar{\sigma}_{G^z IMF_5}^2$	4.00±5.74	19.60±29.12	8.20±11.82	0.46±1.87	0.27±1.03	0.49±2.27
$\bar{\sigma}_{G^z IMF_6}^2$	1.31±2.35	6.23±12.88	4.76±11.43	0.35±2.13	0.19±1.09	0.20±1.16
$\bar{\sigma}_{G^z IMF_7}^2$	0.39±1.09	1.72±2.98	1.02±2.20	0.07±0.45	0.15±0.82	0.23±1.15

Tabla 10: Entropía de Shannon de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN.

Parám	WK	WU	WD	ST	SD	LD
$\vec{H}_{A_x IMF_1}$	1.60±0.19	1.5±0.23	1.62±0.21	1.69±0.27	1.69±0.28	1.66±0.29
$\vec{H}_{A_x IMF_2}$	2.06±0.11	2.02±0.13	2.01±0.13	1.97±0.20	1.95±0.21	1.95±0.20
$\vec{H}_{A_x IMF_3}$	2.20±0.06	2.22±0.07	2.19±0.08	2.08±0.17	2.09±0.15	2.09±0.17
$\vec{H}_{A_x IMF_4}$	2.20±0.08	2.18±0.09	2.18±0.10	2.15±0.13	2.16±0.12	2.11±0.15
$\vec{H}_{A_x IMF_5}$	2.14±0.11	2.11±0.13	2.15±0.10	2.11±0.12	2.13±0.11	2.10±0.12
$\vec{H}_{A_x IMF_6}$	2.1±0.12	2.11±0.13	2.09±0.12	2.12±0.12	2.11±0.12	2.10±0.13
$\vec{H}_{A_x IMF_7}$	2.16±0.11	2.17±0.11	2.16±0.14	2.18±0.13	2.15±0.14	2.16±0.14
$\vec{H}_{A_y IMF_1}$	1.61±0.17	1.57±0.21	1.55±0.21	1.69±0.27	1.67±0.28	1.69±0.27
$\vec{H}_{A_y IMF_2}$	2.08±0.11	2.06±0.12	2.02±0.13	1.93±0.23	1.92±0.23	1.95±0.22
$\vec{H}_{A_y IMF_3}$	2.18±0.08	2.18±0.08	2.16±0.09	2.06±0.18	2.07±0.17	2.08±0.17
$\vec{H}_{A_y IMF_4}$	2.2±0.08	2.20±0.07	2.20±0.08	2.13±0.15	2.15±0.12	2.14±0.12
$\vec{H}_{A_y IMF_5}$	2.16±0.11	2.14±0.10	2.16±0.10	2.10±0.12	2.12±0.1	2.11±0.11
$\vec{H}_{A_y IMF_6}$	2.09±0.12	2.10±0.11	2.1±0.11	2.11±0.13	2.12±0.12	2.12±0.13
$\vec{H}_{A_y IMF_7}$	2.16±0.12	2.12±0.12	2.17±0.14	2.17±0.12	2.17±0.14	2.16±0.14
$\vec{H}_{A_z IMF_1}$	1.58±0.19	1.63±0.21	1.56±0.19	1.72±0.22	1.71±0.24	1.72±0.30
$\vec{H}_{A_z IMF_2}$	2.05±0.10	2.01±0.13	2.00±0.13	1.97±0.18	1.98±0.18	1.99±0.19
$\vec{H}_{A_z IMF_3}$	2.17±0.09	2.14±0.11	2.15±0.10	2.09±0.15	2.08±0.15	2.09±0.15
$\vec{H}_{A_z IMF_4}$	2.20±0.08	2.20±0.06	2.20±0.08	2.15±0.13	2.14±0.12	2.14±0.13
$\vec{H}_{A_z IMF_5}$	2.16±0.09	2.15±0.1	2.15±0.11	2.11±0.11	2.12±0.11	2.11±0.12
$\vec{H}_{A_z IMF_6}$	2.09±0.12	2.09±0.12	2.09±0.12	2.12±0.13	2.12±0.12	2.12±0.13
$\vec{H}_{A_z IMF_7}$	2.11±0.13	2.13±0.12	2.15±0.13	2.15±0.14	2.17±0.12	2.16±0.14

Tabla 11: Entropía de Shannon de las IMF del giroscopio obtenidas con la técnica ICEEMDAN.

Parám	WK	WU	WD	ST	SD	LD
$\vec{H}_{G_{xIMF_1}}$	1.64±0.18	1.57±0.22	1.58±0.22	1.66±0.29	1.64±0.3	1.6±0.32
$\vec{H}_{G_{xIMF_2}}$	2.10±0.1	2.06±0.12	2.02±0.13	1.89±0.26	1.82±0.28	1.88±0.27
$\vec{H}_{G_{xIMF_3}}$	2.16±0.1	2.12±0.12	2.16±0.09	2.05±0.2	2.03±0.19	2.04±0.2
$\vec{H}_{G_{xIMF_4}}$	2.2±0.07	2.18±0.09	2.17±0.09	2.14±0.14	2.15±0.12	2.12±0.15
$\vec{H}_{G_{xIMF_5}}$	2.1±0.13	2.11±0.11	2.1±0.12	2.13±0.11	2.12±0.1	2.12±0.11
$\vec{H}_{G_{xIMF_6}}$	2.11±0.12	2.12±0.12	2.11±0.13	2.11±0.13	2.11±0.13	2.12±0.13
$\vec{H}_{G_{xIMF_7}}$	2.07±0.18	2.16±0.13	2.19±0.09	2.14±0.14	2.17±0.11	2.15±0.15
$\vec{H}_{G_{yIMF_1}}$	1.62±0.2	1.58±0.21	1.6±0.22	1.57±0.29	1.6±0.31	1.56±0.31
$\vec{H}_{G_{yIMF_2}}$	2.06±0.11	2.02±0.14	2.02±0.13	1.9±0.3	1.87±0.3	1.91±0.24
$\vec{H}_{G_{yIMF_3}}$	2.21±0.06	2.14±0.1	2.15±0.1	2.02±0.21	2.03±0.19	2.03±0.21
$\vec{H}_{G_{yIMF_4}}$	2.2±0.06	2.2±0.07	2.2±0.08	2.14±0.13	2.14±0.13	2.11±0.15
$\vec{H}_{G_{yIMF_5}}$	2.11±0.13	2.12±0.12	2.14±0.1	2.12±0.1	2.13±0.1	2.12±0.1
$\vec{H}_{G_{yIMF_6}}$	2.11±0.13	2.1±0.11	2.11±0.11	2.11±0.12	2.1±0.13	2.1±0.13
$\vec{H}_{G_{yIMF_7}}$	2.13±0.14	2.15±0.1	2.16±0.13	2.15±0.12	2.15±0.13	2.12±0.13
$\vec{H}_{G_{zIMF_1}}$	1.62±0.2	1.55±0.21	1.54±0.22	1.65±0.3	1.67±0.29	1.69±0.3
$\vec{H}_{G_{zIMF_2}}$	2.09±0.11	2.03±0.13	2.05±0.12	1.9±0.26	1.89±0.24	1.93±0.23
$\vec{H}_{G_{zIMF_3}}$	2.21±0.06	2.13±0.11	2.17±0.08	2.03±0.21	2.04±0.19	2.07±0.19
$\vec{H}_{G_{zIMF_4}}$	2.2±0.06	2.2±0.07	2.2±0.07	2.13±0.14	2.14±0.12	2.14±0.13
$\vec{H}_{G_{zIMF_5}}$	2.12±0.12	2.12±0.12	2.13±0.11	2.11±0.11	2.12±0.1	2.12±0.11
$\vec{H}_{G_{zIMF_6}}$	2.1±0.12	2.11±0.11	2.1±0.12	2.11±0.13	2.11±0.13	2.11±0.12
$\vec{H}_{G_{zIMF_7}}$	2.12±0.13	2.14±0.13	2.15±0.14	2.16±0.12	2.12±0.16	2.17±0.13

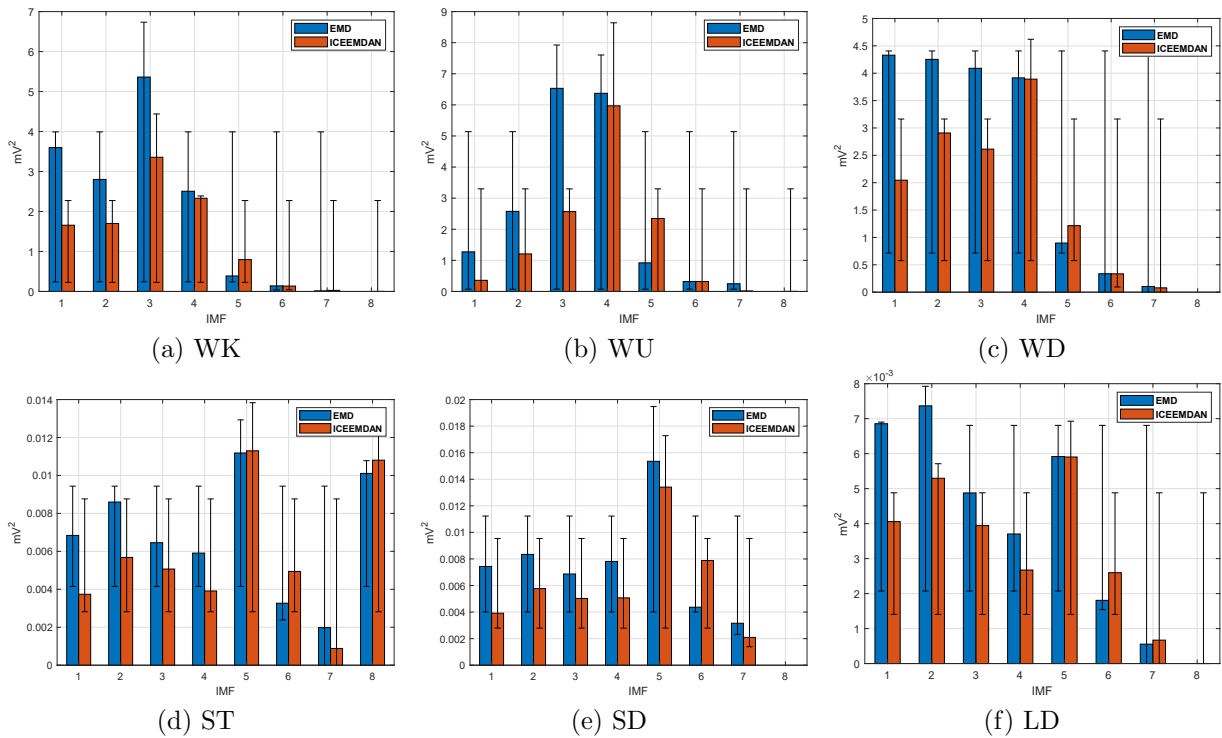


Figura 16: Barras de error de las varianzas de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del acelerómetro en el eje z para cada actividad.

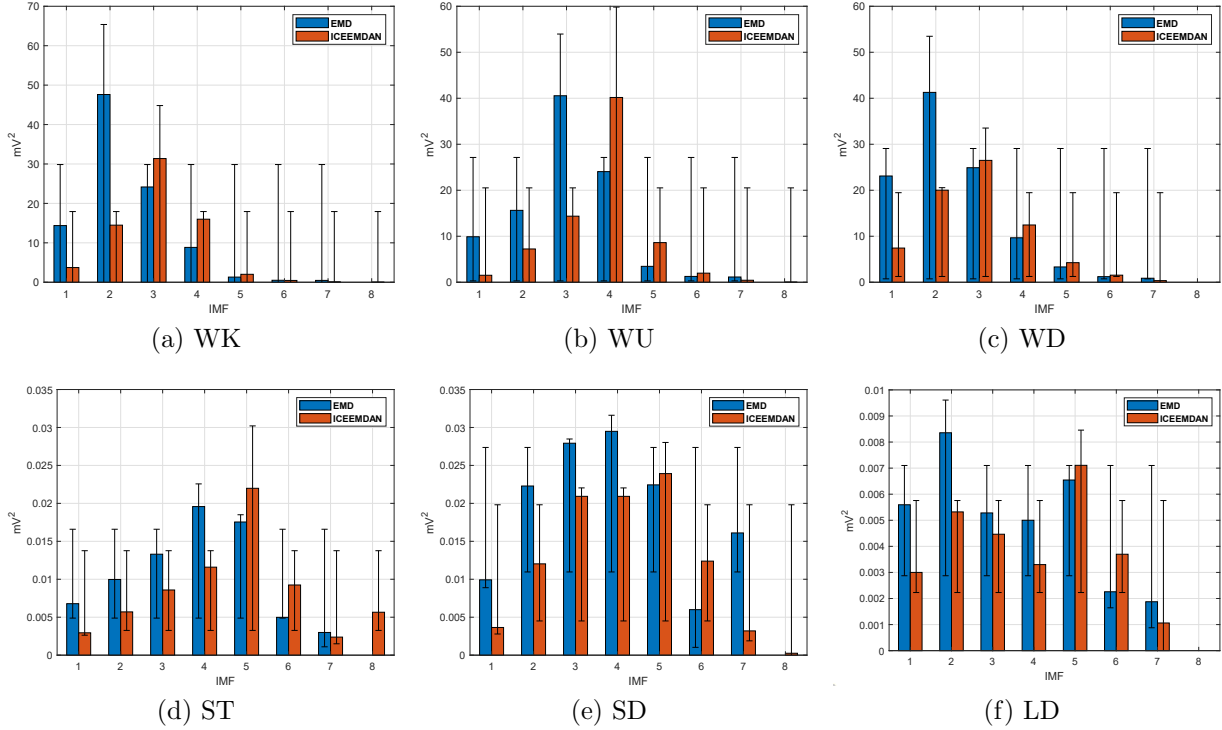


Figura 17: Barras de error de las varianzas de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del giroscopio en el eje z para cada actividad.

usando EMD e ICEEMDAN tienden a aumentar a medida que aumenta el índice de la IMF y que las menores entropías se encuentran en las IMF 1 y 2, mientras que las mayores entropías se encuentran en las IMF 4 y 7.

Rango de frecuencia de las IMF por cada clase

En las tablas 12 a 13, se presentan las bajas y altas frecuencias (f_L y f_H) representativas del acelerómetro y del giroscopio en cada IMF descompuestas con EMD. Del mismo modo, en las tablas 14 a 15, se muestran las frecuencias en cada IMF descompuestas con ICEEMDAN. De estas tablas se desprende que, los rangos de frecuencia son mayores en las primeras IMF y, para IMF mayores, este rango es menor, evidenciándose de esta manera lo que estipula la teoría de descomposición empírica en modos. Por otro lado, las bajas frecuencias son cercanas a cero, excepto en la primera IMF de ICEEMDAN donde son un poco mayores que uno. Como era de esperarse, las últimas IMF no aportan al rango frecuencial ya que corresponden al residuo de la descomposición.

Según la literatura, las actividades humanas se encuentran por debajo de los 18 Hz (Z. Wang y cols., 2016), lo cual concuerda con nuestros resultados, sin embargo, se encuentra que las

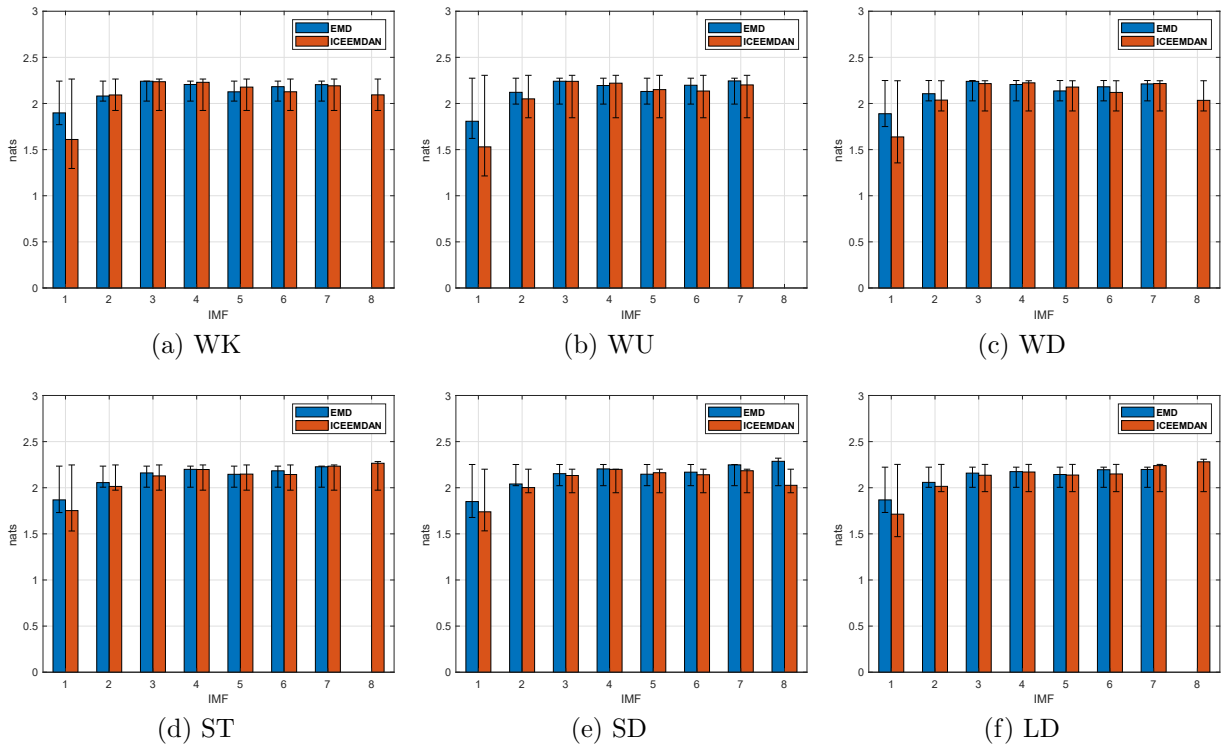


Figura 18: Barras de error de las entropías de Shannon de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del acelerómetro en el eje x para cada actividad.

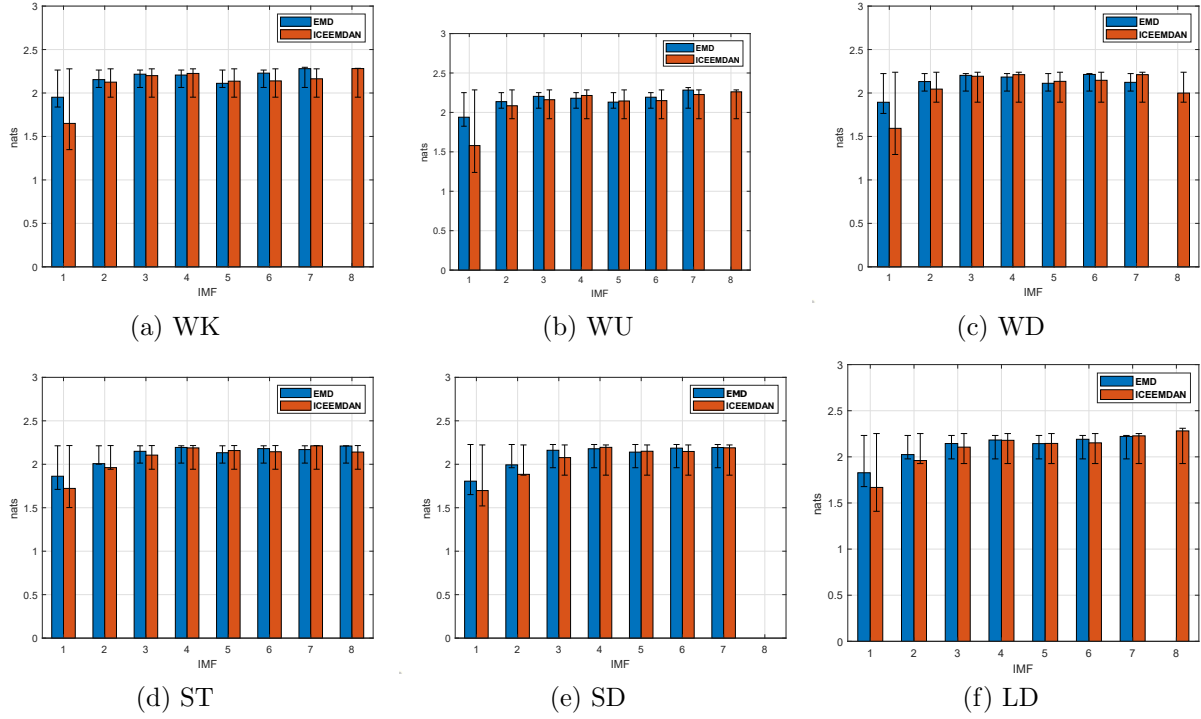


Figura 19: Barras de error de las entropías de Shannon de las IMF descompuestas por EMD (azul) e ICEEMDAN (naranja) del giroscopio en el eje x para cada actividad.

actividades humanas analizadas se encuentran muy por debajo de este rango, ya que tan solo se abarcaron seis actividades humanas. En general, las IMF de la ICEEMDAN contienen más información frecuencial que las de la EMD, por ejemplo, el rango de las IMF 4 llegan hasta 2 Hz mientras que en EMD ya está por debajo de ese valor. Se observa que existe un mayor rango de frecuencia usando el método de EMD debido a que las bajas frecuencias están más cercanas a cero.

Tabla 12: Rango de frecuencia de las IMF del acelerómetro obtenidas con la técnica EMD.

Parám	WK	WU	WD	ST	SD	LD
$\vec{f}_{LA_2IMF_1}, \vec{f}_{HA_2IMF_1}$	0.39±0.49, 7.6±1.45	0.28±0.42, 7.96±1.74	0.37±0.51, 8.67±1.6	0.48±0.60, 9.7±1.25	0.49±0.62, 9.60±1.30	0.48±0.61, 9.71±1.26
$\vec{f}_{LA_2IMF_2}, \vec{f}_{HA_2IMF_2}$	0.15±0.26, 4.0±0.61	0.14±0.23, 3.84±0.72	0.12±0.20, 3.75±0.65	0.35±0.38, 4.71±0.62	0.34±0.38, 4.73±0.61	0.27±0.32, 4.58±0.62
$\vec{f}_{LA_2IMF_3}, \vec{f}_{HA_2IMF_3}$	0.08±0.1, 2.35±0.31	0.06±0.09, 2.21±0.30	0.06±0.09, 2.29±0.30	0.14±0.17, 2.83±0.39	0.13±0.16, 2.82±0.38	0.12±0.16, 2.77±0.39
$\vec{f}_{LA_2IMF_4}, \vec{f}_{HA_2IMF_4}$	0.04±0.08, 1.88±0.39	0.03±0.07, 1.81±0.44	0.04±0.08, 1.84±0.42	0.05±0.08, 1.97±0.37	0.05±0.08, 1.99±0.36	0.04±0.07, 1.85±0.45
$\vec{f}_{LA_2IMF_5}, \vec{f}_{HA_2IMF_5}$	0.01±0.04, 1.51±0.61	0.01±0.04, 1.46±0.63	0.01±0.04, 1.39±0.57	0.01±0.05, 1.58±0.53	0.01±0.05, 1.58±0.52	0.01±0.04, 1.52±0.59
$\vec{f}_{LA_2IMF_6}, \vec{f}_{HA_2IMF_6}$	0.01±0.03, 1.35±0.68	0.00±0.03, 1.30±0.64	0.00±0.02, 1.28±0.62	0.00±0.02, 1.38±0.65	0.00±0.03, 1.38±0.65	0.00±0.02, 1.36±0.65
$\vec{f}_{LA_2IMF_7}, \vec{f}_{HA_2IMF_7}$	0±0, 1.20±0.40	0±0, 1.31±0.46	0±0, 1.05±0.35	0.01±0.04, 1.55±0.84	0.01±0.03, 1.36±0.70	0±0, 1.12±0.51
$\vec{f}_{LA_2IMF_1}, \vec{f}_{HA_2IMF_1}$	0.40±0.50, 7.20±1.39	0.34±0.42, 7.74±1.52	0.35±0.46, 7.56±1.59	0.41±0.55, 9.72±1.37	0.41±0.56, 9.54±1.53	0.40±0.54, 9.73±1.35
$\vec{f}_{LA_2IMF_2}, \vec{f}_{HA_2IMF_2}$	0.24±0.30, 4.04±0.52	0.16±0.23, 3.89±0.59	0.16±0.23, 4.03±0.53	0.26±0.32, 4.39±0.68	0.26±0.32, 4.39±0.68	0.29±0.35, 4.62±0.63
$\vec{f}_{LA_2IMF_3}, \vec{f}_{HA_2IMF_3}$	0.07±0.09, 2.42±0.34	0.06±0.09, 2.24±0.31	0.08±0.11, 2.50±0.35	0.11±0.15, 2.70±0.41	0.10±0.13, 2.64±0.41	0.12±0.16, 2.75±0.39
$\vec{f}_{LA_2IMF_4}, \vec{f}_{HA_2IMF_4}$	0.04±0.08, 1.88±0.36	0.03±0.07, 1.81±0.40	0.04±0.08, 1.90±0.36	0.04±0.08, 1.89±0.40	0.04±0.08, 1.86±0.40	0.04±0.08, 1.92±0.41
$\vec{f}_{LA_2IMF_5}, \vec{f}_{HA_2IMF_5}$	0.01±0.05, 1.55±0.59	0.01±0.03, 1.46±0.62	0.01±0.05, 1.57±0.57	0.01±0.04, 1.49±0.54	0.01±0.04, 1.48±0.57	0.01±0.04, 1.54±0.54
$\vec{f}_{LA_2IMF_6}, \vec{f}_{HA_2IMF_6}$	0.00±0.02, 1.39±0.65	0.00±0.02, 1.38±0.68	0.00±0.02, 1.43±0.68	0.00±0.02, 1.35±0.65	0.00±0.02, 1.37±0.69	0.01±0.03, 1.40±0.67
$\vec{f}_{LA_2IMF_7}, \vec{f}_{HA_2IMF_7}$	0±0, 1.45±0.85	0±0, 1.04±0.29	0.01±0.06, 1.71±0.90	0±0, 1.16±0.55	0.01±0.03, 1.32±0.65	0.00±0.20, 1.31±0.64
$\vec{f}_{LA_2IMF_1}, \vec{f}_{HA_2IMF_1}$	0.39±0.51, 8.36±1.48	0.35±0.48, 9.00±1.52	0.37±0.50, 8.65±1.45	0.41±0.57, 10.09±0.90	0.40±0.54, 10.01±1.00	0.41±0.56, 9.99±1.04
$\vec{f}_{LA_2IMF_2}, \vec{f}_{HA_2IMF_2}$	0.24±0.30, 4.43±0.62	0.18±0.24, 4.20±0.62	0.26±0.31, 4.39±0.55	0.29±0.34, 4.69±0.59	0.30±0.36, 4.73±0.56	0.29±0.35, 4.78±0.57
$\vec{f}_{LA_2IMF_3}, \vec{f}_{HA_2IMF_3}$	0.08±0.11, 2.54±0.39	0.07±0.10, 2.39±0.40	0.10±0.14, 2.66±0.39	0.12±0.15, 2.75±0.39	0.11±0.15, 2.74±0.39	0.12±0.16, 2.81±0.39
$\vec{f}_{LA_2IMF_4}, \vec{f}_{HA_2IMF_4}$	0.04±0.08, 1.84±0.33	0.03±0.07, 1.79±0.37	0.10±0.14, 2.66±0.34	0.01±0.04, 1.49±0.41	0.04±0.08, 1.90±0.40	0.05±0.08, 1.95±0.40
$\vec{f}_{LA_2IMF_5}, \vec{f}_{HA_2IMF_5}$	0.01±0.05, 1.57±0.58	0.01±0.04, 1.52±0.62	0.01±0.05, 1.59±0.55	0.01±0.04, 1.49±0.53	0.01±0.04, 1.50±0.55	0.01±0.04, 1.55±0.56
$\vec{f}_{LA_2IMF_6}, \vec{f}_{HA_2IMF_6}$	0.00±0.02, 1.42±0.65	0.00±0.02, 1.44±0.67	0.01±0.03, 1.46±0.72	0.00±0.02, 1.33±0.64	0.01±0.03, 1.28±0.6	0.00±0.02, 1.38±0.66
$\vec{f}_{LA_2IMF_7}, \vec{f}_{HA_2IMF_7}$	0.01±0.03, 1.4±0.89	0±0, 1.54±0.67	0.01±0.03, 1.42±0.84	0.01±0.05, 1.35±0.8	0±0, 1.35±0.70	0±0, 1.24±0.53

Tabla 13: Rango de frecuencia de las IMF del giroscopio obtenidas con la técnica EMD.

Parám	WK	WU	WD	ST	SD	LD
$\vec{f}_{LG_zIMF_1}, \vec{f}_{HG_zIMF_1}$	0.36±0.52, 7.34±1.67	0.34±0.41, 7.27±1.74	0.34±0.45, 7.88±1.67	0.46±0.60, 9.58±1.47	0.36±0.54, 8.89±2.19	0.44±0.61, 9.58±1.62
$\vec{f}_{LG_zIMF_2}, \vec{f}_{HG_zIMF_2}$	0.22±0.29, 3.95±0.56	0.19±0.30, 3.81±0.60	0.16±0.22, 3.86±0.53	0.26±0.32, 4.45±0.78	0.19±0.30, 3.91±0.80	0.23±0.30, 4.41±0.76
$\vec{f}_{LG_zIMF_3}, \vec{f}_{HG_zIMF_3}$	0.08±0.12, 2.37±0.44	0.06±0.10, 2.28±0.39	0.06±0.11, 2.39±0.38	0.10±0.14, 2.62±0.42	0.07±0.10, 2.34±0.43	0.11±0.15, 2.63±0.43
$\vec{f}_{LG_zIMF_4}, \vec{f}_{HG_zIMF_4}$	0.03±0.07, 1.72±0.42	0.02±0.06, 1.66±0.46	0.02±0.07, 1.72±0.43	0.04±0.08, 1.87±0.42	0.03±0.07, 1.77±0.43	0.04±0.08, 1.86±0.44
$\vec{f}_{LG_zIMF_5}, \vec{f}_{HG_zIMF_5}$	0.01±0.03, 1.47±0.63	0.00±0.02, 1.26±0.56	0.01±0.03, 1.33±0.59	0.01±0.04, 1.46±0.55	0.01±0.04, 1.45±0.59	0.01±0.04, 1.47±0.58
$\vec{f}_{LG_zIMF_6}, \vec{f}_{HG_zIMF_6}$	0.01±0.03, 1.47±0.76	0±0, 1.14±0.55	0.00±0.03, 1.26±0.68	0.00±0.02, 1.3±0.65	0.00±0.02, 1.30±0.62	0.00±0.02, 1.33±0.64
$\vec{f}_{LG_zIMF_7}, \vec{f}_{HG_zIMF_7}$	0±0, 1.31±0.81	0±0, 0.87±0.13	0±0, 0.93±0.08	0±0, 1.13±0.34	0±0, 1.48±0.67	0±0, 1.26±0.60
$\vec{f}_{LG_yIMF_1}, \vec{f}_{HG_yIMF_1}$	0.60±0.63, 7.92±1.42	0.41±0.50, 7.97±1.43	0.45±0.56, 8.36±1.38	0.48±0.60, 9.24±1.50	0.49±0.61, 9.29±1.59	0.65±0.77, 9.09±1.54
$\vec{f}_{LG_yIMF_2}, \vec{f}_{HG_yIMF_2}$	0.17±0.30, 3.92±0.60	0.21±0.28, 4.04±0.61	0.19±0.26, 4.19±0.62	0.26±0.33, 4.45±0.73	0.24±0.31, 4.36±0.73	0.29±0.36, 4.60±0.73
$\vec{f}_{LG_yIMF_3}, \vec{f}_{HG_yIMF_3}$	0.09±0.12, 2.44±0.37	0.06±0.10, 2.32±0.43	0.09±0.13, 2.55±0.37	0.09±0.13, 2.59±0.44	0.09±0.13, 2.58±0.42	0.11±0.15, 2.67±0.43
$\vec{f}_{LG_yIMF_4}, \vec{f}_{HG_yIMF_4}$	0.03±0.07, 1.78±0.37	0.02±0.06, 1.72±0.43	0.04±0.08, 1.86±0.34	0.04±0.08, 1.86±0.42	0.04±0.07, 1.87±0.39	0.04±0.08, 1.88±0.42
$\vec{f}_{LG_yIMF_5}, \vec{f}_{HG_yIMF_5}$	0.01±0.04, 1.5±0.60	0.01±0.03, 1.37±0.59	0.01±0.04, 1.48±0.55	0.01±0.04, 1.46±0.57	0.01±0.04, 1.49±0.56	0.01±0.05, 1.52±0.58
$\vec{f}_{LG_yIMF_6}, \vec{f}_{HG_yIMF_6}$	0.00±0.02, 1.41±0.71	0.01±0.03, 1.27±0.63	0.00±0.02, 1.30±0.65	0.00±0.02, 1.33±0.64	0.00±0.02, 1.29±0.61	0.00±0.02, 1.33±0.64
$\vec{f}_{LG_yIMF_7}, \vec{f}_{HG_yIMF_7}$	0±0, 1.50±0.73	0±0, 0.92±0.18	0±0, 0.97±0.11	0±0, 1.30±0.66	0±0, 1.25±0.62	0.01±0.04, 1.22±0.67
$\vec{f}_{LG_zIMF_1}, \vec{f}_{HG_zIMF_1}$	0.47±0.59, 7.38±1.46	0.38±0.47, 7.26±1.46	0.43±0.54, 7.32±1.48	0.43±0.56, 9.68±1.49	0.41±0.57, 9.28±1.92	0.43±0.58, 9.79±1.46
$\vec{f}_{LG_zIMF_2}, \vec{f}_{HG_zIMF_2}$	0.13±0.20, 3.52±0.61	0.16±0.22, 3.71±0.64	0.16±0.22, 3.75±0.58	0.26±0.33, 4.52±0.76	0.21±0.28, 4.28±0.78	0.23±0.31, 4.56±0.72
$\vec{f}_{LG_zIMF_3}, \vec{f}_{HG_zIMF_3}$	0.07±0.10, 2.24±0.40	0.06±0.09, 2.16±0.39	0.08±0.10, 2.38±0.36	0.09±0.13, 2.59±0.47	0.09±0.12, 2.53±0.43	0.11±0.15, 2.69±0.48
$\vec{f}_{LG_zIMF_4}, \vec{f}_{HG_zIMF_4}$	0.03±0.07, 1.73±0.45	0.02±0.06, 1.69±0.47	0.04±0.08, 1.83±0.38	0.03±0.07, 1.81±0.43	0.04±0.08, 1.85±0.44	0.04±0.08, 1.85±0.44
$\vec{f}_{LG_zIMF_5}, \vec{f}_{HG_zIMF_5}$	0.01±0.04, 1.49±0.62	0.01±0.03, 1.39±0.62	0.01±0.03, 1.45±0.56	0.01±0.04, 1.42±0.56	0.01±0.04, 1.46±0.55	0.01±0.05, 1.48±0.55
$\vec{f}_{LG_zIMF_6}, \vec{f}_{HG_zIMF_6}$	0.00±0.02, 1.43±0.70	0.00±0.03, 1.34±0.69	0.00±0.02, 1.32±0.67	0.00±0.02, 1.31±0.63	0.00±0.02, 1.27±0.61	0.00±0.02, 1.27±0.62
$\vec{f}_{LG_zIMF_7}, \vec{f}_{HG_zIMF_7}$	0±0, 1.28±0.33	0±0, 1.49±0.40	0.01±0.05, 1.43±0.73	0±0, 1.10±0.50	0±0, 1.34±0.62	0.00±0.03, 1.27±0.57

Tabla 14: Rango de frecuencia de las IMF del acelerómetro obtenidas con la técnica ICEEMDAN.

Parám	WK	WU	WD	ST	SD	LD
$\vec{f}_{L_{A_2IMF1}}, \vec{f}_{H_{A_2IMF1}}$	1.27±0.73, 9.7±1.35	1.42±0.8, 10.61±0.85	1.62±0.83, 10.6±0.88	1.35±0.88, 10.21±1.01	1.23±0.86, 10.1±1.05	1.66±0.99, 10.5±0.88
$\vec{f}_{L_{A_2IMF2}}, \vec{f}_{H_{A_2IMF2}}$	0.42±0.47, 5.008±0.32	0.36±0.37, 5.09±0.42	0.3±0.35, 5.16±0.39	0.54±0.48, 5.13±0.52	0.5±0.47, 5.08±0.48	0.56±0.52, 5.36±0.62
$\vec{f}_{L_{A_2IMF3}}, \vec{f}_{H_{A_2IMF3}}$	0.06±0.09, 2.6±0.3	0.04±0.08, 2.5±0.29	0.03±0.08, 2.66±0.3	0.18±0.2, 3.04±0.28	0.16±0.19, 2.98±0.29	0.3±0.26, 3.11±0.3
$\vec{f}_{L_{A_2IMF4}}, \vec{f}_{H_{A_2IMF4}}$	0.06±0.09, 2.01±0.29	0.05±0.08, 1.95±0.31	0.06±0.09, 2.04±0.27	0.06±0.09, 2.09±0.32	0.06±0.09, 2.09±0.3	0.06±0.09, 2.12±0.38
$\vec{f}_{L_{A_2IMF5}}, \vec{f}_{H_{A_2IMF5}}$	0.02±0.06, 1.72±0.51	0.01±0.05, 1.58±0.56	0.02±0.06, 1.64±0.49	0.02±0.06, 1.59±0.47	0.01±0.05, 1.63±0.46	0.01±0.05, 1.56±0.49
$\vec{f}_{L_{A_2IMF6}}, \vec{f}_{H_{A_2IMF6}}$	0.004±0.02, 1.35±0.64	0.002±0.02, 1.26±0.63	0.002±0.02, 1.07±0.41	0.003±0.02, 1.42±0.64	0.006±0.03, 1.43±0.67	0.005±0.03, 1.45±0.7
$\vec{f}_{L_{A_2IMF7}}, \vec{f}_{H_{A_2IMF7}}$	0.005±0.03, 1.35±0.78	0±0, 1.13±0.51	0±0, 0.99±0.39	0.005±0.03, 1.25±0.63	0.005±0.03, 1.38±0.72	0±0, 1.13±0.53
$\vec{f}_{L_{A_2IMF1}}, \vec{f}_{H_{A_2IMF1}}$	1.16±0.69, 9.05±1.65	1.32±0.76, 10.32±1.14	1.08±0.69, 9.64±1.42	1.43±0.96, 10.5±0.96	1.5±0.92, 10.54±0.99	1.38±0.93, 10.46±0.9
$\vec{f}_{L_{A_2IMF2}}, \vec{f}_{H_{A_2IMF2}}$	0.48±0.42, 4.83±0.34	0.47±0.42, 4.97±0.38	0.39±0.39, 4.81±0.35	0.52±0.49, 5.17±0.55	0.5±0.46, 5.11±0.55	0.51±0.48, 5.17±0.58
$\vec{f}_{L_{A_2IMF3}}, \vec{f}_{H_{A_2IMF3}}$	0.06±0.1, 2.77±0.28	0.06±0.1, 2.7±0.27	0.09±0.15, 2.82±0.24	0.18±0.2, 3.004±0.32	0.16±0.18, 2.97±0.3	0.19±0.21, 3±0.31
$\vec{f}_{L_{A_2IMF4}}, \vec{f}_{H_{A_2IMF4}}$	0.06±0.09, 2.02±0.28	0.05±0.08, 1.95±0.29	0.06±0.09, 2.02±0.28	0.06±0.09, 2.1±0.32	0.06±0.09, 2.06±0.3	0.06±0.09, 2.11±0.34
$\vec{f}_{L_{A_2IMF5}}, \vec{f}_{H_{A_2IMF5}}$	0.03±0.07, 1.73±0.46	0.02±0.06, 1.7±0.51	0.02±0.06, 1.73±0.44	0.01±0.05, 1.55±0.47	0.01±0.05, 1.55±0.47	0.01±0.05, 1.55±0.45
$\vec{f}_{L_{A_2IMF6}}, \vec{f}_{H_{A_2IMF6}}$	0.005±0.03, 1.39±0.65	0.005±0.03, 1.38±0.64	0.007±0.03, 1.48±0.66	0.004±0.02, 1.37±0.65	0.002±0.02, 1.35±0.63	0.003±0.02, 1.41±0.64
$\vec{f}_{L_{A_2IMF7}}, \vec{f}_{H_{A_2IMF7}}$	0.01±0.05, 1.34±0.74	0±0, 1.21±0.66	0.002±0.02, 1.29±0.66	0±0, 1.13±0.51	0.003±0.02, 1.27±0.67	0±0, 1.21±0.56
$\vec{f}_{L_{A_2IMF1}}, \vec{f}_{H_{A_2IMF1}}$	1.36±0.74, 9.55±1.35	1.73±0.89, 10.66±0.99	1.17±0.78, 9.87±1.09	1.58±0.93, 10.61±0.72	1.52±0.92, 10.57±0.67	1.46±0.9, 10.52±0.7
$\vec{f}_{L_{A_2IMF2}}, \vec{f}_{H_{A_2IMF2}}$	0.58±0.52, 5.14±0.33	0.55±0.48, 5.4±0.43	0.47±0.42, 5.05±0.34	0.55±0.47, 5.31±0.49	0.6±0.49, 5.35±0.46	0.56±0.48, 5.22±0.46
$\vec{f}_{L_{A_2IMF3}}, \vec{f}_{H_{A_2IMF3}}$	0.08±0.13, 2.9±0.3	0.09±0.14, 2.91±0.29	0.13±0.19, 2.98±0.3	0.17±0.2, 3.03±0.29	0.16±0.19, 3.04±0.28	0.18±0.19, 3.02±0.29
$\vec{f}_{L_{A_2IMF4}}, \vec{f}_{H_{A_2IMF4}}$	0.05±0.08, 2.01±0.31	0.05±0.08, 1.89±0.27	0.05±0.08, 2.04±0.32	0.06±0.09, 2.09±0.31	0.05±0.09, 2.07±0.31	0.06±0.09, 2.09±0.3
$\vec{f}_{L_{A_2IMF5}}, \vec{f}_{H_{A_2IMF5}}$	0.02±0.06, 1.72±0.45	0.02±0.06, 1.7±0.49	0.02±0.06, 1.71±0.46	0.01±0.05, 1.52±0.45	0.01±0.05, 1.54±0.46	0.01±0.05, 1.55±0.46
$\vec{f}_{L_{A_2IMF6}}, \vec{f}_{H_{A_2IMF6}}$	0.006±0.03, 1.43±0.67	0.004±0.03, 1.44±0.67	0.006±0.03, 1.42±0.64	0.004±0.02, 1.43±0.66	0.003±0.02, 1.35±0.63	0.004±0.027, 1.43±0.66
$\vec{f}_{L_{A_2IMF7}}, \vec{f}_{H_{A_2IMF7}}$	0.003±0.02, 1.32±0.62	0±0, 1.19±0.55	0.01±0.05, 1.35±0.68	0±0, 1.2±0.55	0.006±0.03, 1.31±0.74	0±0, 1.22±0.58

Tabla 15: Rango de frecuencia de las IMF del giroscopio obtenidas con la técnica ICEEMDAN.

Parám	WK	WU	WD	ST	SD	LD
$\tilde{f}_{LG_zIMF_1}, \tilde{f}_{HG_zIMF_1}$	1.24±0.71, 9.54±1.65	1.35±0.84, 10.35±1.37	1.39±0.85, 10.38±1.15	1.52±0.94, 10.46±1.02	1.57±1.08, 10.86±1.08	1.63±1.05, 10.62±0.95
$\tilde{f}_{LG_zIMF_2}, \tilde{f}_{HG_zIMF_2}$	0.42±0.4, 4.88±0.38	0.57±0.44, 5.05±0.43	0.47±0.43, 5.04±0.43	0.5±0.49, 5.27±0.58	0.4±0.41, 5.02±0.8	0.53±0.51, 5.24±0.74
$\tilde{f}_{LG_zIMF_3}, \tilde{f}_{HG_zIMF_3}$	0.13±0.16, 2.97±0.29	0.13±0.15, 2.93±0.29	0.16±0.18, 2.87±0.23	0.16±0.18, 2.94±0.33	0.12±0.15, 2.83±0.33	0.16±0.19, 2.92±0.34
$\tilde{f}_{LG_zIMF_4}, \tilde{f}_{HG_zIMF_4}$	0.04±0.08, 1.86±0.3	0.04±0.08, 1.96±0.31	0.04±0.08, 1.97±0.31	0.06±0.09, 2.07±0.35	0.05±0.08, 1.99±0.29	0.06±0.09, 2.07±0.36
$\tilde{f}_{LG_zIMF_5}, \tilde{f}_{HG_zIMF_5}$	0.01±0.05, 1.65±0.59	0.01±0.04, 1.46±0.5	0.01±0.05, 1.47±0.51	0.01±0.05, 1.56±0.49	0.02±0.06, 1.59±0.49	0.01±0.053, 1.52±0.48
$\tilde{f}_{LG_zIMF_6}, \tilde{f}_{HG_zIMF_6}$	0.003±0.02, 1.4±0.66	0.002±0.02, 1.21±0.58	0.003±0.02, 1.29±0.65	0.005±0.03, 1.34±0.64	0.005±0.03, 1.38±0.68	0.005±0.03, 1.34±0.64
$\tilde{f}_{LG_zIMF_7}, \tilde{f}_{HG_zIMF_7}$	0.01±0.04, 1.47±0.81	0±0, 1.12±0.49	0.005±0.03, 1.3±0.67	0±0, 1.16±0.58	0±0, 1.01±0.35	0.002±0.02, 1.29±0.69
$\tilde{f}_{LG_yIMF_1}, \tilde{f}_{HG_yIMF_1}$	1.5±0.74, 8.94±1.6	1.47±0.87, 9.85±1.45	1.29±0.79, 9.43±1.33	1.63±0.93, 10.3±1.2	1.55±0.92, 10.28±1.14	1.78±0.98, 10.02±1.25
$\tilde{f}_{LG_yIMF_2}, \tilde{f}_{HG_yIMF_2}$	0.39±0.43, 5.03±0.47	0.61±0.5, 5.31±0.42	0.38±0.41, 5.07±0.43	0.62±0.54, 5.36±0.52	0.48±0.46, 5.22±0.61	0.64±0.55, 5.45±0.64
$\tilde{f}_{LG_yIMF_3}, \tilde{f}_{HG_yIMF_3}$	0.13±0.16, 2.74±0.24	0.12±0.15, 2.95±0.28	0.14±0.17, 2.83±0.24	0.15±0.17, 2.99±0.33	0.14±0.17, 2.93±0.3	0.17±0.2, 3.04±0.32
$\tilde{f}_{LG_yIMF_4}, \tilde{f}_{HG_yIMF_4}$	0.03±0.07, 1.85±0.32	0.03±0.07, 1.86±0.31	0.05±0.08, 2.003±0.3	0.06±0.09, 2.06±0.31	0.05±0.08, 2.03±0.29	0.05±0.08, 2.06±0.32
$\tilde{f}_{LG_yIMF_5}, \tilde{f}_{HG_yIMF_5}$	0.01±0.05, 1.62±0.55	0.01±0.05, 1.58±0.55	0.02±0.06, 1.64±0.5	0.01±0.05, 1.57±0.49	0.02±0.06, 1.59±0.46	0.01±0.05, 1.58±0.47
$\tilde{f}_{LG_yIMF_6}, \tilde{f}_{HG_yIMF_6}$	0.007±0.03, 1.49±0.71	0.004±0.03, 1.29±0.61	0.004±0.03, 1.36±0.68	0.003±0.02, 1.37±0.64	0.005±0.03, 1.35±0.63	0.005±0.03, 1.37±0.64
$\tilde{f}_{LG_yIMF_7}, \tilde{f}_{HG_yIMF_7}$	0±0, 1.13±0.44	0.007±0.03, 1.08±0.51	0±0, 1.08±0.49	0±0, 1.18±0.51	0.001±0.01, 1.24±0.62	0.003±0.02, 1.27±0.65
$\tilde{f}_{LG_zIMF_1}, \tilde{f}_{HG_zIMF_1}$	1.48±0.79, 9.46±1.66	1.4±0.87, 10.17±1.48	1.28±0.75, 9.17±1.56	1.62±1.006, 10.61±1.15	1.57±0.97, 10.56±1.13	1.56±0.97, 10.62±0.95
$\tilde{f}_{LG_zIMF_2}, \tilde{f}_{HG_zIMF_2}$	0.31±0.41, 4.9±0.49	0.6±0.46, 5.21±0.47	0.29±0.35, 4.86±0.4	0.63±0.53, 5.43±0.59	0.46±0.45, 5.16±0.62	0.5±0.5, 5.26±0.61
$\tilde{f}_{LG_zIMF_3}, \tilde{f}_{HG_zIMF_3}$	0.13±0.16, 2.64±0.22	0.12±0.15, 2.9±0.31	0.11±0.16, 2.78±0.23	0.15±0.18, 3.03±0.37	0.14±0.17, 2.9±0.3	0.2±0.22, 3.01±0.33
$\tilde{f}_{LG_zIMF_4}, \tilde{f}_{HG_zIMF_4}$	0.04±0.08, 1.87±0.32	0.04±0.07, 1.87±0.3	0.05±0.08, 2.02±0.31	0.05±0.08, 2.03±0.35	0.05±0.08, 2.03±0.31	0.05±0.09, 2.03±0.4
$\tilde{f}_{LG_zIMF_5}, \tilde{f}_{HG_zIMF_5}$	0.02±0.06, 1.66±0.54	0.01±0.05, 1.56±0.54	0.02±0.06, 1.66±0.49	0.01±0.05, 1.51±0.49	0.01±0.05, 1.56±0.49	0.01±0.05, 1.52±0.48
$\tilde{f}_{LG_zIMF_6}, \tilde{f}_{HG_zIMF_6}$	0.005±0.03, 1.42±0.67	0.004±0.03, 1.35±0.67	0.005±0.03, 1.35±0.64	0.005±0.03, 1.34±0.64	0.004±0.02, 1.33±0.62	0.004±0.02, 1.31±0.63
$\tilde{f}_{LG_zIMF_7}, \tilde{f}_{HG_zIMF_7}$	0.01±0.04, 1.28±0.72	0.006±0.03, 1.33±0.71	0±0, 1.31±0.7	0.002±0.02, 1.21±0.62	0±0, 1.15±0.53	0.004±0.02, 1.25±0.66

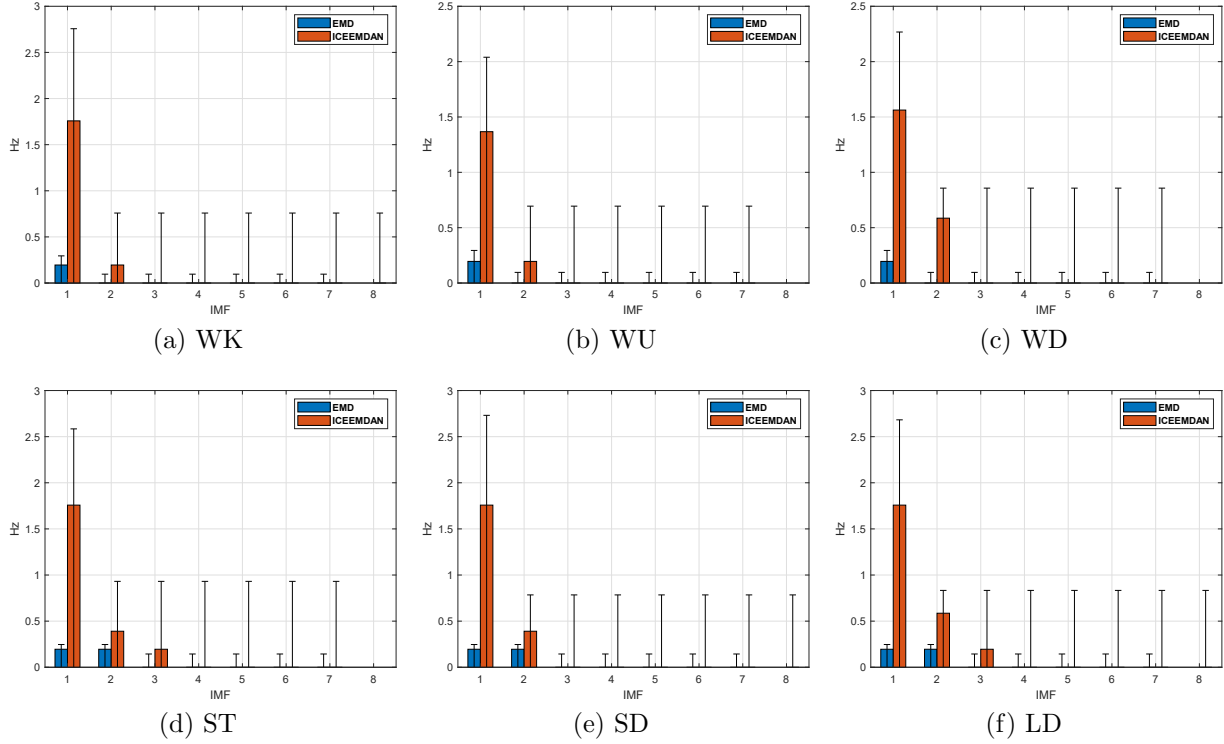


Figura 20: Barras de error de f_L del giroscopio en el eje z .

5.2. Clasificación de las actividades humanas

A continuación se muestran los resultados de la clasificación de las seis actividades humanas, tal como se describió en la sección 4.3.

5.2.1. Representación en tiempo-frecuencia

La figura 22 muestra un ejemplo de descomposición del acelerómetro en el eje x en seis IMF usando la técnica ICEEDMAN. Se observa que las oscilaciones más rápidas aparecen en las primeras IMF y, a medida que aumentan las IMF, las oscilaciones se hacen más lentas. Además, la última IMF, es decir, la IMF_6 no posee más de dos extremos, obedeciendo a la teoría, según lo indicado en la sección 2.2.2.

Para el modelo de clasificación se utilizaron las nueve señales disponibles en la base de datos. No obstante, la cantidad de IMF obtenidas para cada señal es variable, ya que como se mencionó, el número de IMF depende únicamente de la misma señal, consiguiendo señales con un rango de IMF desde tres hasta ocho. La tabla 16 muestra el número de registros

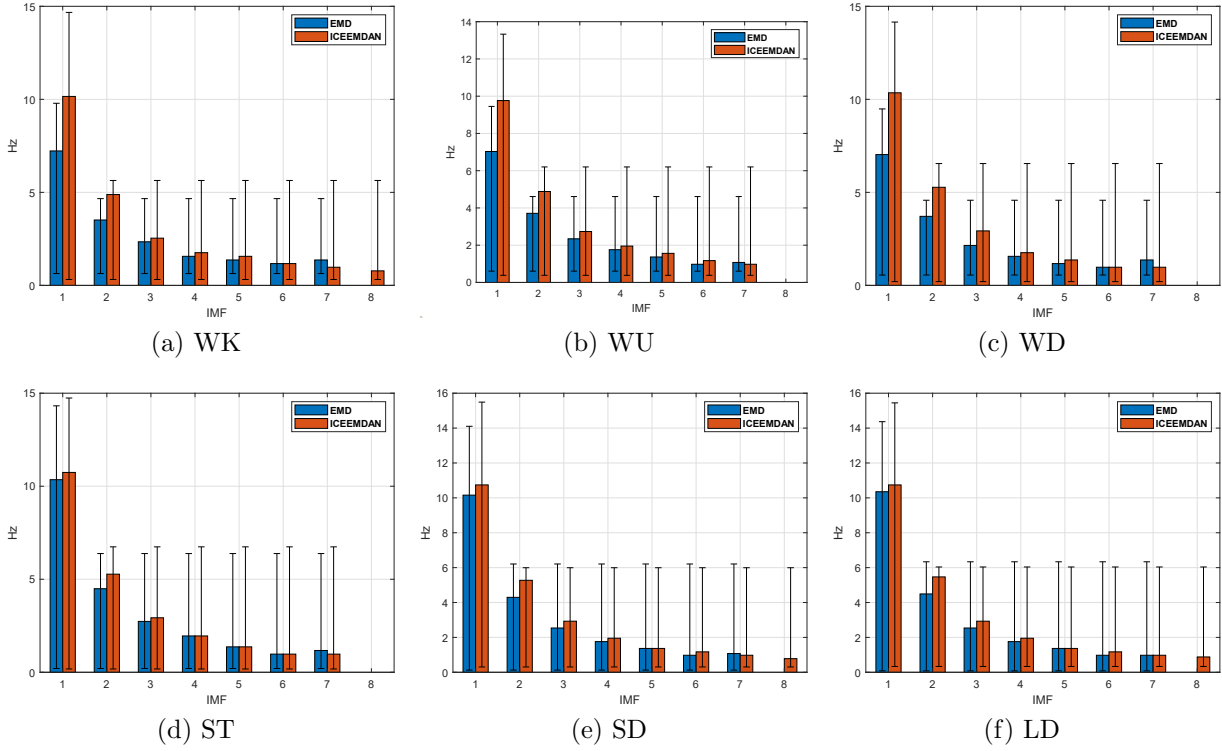


Figura 21: Barras de error de f_H del giroscopio en el eje z .

de la base de datos que pueden descomponerse hasta ese máximo número de IMF usando ICEEMDAN.

Debido a que a la entrada del modelo de clasificación propuesto se debe ingresar el mismo número de observaciones para cada señal, se optó por limitar al número máximo de IMF a cinco, teniendo en cuenta los resultados obtenidos en la sección anterior en cuanto a varianza y entropía de Shannon, ver tablas 8 y 10, dado que la información que poseían las últimas IMF no aportaban mucho a la actividad realizada. En ese sentido, se procedió a eliminar los registros que produjeran menos de cinco IMF. De la tabla 16 se observa que se eliminarán 340 registros del subconjunto de entrenamiento y 113 registros del subconjunto de prueba que no alcanzan a descomponerse en cinco IMF. Se realizó entonces el entrenamiento del clasificador en 7012 registros y se probó en 2834 registros.

La figura 23 muestra un ejemplo de la descomposición en cinco IMF de la señal mostrada en la figura 22. Esta señal ha sido estandarizada (puntajes Z), lo cual produce una señal escalada, con media igual a cero y desviación estándar unitaria. Esto beneficia el proceso de clasificación al evitar el sesgo producido por señales con diferentes rangos de amplitud.

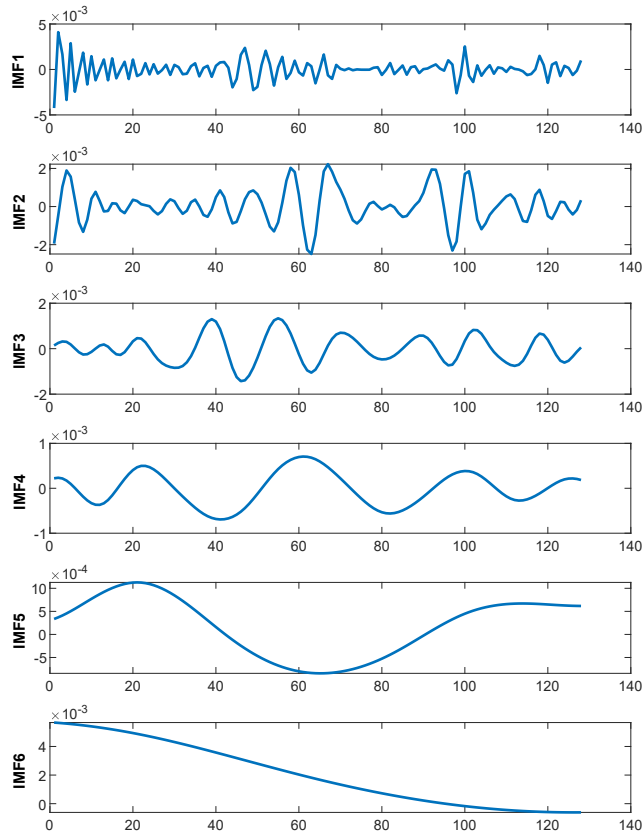


Figura 22: Descomposición en seis IMF, por medio de la técnica ICEEMDAN, de una señal del acelerómetro en el eje x .

Tabla 16: Cantidad de registros que pueden descomponerse hasta el máximo número de IMF indicado usando la técnica ICEEMDAN.

ICEEMDAN		
IMF máxima	Entrenamiento	Prueba
1	0	0
2	0	0
3	7	2
4	340	113
5	6617	2659
6	7352	2947
7	7352	2947
8	7352	2947

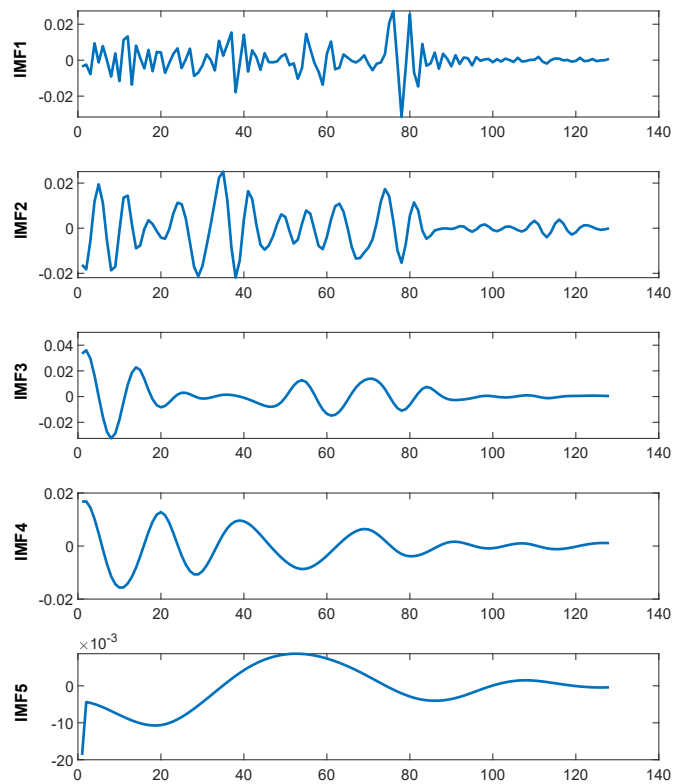


Figura 23: Ejemplo de representación de una señal del acelerómetro en el eje x en cinco IMF. Las IMF están estandarizadas.

5.2.2. Segmentación

Como se mencionó en la subsección 4.3.3, se dividió cada IMF, de ambos conjuntos, en nueve partes, cada una de 1.28 s, correspondiente a 64 muestras, con un solapamiento entre ventanas de observación de 12.5 % (desplazamiento de 8 muestras). Esto produjo un total de 63108 señales para el conjunto de entrenamiento y 25506 para el conjunto de prueba.

Un ejemplo de la segmentación de una señal se aprecia en la figura 24. Esta etapa nos permitió aumentar la cantidad de observaciones que se tienen, dando como resultado un mejor entrenamiento por parte de la red, lo que produce una mayor exactitud.

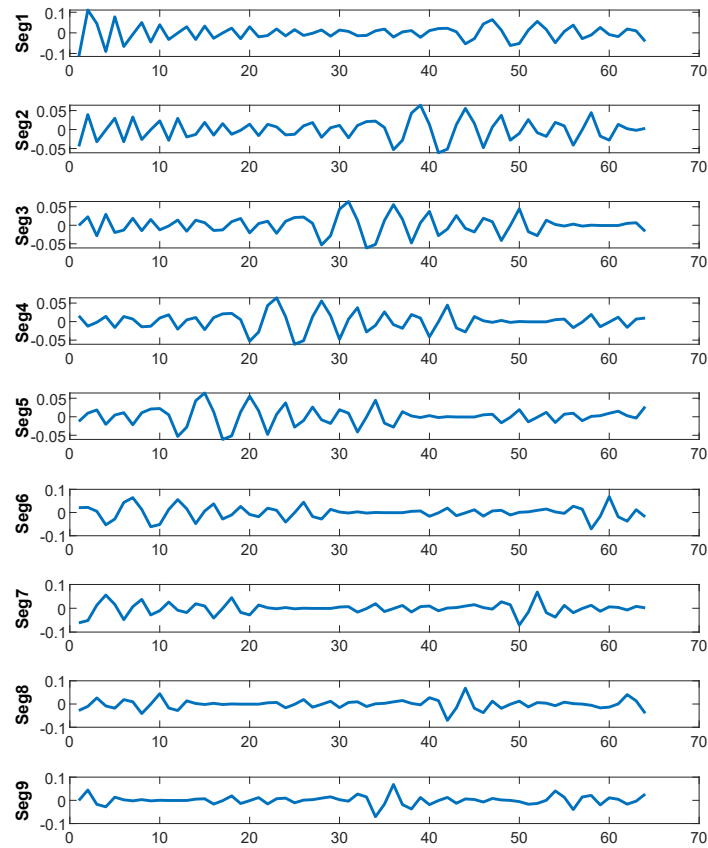


Figura 24: Segmentación en ventanas de 64 muestras de la IMF 1 de una señal del acelerómetro en el eje x .

5.2.3. Clasificación

Primer modelo de clasificación

El primer modelo de clasificación está compuesto por cuatro salidas, tres de las cuales fueron entrenadas para identificar las actividades dinámicas (WK, WU, WD), y la cuarta salida fue entrenada para identificar las actividades estáticas (ST, SD, LD) como una única clase denominada *statics*. La tabla 17 muestra la matriz de confusión de esta primera etapa de clasificación. En esta tabla también se aprecia la exactitud, y la precisión y la exhaustividad de cada clase.

Tabla 17: Matriz de confusión del primer modelo de clasificación.

Clase verdadera	Predicción				Rec (%)
	WK	WU	WD	Statics	
WK	469	1	18	0	96.1
WU	1	462	0	0	99.8
WD	1	9	406	0	97.6
Statics	0	1	1	1465	99.9
Pre (%)	99.6	97.7	95.5	100	98.87

Se puede observar que todas las métricas de desempeño del clasificador se encuentran por encima del 95 %. Se destaca asimismo una exactitud del 98.87 %, lo cual es un resultado bastante prometedor: tan solo se cometieron 32 errores en este primer modelo de clasificación. Además, el modelo tiene una alta capacidad de distinguir entre una actividad dinámica y una estática, ya que se cometieron solo dos errores: una actividad estática fue clasificada como WU y otra como WD. El mayor error fue cometido con la actividad caminando, ya que 18 de estos registros fueron clasificados como WU. Este resultado se debe, en gran medida, a la segmentación de las señales en ventanas de 1.28 s, ya que la actividad dinámica, ya sea caminando o bajando o subiendo escaleras, tiende a ser repetitiva y suele completarse un ciclo en menos de 1 segundo. Esta información a priori nos permitió entonces este elevado desempeño y corresponde a una de las principales contribuciones de este trabajo.

Es de recordar que, si bien la entrada del modelo de clasificación son segmentos cortos de las señales inerciales, el tipo de actividad se selecciona como la clase más probable entre los nueve segmentos clasificados. Además, debido a que son nueve segmentos por registro y cuatro clases (salidas) posibles, siempre va a existir una clase más probable.

Segundo modelo de clasificación

Una vez obtenidos los resultados de clasificación del primer modelo, los registros clasificados como *statics* fueron nuevamente clasificados por el segundo modelo con el fin de identificar el tipo de actividad estática que realizan. Por lo tanto, este segundo modelo de clasificación consta de tres salidas, una para cada clase estática.

Las señales de entrada de este modelo corresponden a las señales IMF estandarizadas, pero sin segmentar. Dado que las actividades estáticas no son repetitivas, la segmentación no aporta información a priori que pueda ayudar al clasificador a identificar la actividad. En la tabla 18 se aprecia la matriz de confusión de este modelo.

Tabla 18: Matriz de confusión del segundo modelo de clasificación.

Clase verdadera	Predicción			<i>Rec</i> (%)
	ST	SD	LD	
ST	351	107	0	76.6
SD	39	481	0	92.5
LD	0	23	464	95.3
<i>Pre</i> (%)	90	78.7	100	88.46

Los desempeños de este modelo tienden a ser variables, sin embargo, todas las métricas son superiores al 75 %. La exactitud es de 88.46 %. La actividad mejor identificada fue la de estar acostado, donde el modelo cometió 23 errores, clasificando esos registros como estar parado. Se observa además que al modelo se le dificulta distinguir las clases estar sentado y estar parado debido a que son movimientos muy similares.

Clasificación final

Finalmente, al combinar los resultados de ambos modelos, se obtiene la matriz de confusión que se muestra en la tabla 19. La exactitud total que se logró para identificar las seis clases fue de 92.91 %. Este resultado es comparable con los resultados obtenidos por otros autores usando otros enfoques.

Este importante desempeño se debe a tres estrategias que se aplicaron, las cuales corresponden a las principales contribuciones de este trabajo. En primer lugar, la realización de una clasificación en cascada de actividades estáticas y dinámicas, ya que las características de ambos tipos de actividades son diferentes (en las dinámicas se observa un patrón repetitivo que en las estáticas no existe). En segundo lugar, el incremento de la observabilidad del sistema a través de una representación en tiempo-frecuencia, en este caso las IMF derivadas

Tabla 19: Matriz de confusión del modelo de clasificación de las seis actividades humanas

Clase verdadera	Predicción						<i>Rec (%)</i>
	WK	WU	WD	ST	SD	LD	
WK	469	1	18	0	0	0	96.1
WU	1	462	0	0	0	0	99.8
WD	1	9	406	0	0	0	97.6
ST	0	1	0	351	107	0	76.5
SD	0	0	0	39	481	0	92.52
LD	0	0	1	0	23	464	95.1
<i>Pre (%)</i>	99.6	97.7	95.5	90	78.7	100	92.91

de la técnica ICEEMDAN, de las nueve señales inerciales. Y, en tercer lugar, la segmentación de las IMF en ventanas de observación más cortas que la señal original, para realzar el patrón característico de las actividades dinámicas y, a su vez, incrementar aún más la observabilidad del sistema.

CAPÍTULO 6

CONCLUSIONES Y RECOMENDACIONES

En este capítulo se presentan las conclusiones del trabajo realizado, haciendo un énfasis especial en el cumplimiento detallado de los objetivos y en el impacto de nuestro trabajo.

El objetivo general de este trabajo consistió en identificar seis actividades humanas: caminar, subir y bajar escaleras, estar sentado, estar parado y estar acostado. Esto se logró usando una base de datos comúnmente utilizada por la comunidad científica y en la que se encuentran disponibles de forma gratuita señales inerciales (acelerómetro y giroscopio) que representan estas actividades. Las señales, de duración 2.56 s, fueron recolectadas por un teléfono inteligente a un grupo de treinta voluntarios. Se desarrolló un modelo de clasificación basado en aprendizaje profundo, específicamente las redes neuronales recurrentes del tipo LSTM bidireccionales, recientemente concebidas y con una gran aceptación para este tipo de tareas debido a su gran versatilidad para procesar series de tiempo. El desempeño de la clasificación de las actividades humanas logrado en este trabajo es prometedor en cuanto al alto desempeño obtenido e innovador en cuanto al uso de la técnica de descomposición empírica en modos por conjunto completo con ruido adaptativo mejorada (ICEEMDAN) para representar las señales inerciales en tiempo-frecuencia como funciones de modo intrínseco (IMF). El análisis de las señales inerciales y la clasificación de las actividades humanas realizadas en este trabajo podrían ser útiles en el desarrollo de aplicaciones que utilicen el movimiento de las personas para interactuar con el ambiente, tal como en la realidad virtual y aumentada, y en el seguimiento de su condición de salud, como por ejemplo para combatir el sedentarismo, y de esta manera aportar en la reducción de obesidad y sus complicaciones asociadas.

El análisis de las señales inerciales por medio de descriptores estadísticos tuvo como fin ofrecer una caracterización exhaustiva de éstas, y permitió obtener información importante que se encuentra implícita en las señales y que es imposible de observar a simple vista. Este análisis se llevó a cabo no solamente en las señales inerciales crudas (sin procesar) sino en su representación tiempo-frecuencia en funciones de modo intrínseco obtenidas tanto por EMD como por ICEEMDAN. Entre los principales hallazgos se pueden destacar que: (i) la mayor

varianza y entropía de Shannon siempre se encontraba en las primeras IMF de las señales inerciales, específicamente entre las IMF 2 y 4, y (ii) las actividades dinámicas (WK, WU y WD) arrojaban los mayores valores de varianza y entropía de Shannon en comparación con los de las estáticas (ST, SD y LD), diferenciando así estos tipos de actividades. En ese sentido, la obtención de características relevantes de las señales inerciales que representan las actividades humanas es también una contribución importante de este trabajo, ya que, en un enfoque de aprendizaje automático supervisado, la selección de características es una tarea tediosa y poco realizada.

La utilización de una versión mejorada de la descomposición empírica en modos, específicamente la ICEEMDAN, permitió realizar una representación tiempo-frecuencia con IMF, en una base completa y casi ortogonal, con frecuencias variables, con menor error de reconstrucción y sin modos que se mezclan entre sí, en la que, a diferencia de otros tipos de representación tiempo-frecuencia como la transformada de ondículas, son las mismas señales las que manejan el proceso de descomposición y el nivel a descomponerse. Esta descomposición no solo nos permitió incrementar la observabilidad del sistema, al pasar de, por ejemplo, una señal en el dominio del tiempo a ocho señales en el mismo dominio del tiempo, pero con diferentes rangos de frecuencia, sino también conocer su información frecuencial relevante en varios intervalos de frecuencia. Además, esta disgregación fue particularmente útil en el proceso de clasificación de las actividades humanas ya que se contó con una mayor cantidad de información a explotar.

Finalmente, el enfoque de clasificación propuesto, compuesto de dos redes LSTM bidireccional en cascada y con la misma arquitectura, que procesa directamente la información contenida en las IMF, sin necesidad de extraer características de éstas, fue capaz de diferenciar las seis actividades humanas con muy pocos errores: 201 registros incorrectamente identificados de un universo de 2834 registros, es decir, apenas un 7 % de error. Las actividades dinámicas, como caminar y subir y bajar escaleras, fueron fácilmente identificadas dado que contienen un patrón típico que fue aprovechado por el clasificador gracias a un proceso de segmentación de las IMF: 30 registros con actividades dinámicas fueron incorrectamente identificados de un universo de 1367 registros, es decir, apenas un 2.20 % de error. A pesar de que las actividades estáticas, como estar sentado, estar parado o estar acostado, no contienen un patrón característico relevante, se pudo identificar correctamente la gran mayoría de estas actividades: 169 registros con actividades estáticas fueron incorrectamente identificados de un universo de 1467 registros, es decir, un 11.52 % de error.

Las contribuciones principales de este trabajo corresponden a:

1. La descomposición de las señales inerciales en IMF con frecuencias variables permitió incrementar la observabilidad del sistema, caracterizar la representación tiempo-

frecuencia de las actividades humanas y reconocer seis actividades humanas.

2. La segmentación de las IMF de las señales inerciales permitió incrementar aún más la observabilidad del sistema y diferenciar las actividades dinámicas de las estáticas.
3. El uso de dos clasificadores en cascada permitió clasificar secuencialmente las actividades dinámicas y estáticas, ahorrar tiempo de cómputo y mejorar el desempeño global del clasificador.
4. La utilización de redes neuronales del tipo profundo, específicamente la LSTM bidireccional, permitió crear un modelo que aprende directamente de los datos de entrada, sin necesidad de extraer características de las señales manualmente.

Entre las limitaciones de este trabajo se pueden mencionar que se reconocieron tan solo seis actividades humanas de 30 sujetos. En trabajos futuros sería interesante reconocer una mayor cantidad de actividades humanas a un mayor número de sujetos. Esto tendría un impacto en la red que se usó, ya que a mayor cantidad de datos la red se comporta mejor. También se debe mencionar que las señales son de corta duración (2.56 s), lo que afectó el número de IMF en la que podría ser descompuesta. Es interesante entonces que la actividad se realiza durante un tiempo mayor, de manera tal de poder obtener un mayor número de IMF, y de esta forma incrementar aún más la observabilidad del sistema, y a su vez mejorar el reconocimiento del clasificador. Por otro lado, en el análisis de las señales se usaron la varianza y entropía de Shannon; sería interesante analizar las señales usando otras métricas como la curtosis, asimetría, análisis de fluctuación de tendencia, y exponentes de Lyapunov, y seguir comprendiendo la naturaleza de las señales desde diferentes panoramas.

Una dificultad mayor que se encontró fue el reconocer actividades similares como estar sentado y estar parado; en este sentido, se podría usar la información de otro sensor que ayude a diferenciarlas, como puede ser la estimación de la altitud usando un barómetro, ya que ésta es una variable que es diferente entre las dos actividades mencionadas y podría ayudar al clasificador a reconocerlas automáticamente. Finalmente, consideramos la posición y ubicación del teléfono inteligente al momento de adquirir los datos, colocado en una posición fija en la cintura del sujeto, no representa un ambiente real debido a que el teléfono suele ser transportado en el bolsillo del pantalón, en la mano o en una cartera; esto sería un reto interesante del clasificador ya que se enfrentaría a una gran variabilidad de las señales adquiridas.

Cabe resaltar que este proyecto es la base para el desarrollo de futuras investigaciones. Además, es un precedente del grupo de investigación en Bioingeniería, Señales y Microelectrónica (BISEMIC) de la Universidad Pontificia Bolivariana seccional Bucaramanga, el

cual hemos formado parte desde hace más de dos años. Actualmente, se están realizando en paralelo otras investigaciones, como el reconocimiento de actividades humanas usando señales electromiográficas por el estudiante Mario Otero, donde se desea reconocer los movimientos del brazo y de la mano, y el desarrollo de una aplicación en un teléfono inteligente para la clasificación en tiempo real de las actividades usando los sensores embebidos en el teléfono, realizado por los estudiantes Gabriel González y Manuel Díaz.

Respecto a la línea de investigación y los resultados obtenidos en este trabajo se envió una publicación sobre clasificación del reconocimiento de las actividades humanas en ASOVAC, una convención anual que se lleva a cabo en Venezuela, la cual ha sido aceptada por los organizadores del evento. Además, se ha enviado un artículo de investigación a una revista científica de alto impacto (Computers & Electrical Engineering de la editorial Elsevier Ltd.).

BIBLIOGRAFÍA

- Alonso, A. B., Artime, I. F., Rodríguez, M. Á., y Baniello, R. G. (2011). Dispositivos móviles. *EPSIG Ing. Telecomunicación Universidad de Oviedo*.
- Altun, K., y Barshan, B. (2010). Human activity recognition using inertial/magnetic sensor units. En *International workshop on human behavior understanding* (pp. 38–51).
- Anguita, D., Ghio, A., Oneto, L., Parra, X., y Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. En *Esann*.
- Ann, O. C., y Theng, L. B. (2014). Human activity recognition: a review. En *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)* (pp. 389–393).
- Arnott, P., Pfeiffer, G., y Tavel, M. (1984). Spectral analysis of heart sounds: Relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives. *Journal of Biomedical Engineering*, 6(2), 121–128. doi: 10.1016/0141-5425(84)90054-2
- Banos, O., Galvez, J.-M., Damas, M., Pomares, H., y Rojas, I. (2014). Window size impact in human activity recognition. *Sensors*, 14(4), 6474–6499.
- Bao, L., y Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. En *International conference on pervasive computing* (pp. 1–17).
- Bidargaddi, N., Sarela, A., Klingbeil, L., y Karunanithi, M. (2007). Detecting walking activity in cardiac rehabilitation by using accelerometer. En *2007 3rd international conference on intelligent sensors, sensor networks and information* (pp. 555–560).
- Bulling, A., Blanke, U., y Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3), 33.
- Chahuara, P., Fleury, A., Portet, F., y Vacher, M. (2012). Using markov logic network for on-line activity recognition from non-visual home automation sensors. En *International joint conference on ambient intelligence* (pp. 177–192).
- Chang Francis, H., Ping-Yen, L., Hsuan-Hao, C., Long, H., y Chi, S. (2019). Average entropy: Measurement of disorder for cardiac rr interval signals. *Physica A: Statistical Mechanics and its Applications*, 121533. doi: <https://doi.org/10.1016/j.physa.2019.121533>
- Chen, L., y Nugent, C. (2009). Ontology-based activity recognition in intelligent pervasive

- environments. *International Journal of Web Information Systems*, 5(4), 410–430.
- Colah's Blog. (2015). *Understanding LSTM Networks*. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. (Accedido: 2018-11-26)
- Colominas, M. A. (2016). Métodos guiados por los datos para el análisis de señales: contribuciones a la descomposición empírica en modos.
- Colominas, M. A., Schlotthauer, G., y Torres, M. E. (2014). Improved complete ensemble emd: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14, 19–29.
- Fawcett, T. (2006). An introduction to roc analysis. *Pattern recognition letters*, 27(8), 861–874.
- Ferrer Cid, P. (2017). *Comparativa de métodos de localización con smartphones* (B.S. thesis). Universitat Politècnica de Catalunya.
- Ferrer Mínguez, G. (2009). Integración kalman de sensores inerciales ins con gps en un uav.
- Guo, P., y Miao, Z. (2006). Projection histogram based human posture recognition. En *2006 8th international conference on signal processing* (Vol. 2).
- Hassan, A. R., y Subasi, A. (2016). Automatic identification of epileptic seizures from eeg signals using linear programming boosting. *computer methods and programs in biomedicine*, 136, 65–77. doi: 10.1016/j.cmpb.2016.08.013
- Heinz, E. A., Kunze, K. S., Gruber, M., Bannach, D., y Lukowicz, P. (2006). Using wearable sensors for real-time recognition tasks in games of martial arts-an initial experiment. *CIG, 2006*, 98–102.
- Hernández, F., Suárez, L. F., Villamizar, J., y Altuve, M. (2019). Human activity recognition on smartphones using a bidirectional lstm network. En *2019 xxii symposium on image, signal processing and artificial vision (stsva)* (pp. 1–5).
- Hochreiter, S., y Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., . . . Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903–995.
- Huang, N. E., y Wu, Z. (2008). A review on hilbert-huang transform: Method and its applications to geophysical studies. *Reviews of geophysics*, 46(2). doi: 10.1029/2007RG000228
- Kærgaard, K., Jensen, S. H., y Puthusserypady, S. (2016). A comprehensive performance analysis of eemd-blms and dwt-nn hybrid algorithms for ecg denoising. *Biomedical Signal Processing and Control*, 25, 178–187. doi: 10.1016/j.bspc.2015.11.012

- Karantonis, D. M., Narayanan, M. R., Mathie, M., Lovell, N. H., y Celler, B. G. (2006). Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE transactions on information technology in biomedicine*, 10(1), 156–167.
- Khan, A. M., Lee, Y.-K., Lee, S.-Y., y Kim, T.-S. (2010). Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. En *2010 5th international conference on future information technology* (pp. 1–6).
- Kwapisz, J. R., Weiss, G. M., y Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2), 74–82.
- Labate, D., La Foresta, F., Occhiuto, G., Morabito, F. C., Lay-Ekuakille, A., y Vergallo, P. (2013). Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ecg: A comparison. *IEEE Sensors Journal*, 13(7), 2666–2674. doi: 10.1109/JSEN.2013.2257742
- Laverde, R., Rueda, C., Amado, L., Rojas, D., y Altuve, M. (2018). Artificial neural network for laparoscopic skills classification using motion signals from apple watch. En *2018 40th annual international conference of the ieee engineering in medicine and biology society (embc)* (pp. 5434–5437).
- Lee, Y.-S., y Cho, S.-B. (2011). Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer. En *International conference on hybrid artificial intelligence systems* (pp. 460–467).
- Lin, W., Sun, M.-T., Poovendran, R., y Zhang, Z. (2008). Activity recognition using a combination of category components and local models for video surveillance. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(8), 1128–1139.
- Marcos Plaza, P. (2017). Uso de los smartphones en los laboratorios de prácticas de física.
- Mathie, M., Lovell, N. H., Coster, A., y Celler, B. (2002). Determining activity using a triaxial accelerometer. En *Proceedings of the second joint 24th annual conference and the annual fall meeting of the biomedical engineering society*[*engineering in medicine and biology* (Vol. 3, pp. 2481–2482).
- MathWorks. (2018). *What is Deep Learning? 3 things you need to know*. <https://la.mathworks.com/discovery/deep-learning.html>. (Accedido: 2018-11-25)
- MathWorks. (2019). *Options for training deep learning neural network*. <https://la.mathworks.com/help/deeplearning/ref/trainingoptions.html>. (Accedido: 2019-2-7)
- Miranda Calle, J. D., y Gamboa Entralgo, C. A. (2018). *Desarrollo de un sistema para la detección de movimientos sísmicos usando redes neuronales artificiales* (Bachelor's Thesis). Universidad Pontificia Bolivariana.

- Motin, M. A., Karmakar, C. K., y Palaniswami, M. (2019). Selection of empirical mode decomposition techniques for extracting breathing rate from ppg. *IEEE Signal Processing Letters*, 26(4), 592–596. doi: 10.1109/LSP.2019.2900923
- Organista-Sandoval, J., McAnally-Salas, L., y Lavigne, G. (2013). El teléfono inteligente (smartphone) como herramienta pedagógica. *Apertura*, 5(1), 6–19.
- Palencia-Vizcarra, R. d. J., y Palencia-Díaz, R. (2013). Teléfonos inteligentes y tabletas: ¿una herramienta o una barrera en la atención del paciente? *Medicina Interna de México*, 29(4), 404–409.
- Piergiovanni, A., y Ryoo, M. S. (2018). Fine-grained activity recognition in baseball videos. En *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 1740–1748).
- Pirttikangas, S., Fujinami, K., y Nakajima, T. (2006). Feature selection and activity recognition from wearable sensors. En *International symposium on ubiquitous computing systems* (pp. 516–527).
- Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Qi, W., Su, H., Yang, C., Ferrigno, G., De Momi, E., y Aliverti, A. (2019). A fast and robust deep convolutional neural networks for complex human activity recognition using smartphone. *Sensors*, 19(17), 3731.
- Recker, W. W., McNally, M. G., y Root, G. S. (1985). Travel/activity analysis: pattern recognition, classification and interpretation. *Transportation Research Part A: General*, 19(4), 279–296.
- Ren, Y., Suganthan, P., y Srikanth, N. (2014). A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. *IEEE Transactions on Sustainable Energy*, 6(1), 236–244. doi: 10.1109/TSTE.2014.2365580
- Satava, R. M. (2006). Robotics in colorectal surgery: telemonitoring and telerobotics. *Surgical Clinics*, 86(4), 927–936.
- Sun, L., Zhang, D., Li, B., Guo, B., y Li, S. (2010). Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. En *International conference on ubiquitous intelligence and computing* (pp. 548–562).
- Suárez Mantilla, L. F., y Ardila Navas, J. A. (2019). *Clasificación de registros fonocardiográficos usando descomposición empírica en modos y redes de gran memoria de corto plazo* (Bachelor’s Thesis). Universidad Pontificia Bolivariana.
- Torres, M. E., Colominas, M. A., Schlotthauer, G., y Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. En *2011 IEEE International conference on acoustics, speech and signal processing (icassp)* (pp. 4144–4147).

- Tsunoda, T., Komori, Y., Matsugu, M., y Harada, T. (2017). Football action recognition using hierarchical lstm. En *Proceedings of the ieee conference on computer vision and pattern recognition workshops* (pp. 99–107).
- Villavicencio, J. (2010). Introducción a series de tiempo. *Metodología de series de tiempo*, 4.
- Vogt, P., y Kuhn, J. (2012). Analyzing free fall with a smartphone acceleration sensor. *The Physics Teacher*, 50(3), 182–183.
- Vrigkas, M., Nikou, C., y Kakadiaris, I. A. (2015). A review of human activity recognition methods. *Frontiers in Robotics and AI*, 2, 28.
- Wang, L., Li, X., Ma, C., y Bai, Y. (2019). Improving the prediction accuracy of monthly streamflow using a data-driven model based on a double-processing strategy. *Journal of Hydrology*, 573, 733–745. doi: 10.1016/j.jhydrol.2019.03.101
- Wang, Z., Wu, D., Chen, J., Ghoneim, A., y Hossain, M. A. (2016). A triaxial accelerometer-based human activity recognition via eemd-based features and game-theory-based feature selection. *IEEE Sensors Journal*, 16(9), 3198–3207.
- Wei, W. W. (2006). Time series analysis. En *The oxford handbook of quantitative methods in psychology: Vol. 2*. Oxford University Press, USA.
- Wilmshurst, T. (2001). *An introduction to the design of small-scale embedded systems*. Palgrave.
- Wu, Z., y Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01), 1–41.
- Yu, S., y Qin, L. (2018). Human activity recognition with smartphone inertial sensors using bidir-lstm networks. En *2018 3rd international conference on mechanical, control and computer engineering (icmccce)* (pp. 219–224).
- Zhan, L., y Li, C. (2016). A comparative study of empirical mode decomposition-based filtering for impact signal. *Entropy*, 19(1), 13. doi: 10.3390/e19010013

ANEXO A

ARTÍCULO PRESENTADO EN UN EVENTO INTERNACIONAL

Se presenta a continuación el artículo titulado “Reconocimiento de actividades humanas usando EMD por conjuntos completa con ruido adaptativo mejorada y redes bidireccionales LSTM, el cual será presentado en la LXIX Convención Anual de la AsoVAC, a celebrarse en la Universidad Católica Andrés Bello, Caracas, Venezuela, del 20 al 22 de noviembre de 2019.”



RECONOCIMIENTO DE ACTIVIDADES HUMANAS USANDO EMD POR CONJUNTOS COMPLETA CON RUIDO ADAPTATIVO MEJORADA Y REDES BIDIRECCIONALES LSTM (HUMAN ACTIVITY RECOGNITION USING IMPROVED COMPLEMENTARY ENSEMBLE EMD WITH ASSISTED NOISE AND BIDIRECTIONAL LSTM NETWORKS)

Javier Villamizar, Paula Lizarazo, Miguel Altuve

Facultad de Ingeniería Eléctrica y Electrónica, Universidad Pontificia Bolivariana, Bucaramanga, Colombia

javier.villamizar.2015@upb.edu.co, paula.lizarazo.2015@upb.edu.co, miguel.altuve@upb.edu.co

El reconocimiento de la actividad humana usando la información brindada por sensores inerciales es un tema de mucha relevancia en los últimos años, debido a su amplia gama de aplicaciones, como el seguimiento de pacientes y el mejoramiento de actividades deportivas. En la actualidad, debido a la gran cantidad de dispositivos que nos monitorean constante y a la alta velocidad de procesamiento de computadores y teléfonos celulares, la disponibilidad y el procesamiento de los datos no es un problema, siendo uno de los retos más importantes, y el tema principal de este trabajo, el desarrollo de algoritmos para detectar las actividades que realizan las personas. Diferentes enfoques han sido propuestos para detectar las actividades humanas, desde la comparación de señales con umbrales fijos y adaptativos, hasta la aplicación de técnicas de aprendizaje automático y profundo. En particular, las técnicas basadas en aprendizaje profundo, además de superar a otras técnicas de aprendizaje automático en varios campos, presentan una serie de ventajas que las hacen idóneas para este tipo de problemas. En este trabajo realizamos la detección de seis actividades humanas usando redes LSTM bidireccionales y, como representación de las señales inerciales a ser explotadas por estas redes, e incrementar la observabilidad del sistema, se propone la representación de las señales en funciones de modo intrínseco (IMF), con oscilaciones de alta y baja frecuencias, usando la descomposición empírica en modos por conjuntos completa con ruido adaptativo mejorada (ICEEMDAN). Las actividades clasificadas en este trabajo fueron: caminar (W), subir escaleras (WU), bajar escaleras (WD), estar sentado (ST), estar parado (SD) y estar acostado (LD), a través nueve señales de sensores inerciales (acelerómetro y giroscopio) de 2.56 segundos (128 muestras) recolectadas por un teléfono inteligente a 30 voluntarios, y tomadas del UCI Machine Learning Repository <https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>. En primer lugar, las señales inerciales fueron estandarizadas para llevarlas a una misma escala usando puntajes Z y fueron descompuestas en IMF usando la ICEEMDAN. Posteriormente, las IMF fueron segmentadas (divididas) en nueve segmentos de 64 muestras con solapamiento de 8 muestras. Estos segmentos fueron utilizados como entrada a una red LSTM para el respectivo entrenamiento y prueba. Se utilizaron dos redes en cascada para detectar las seis actividades: una primera red con cuatro salidas para identificar las actividades dinámicas (W, WU y WD) y las estáticas como una sola clase denominada “statics” (ST, SD, LD), obteniendo una exactitud del 98.86%. Los registros asignados a la clase de statics fueron introducidos a una segunda red para clasificar sus tres actividades, dando una exactitud de 88.46%. La exactitud total que se logró para clasificar las seis actividades fue 92.91%. Este desempeño fue, en gran parte logrado, gracias a que la ICEEMAN nos permitió extraer información que se encontraba implícita en la señal (invisible a simple vista), la segmentación de las señales permitió que la red pudiera discriminar correctamente entre actividades estáticas y dinámicas. No obstante, observamos que la red tuvo dificultad en clasificar las actividades ST y SD debido a que son actividades muy similares entre sí.



Palabras clave: Sensores inerciales, aprendizaje automático, aprendizaje profundo, descomposición empírica en modos, reconocimiento de actividades humanas.

ANEXO B

ARTÍCULO ENVIADO A UNA REVISTA

A continuación se muestra el artículo de investigación titulado “Human Activity Recognition using Improved Complete Ensemble EMD with Adaptive Noise and Long Short-Term Memory Neural Networks”, enviado a la revista *Computers & Electrical Engineering* de la editorial Elsevier Ltd., la cual tiene un factor de impacto de 2.189, CiteScore de 3.02, Source-Normalized Impact per Paper (SNIP) de 1.395, SCImago Journal Rank (SJR) de 0.443, H Index de 49, y tasa de aceptación de 11 %.

Human Activity Recognition using Improved Complete Ensemble EMD with Adaptive Noise and Long Short-Term Memory Neural Networks

Miguel Altuve*, Javier Villamizar, Paula Lizarazo

*Faculty of Electrical and Electronic Engineering, Pontifical Bolivarian University,
Bucaramanga, Colombia*

Abstract

The recognition of human activities is a topic of great relevance due to its wide range of applications. Different approaches have been proposed to recognize human activities, ranging from the comparison of the signals with thresholds to the application of deep and machine learning techniques. In this work, the classification of six human activities is performed using bidirectional LSTM networks and, as a representation of the inertial signals to be exploited by these networks, the representation of these signals in intrinsic mode functions (IMF) is proposed using improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN). The classified activities were: walking, walking downstairs, walking upstairs, standing, sitting, and lying down. Records with inertial signals (accelerometer and gyroscope) of 2.56 seconds were collected by a smartphone to 30 volunteers and are available in the UCI Machine Learning Repository. First, inertial signals were standardized to take them to the same scale and were decomposed into IMF using ICEEMDAN. IMF were then segmented (split) into nine segments of 1.28 s with 12.5% overlap and introduced to a first network with four outputs to identify the dynamic activities and the statics as a single class called “statics”, giving 98.86% accuracy. Then, the non-

*Corresponding author

Email addresses: miguel.altuve@upb.edu.co (Miguel Altuve),
javier.villamizar.2015@upb.edu.co (Javier Villamizar),
paula.lizarazo.2015@upb.edu.co (Paula Lizarazo)

segmented IMF of the records assigned to the statics class were introduced to a second network to classify their three activities, giving an accuracy of 88.46%. In total, 92.91% accuracy was obtained to classify the six human activities. This performance is because ICEEMAN allowed the extraction of information that was embedded in the signal, and the segmentation of the signals allowed the network to discriminate between static and dynamic activities.

Keywords: Empirical Mode Decomposition, Intrinsic mode function, Improved Complete Ensemble EMD, Long Short-Term Memory, Human Activity Recognition

1. Introduction

Human beings have the ability to perform activities, even simultaneously, as simple as walking or as complex as sewing while talking. However, technological development and high availability of resources have negatively influenced our daily activities, favoring a sedentary lifestyle, bad postures, and an unbalanced diet. Automatically identifying and monitoring the activities we carry out have been of increasing interest given the large number of applications that it may have, for example, fall detection [1], recognition of sedentary behavior [2], video surveillance [3], comfort in smart homes [4] and for driver assistance and situational awareness in intelligent vehicles [5].

The complexity and diversity of activities that humans can perform, and the high dimensionality of the collected data make recognizing human activities a challenging but promising task [6]. In addition, intraclass variability (the same activity may vary from one subject to another) and interclass similarity (different activities may express similar forms) make it even more difficult the recognition of human activities. The identification of human activities can be achieved through the processing of signals or sequence of images collected from one or more sensors, for example, by analyzing videos and images from cameras or by analyzing motion data from inertial sensors (accelerometer, gyroscope, magnetometers). Moreover, the incorporation of inertial sensors in smartphones

has allowed the recognition of human activities with applications related to well-being, with a single small portable device that goes unnoticed to the user [7, 8, 9, 10].

A machine learning-based human activity recognition system is usually composed of (i) a preprocessing step that looks to enhance some waveform or dynamic in the signal and to reduce noise and interference; transformation of the signal and its representation in other domains are common practices as preprocessing steps; (ii) a feature extraction step that extracts features from the original, transformed or mapped signals using descriptive statistics and non-linear measures; and (iii) a classification step that makes the decision by exploiting a feature set. The performance of a machine learning-based classification system depends heavily on the quantity and quality of the feature set. In this work, we focus our attention on the representation of inertial signals on a time-frequency domain that can be used in an automatic human activity recognition system.

Techniques for decomposing a signal into several components, such as the Fourier series, the wavelet transform, the Karhunen-Loève transform and the empirical mode decomposition (EMD), are powerful tools for eliminating undesirable components, identifying desirable components and reducing the amount of data necessary for storage and transmission. In particular, the EMD is an adaptive and data-driven, time-frequency analysis technique that decomposes a time domain signal into intrinsic mode functions (IMF), with the same length as the original signal and that reside in the time domain but have varying frequencies [11]. Noise-assisted versions of the EMD, such as the Ensemble EMD (EEMD) [12], the Complete EEMD with Adaptive Noise (CEEMDAN) [13] and the Improved CEEMDAN (ICEEMDAN) [14], have been proposed to address some of issues of the EMD, such as the mode mixing problem in which oscillations with different time scales reside in the same IMF or that oscillations with similar time scale reside in different IMF [11, 12]. EMD family techniques break down non-linear and non-stationary signals without defining a basis function and without establishing a priori the level of decomposition [15, 16] and have outperformed other techniques in several domains, such as in the analysis

of electroencephalogram with epileptic seizures [17], to improve the prediction accuracy of monthly streamflow [18], and to extract the respiratory rate from the photoplethysmogram [16]. These properties made EMD-based techniques particularly suitable for decomposing inertial signals into IMF components that can be exploited by a machine learning technique to classify human activities. The aim of this paper is to decompose gyroscopes and accelerometers signals into IMF issued from the ICEEMDAN technique and to exploit these IMF by bidirectional LSTM recurrent neural networks to identify six human activities, three statics (sitting, standing, and lying down) and three dynamics (walking, walking upstairs, walking downstairs). Data were collected from 30 volunteers using inertial sensors embedded in a smartphone that was worn on the subjects' waist.

2. Materials and methods

2.1. Inertial signals

Inertial sensor signals from the UCI machine learning repository¹ were used in this work. The dataset is composed of 10299 recordings of 2.56 s duration of nine signals: three linear accelerations, three angular velocities, and three estimated body accelerations. Signals were acquired at a sampling rate of 50 Hz using accelerometer and gyroscope inertial sensors embedded in a Samsung Galaxy S II smartphone from 30 volunteers (ages ranging from 19 to 48 years) while wearing the smartphone on the waist [19]. Six activities were performed by the participants and are included in the dataset with an appropriate label: standing (ST), sitting (SD), lying down (LD), walking (WK), walking downstairs (WD), and walking upstairs (WU). We denote dynamic activities to WK, WD and WU, and static activities to ST, SD and LD. The dataset in the repository was randomly partitioned into training (70%) and test (30%) sets. We

¹<https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

denoted A_x the linear acceleration from the accelerometer in the x axis, G_x the angular velocity from the gyroscope in the x axis, and TA_x the estimated body
 80 accelerations in the x axis; similarly for A_y , A_z , G_y , G_z , TA_y and TA_z .

2.2. Decomposition of the inertial signals into IMF

Inertial signals were decomposed into IMF using the ICEEDMAN technique. This decomposition looks to increase the observability of the system by representing each signal into a finite set (M) of fast and slow oscillations components,
 85 $IMF_1(t), IMF_2(t), \dots, IMF_M(t)$. The original signal is expressed as the sum of all its IMF plus a residue $r(t)$ as: $x(t) = \sum_{i=1}^M IMF_i(t) + r(t)$ [11, 20]. IMF are sorted by frequency: the first ones capture rapid oscillations in the signal while the last ones represent slow oscillations.

Given that ICEEMDAN is an adaptive and data-driven technique, a different
 90 number of IMF components was obtained for every record, ranging from three to eight IMF. In this sense, we chose to decompose the inertial signals into five IMF components. Those signals that were not possible to decompose into five IMF were not used in the classification task. 7012 records were thus considered to train the classifier and 2834 records were used to test the classifier.

95 2.3. Classification model

LSTM networks are considered to exploit the IMF given that these networks have memory cells to store information over time. Moreover, an approach based on bidirectional LSTM networks was employed to consider future and past information of the IMF in the classification task. Two bidirectional LSTM classification models in cascade were proposed: one to classify dynamic activities
 100 (WK, WU and WD) and the other to classify static activities (ST, SD and LD). The first model considers all records to produce four outputs: three to predict the dynamic activity classes and another output to predict the static classes as a single class. Then, records classified as statics are passed through the second
 105 model, with three outputs, to predict the static classes. The architecture of both models is the same, however, they differ in their input data and parameters.

The IMF signals input to the classifiers were standardized using the Z-score. To assure the generalization of the models, the mean and standard deviation of the IMF signals were computed in the training set; these parameters were then
110 used to standardize the IMF signals in the training and test sets.

Given that dynamic activities involve repetitive movements to perform such activities (a gait cycle lasts around 1 s [21]) and have patterns that are not observable in the signals of static activities, a segmentation procedure was performed to the IMF signals of dynamic activities. Specifically, IMF signals were
115 segmented (split) into segments of 1.28 s (64 samples) with 50% overlap between segments; IMF signals were split into nine parts. The duration of segments is enough to accommodate a complete cycle of the dynamic activities that we want to classify. Segmentation ensures that the 2.56 s signal (128 samples) is observed in nine consecutive segments.

The input of the first classification model is the standardized and segmented
120 IMF signals. The input layer of this model is composed of 9 (inertial) signals \times 5 IMF/signals \times 9 segments/IMF = 405 nodes. The input of the second classification model is the standardized IMF signals of the records classified by the first model as static activities. The input layer of this model is composed
125 of 9 (inertial) signals \times 5 IMF/signals = 45 nodes. The first classifier processes arrays of size 405×64 while the second classifier processes arrays of size 45×128 (features \times samples).

The output of the first model is taken as the probability for each of the three dynamic classes and the static classes as one single class. The output
130 of the second model is taken as the probability for each of the three static classes. We used the cross-entropy as loss function and the Adam optimization algorithm. The number of epochs was set to 40, the learning rate was initially set to 0.001 with a drop factor of 0.1 every 50 epochs. The batch size was set to 1098 in the first classifier and to 128 in the second classifier. The hyperbolic
135 tangent function (tanh) was used as the activation function to update the cell and hidden state, and the sigmoid function was used as the activation function to apply to the gates.

This classification scheme is schematized in Figure 1.

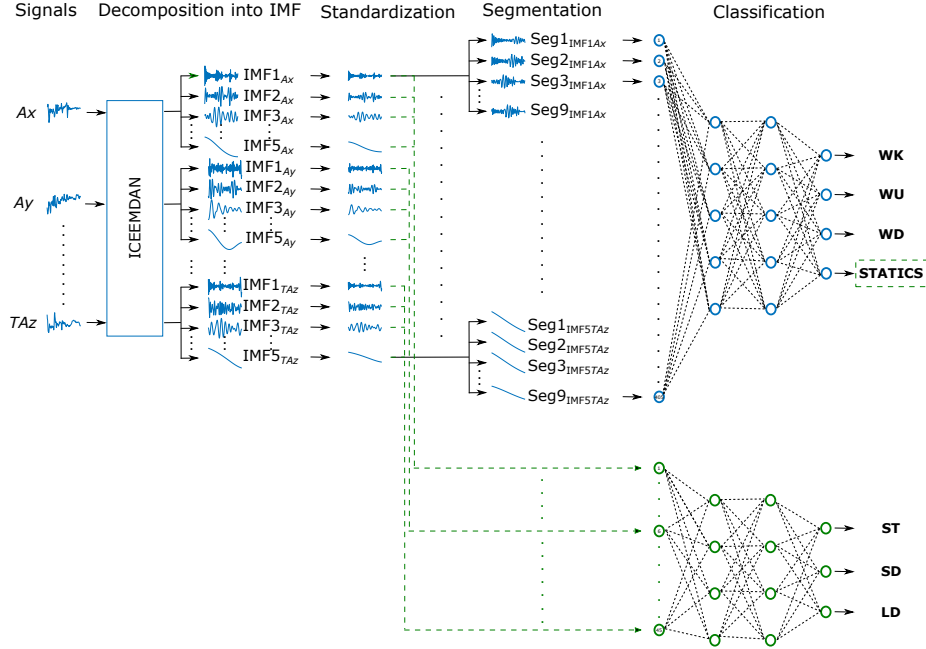


Figure 1: Human activity classification scheme.

2.4. Classification evaluation metrics

140 A confusion matrix was used to visualize the discrepancies between the predicted and the true classes, as depicted in Table 1, where the rows correspond to the true class (output class) and the columns correspond to the predicted class (target class). The diagonal corresponds to the correct classifications made by the classifier. From the confusion matrix, accuracy (Acc), precision (Pre) and
 145 recall (Rec) were computed as shown in the Table.

3. Results

Figure 2 shows an example of the accelerometer and gyroscope signals in the x axis during walking. Figure 3 shows examples of the decomposition of the signals shown in Figure 2 into five IMF. In this figure, we can clearly observe
 150 that the signal's content is handed out in different ways in the IMF.

Table 1: Confusion matrix. Equations for precision (Pre), recall (Rec), and accuracy (Acc) are also shown for reference.

True class	Predicted class				Rec
	$Class_1$	$Class_2$	\dots	$Class_n$	
$Class_1$	a_{11}	a_{12}	\dots	a_{1n}	$\frac{a_{11}}{\sum_{j=1}^n a_{1j}}$
$Class_2$	a_{21}	a_{22}	\dots	a_{2n}	$\frac{a_{22}}{\sum_{j=1}^n a_{2j}}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Class_n$	a_{n1}	a_{n2}	\dots	a_{nn}	$\frac{a_{nn}}{\sum_{j=1}^n a_{nj}}$
Pre	$\frac{a_{11}}{\sum_{i=1}^n a_{i1}}$	$\frac{a_{22}}{\sum_{i=1}^n a_{i2}}$	\dots	$\frac{a_{nn}}{\sum_{i=1}^n a_{in}}$	$Acc = \frac{\sum_{i=1}^n a_{ii}}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}}$

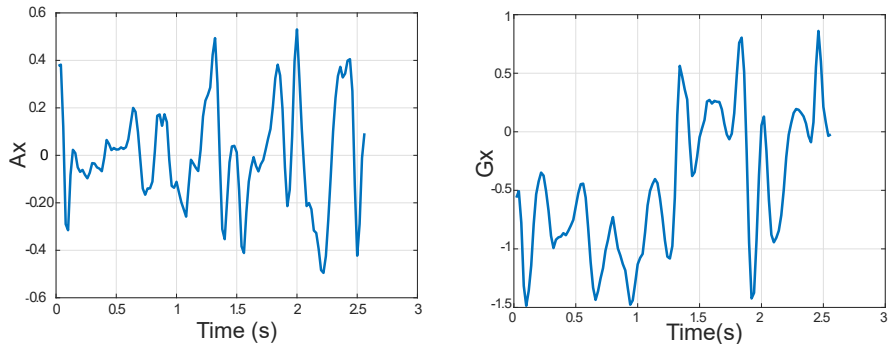


Figure 2: Examples of accelerometer (left) and gyroscope (right) signals in the x axis during walking.

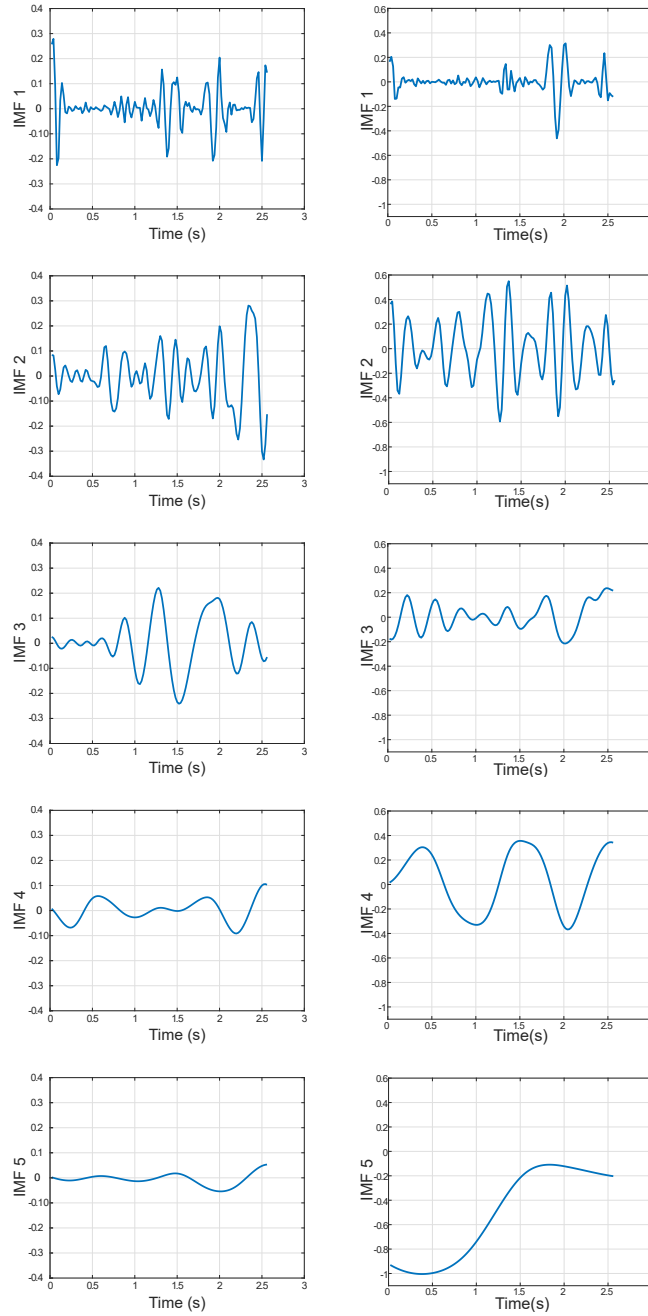


Figure 3: Decomposition of the accelerometer (left) and gyroscope (right) signals shown in Figure 2 into five IMF (rows).

Figure 4 shows an example of the segmentation carried out to an IMF.

Tables 2 and 3 show confusion matrices for the first and second classification models, respectively. Table 4 shows the complete confusion matrix that results from the combination of the confusion matrices of the first and second classification models.

155

Table 2: Confusion matrix of the first classification model.

True class	Predicted class				<i>Rec (%)</i>
	WK	WU	WD	Statics	
WK	469	1	18	0	96.1
WU	1	462	0	0	99.8
WD	1	9	406	0	97.6
Statics	0	1	1	1465	99.9
<i>Pre (%)</i>	99.6	97.7	95.5	100	98.87

Table 3: Confusion matrix of the second classification model.

True class	Predicted class			<i>Rec (%)</i>
	ST	SD	LD	
ST	351	107	0	76.6
SD	39	481	0	92.5
LD	0	23	464	95.3
<i>Pre (%)</i>	90	78.7	100	88.46

4. Discussion

The proposed approach correctly discriminates between a dynamic activity and a static one, producing an incorrect classification in only two static activity records as dynamic activities. The first classification model almost correctly recognizes all records, even activity with similar patterns (WU vs WD). This first classifier incorrectly assigned 18 walking activity records to the activity

160

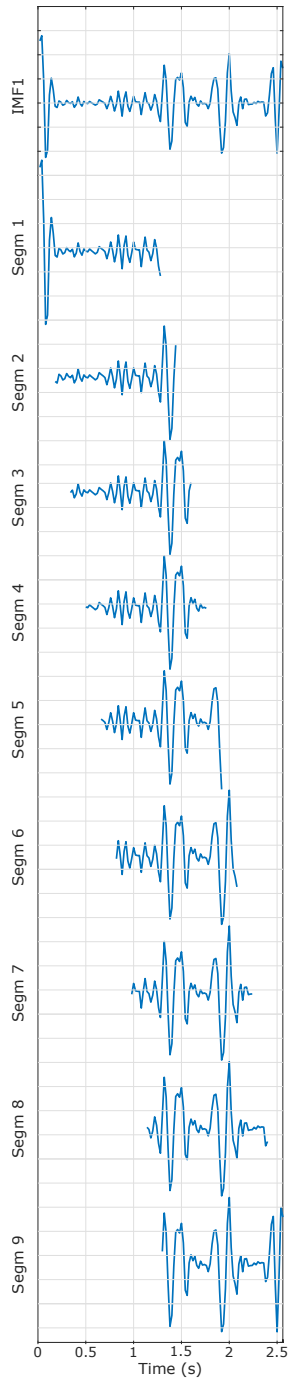


Figure 4: Example of segmentation of an IMF into nine segments.

Table 4: Confusion matrix of the final classification model.

True class	Predicted class						<i>Rec</i> (%)
	WK	WU	WD	ST	SD	LD	
WK	469	1	18	0	0	0	96.1
WU	1	462	0	0	0	0	99.8
WD	1	9	406	0	0	0	97.6
ST	0	1	0	351	107	0	76.5
SD	0	0	0	39	481	0	92.5
LD	0	0	1	0	23	464	95.1
<i>Pre</i> (%)	99.6	97.7	95.5	90	78.7	100	92.91

of walking down. Concerning the second classifier, the model had difficulty to differentiate when a person is sitting or standing. Moreover, 23 records belonging to lying down were assigned to standing. The proposed classification approach was able to differentiate the six human activities with very few errors: 201 incorrectly classified records from a universe of 2834 records (7% error). Dynamic activities, such as walking, walking up and walking down, were easily identified since they contain patterns that were exploited by the first classifier thanks to the segmentation of IMF: 30 records with dynamic activities were incorrectly identified from a universe of 1367 records (2.20% error). Although static activities, such as sitting, standing or lying down, do not contain a relevant characteristic pattern, the vast majority of these activities were correctly identified: 169 records with static activities were incorrectly identified from a universe of 1467 records (11.52% error).

Results from this work are comparable with the best-related works using the same database. For instance, using a convolutional neural network, an accuracy of 94.79% was obtained by processing the raw signals and 95.75% by processing the Fourier transform of the signals [22], an accuracy of 91.38% was obtained using the 2D wavelet transform and 95.18% using the discrete Fourier transform [23], an accuracy of 91.76% was obtained using hidden Markov models [24],

an accuracy of 83.43% was obtained using a recurrent neural network [25], and an accuracy of 93.79% was obtained using bidirectional LSTM networks [26]. In a recent work, we have obtained accuracy of 92.67% using a bidirectional LSTM network that exploit the raw signals [27].

185 5. Conclusions

This work focused on the recognition of six human activities (walking, walking up, walking down, sitting, standing and lying down) by exploiting nine inertial signals (derived from gyroscope and accelerometer) decomposed into intrinsic mode function by two classification models based on bidirectional LSTM
190 networks, without manually extracting features from the signals. The easiest activities to recognize were walking up ($Rec = 99.8\%$, $Pre = 97.7\%$) and walking ($Rec = 96.1\%$, $Pre = 99.6\%$) while the hardest activities to recognize were sitting ($Rec = 76.5\%$, $Pre = 90.0\%$) and standing ($Rec = 92.5\%$, $Pre = 78.7\%$).

The performance of the classification of human activities achieved in this
195 work is promising in terms of the high performance obtained and innovative in terms of the use of the improved complete ensemble empirical mode decomposition with adaptive noise to map the inertial signals into the intrinsic mode functions (time-frequency) domain. This decomposition allowed us to increase the observability of the system (one signal in the time domain was mapped to
200 five signals in the same time domain but with varying frequencies), and was particularly useful in the process of classification of human activities since there was a greater amount of information to be exploited by the bidirectional LSTM networks.

Conflict of Interest

205 The authors have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Author Contributions

Miguel Altuve contributed to the conception of the study, the supervision of
210 the experiments, the discussion of the results and the writing and revising of the
manuscript. Javier Villamizar and Paula Lizarazo carried out the experiment
and contributed to the analysis of the results and the writing and revising of
the manuscript.

References

- 215 [1] G. Mastorakis, D. Makris, Fall detection system using kinect's infrared
sensor, *Journal of Real-Time Image Processing* 9 (4) (2014) 635–646.
- [2] E. Kańtoch, Recognition of sedentary behavior by machine learning anal-
ysis of wearable sensors during activities of daily living for telemedical
assessment of cardiovascular risk, *Sensors* 18 (10) (2018) 3219.
- 220 [3] W. Lin, M.-T. Sun, R. Poovendran, Z. Zhang, Activity recognition using
a combination of category components and local models for video surveil-
lance, *IEEE Transactions on Circuits and Systems for Video Technology*
18 (8) (2008) 1128–1139.
- [4] L. Chen, C. D. Nugent, H. Wang, A knowledge-driven approach to activity
225 recognition in smart homes, *IEEE Transactions on Knowledge and Data
Engineering* 24 (6) (2012) 961–974.
- [5] E. Ohn-Bar, A. Tawari, S. Martin, M. M. Trivedi, On surveillance for safety
critical events: In-vehicle video networks for predictive driver assistance
systems, *Computer Vision and Image Understanding* 134 (2015) 130–140.
- 230 [6] S. Zhang, Z. Wei, J. Nie, L. Huang, S. Wang, Z. Li, A review on human
activity recognition using vision-based method, *Journal of healthcare engi-
neering* 2017 (2017) 1–31. doi:10.1155/2017/3090343.

- [7] L. Sun, D. Zhang, B. Li, B. Guo, S. Li, Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations, in: International conference on ubiquitous intelligence and computing, Springer, 2010, pp. 548–562. 235
- [8] N. Roy, A. Misra, D. Cook, Ambient and smartphone sensor assisted activity recognition in multi-inhabitant smart environments, Journal of ambient intelligence and humanized computing 7 (1) (2016) 1–19. doi:10.1007/s12652-015-0294-7. 240
- [9] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, D. Anguita, Transition-aware human activity recognition using smartphones, Neurocomputing 171 (2016) 754–767. doi:10.1016/j.neucom.2015.07.085.
- [10] M. M. Hassan, M. Z. Uddin, A. Mohamed, A. Almogren, A robust human activity recognition system using smartphone sensors and deep learning, Future Generation Computer Systems 81 (2018) 307–313. doi:10.1016/j.future.2017.11.029. 245
- [11] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454 (1971) (1998) 903–995. doi:10.1098/rspa.1998.0193. 250
- [12] Z. Wu, N. E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in adaptive data analysis 1 (01) (2009) 1–41. doi:10.1142/S1793536909000047. 255
- [13] M. E. Torres, M. A. Colominas, G. Schlotthauer, P. Flandrin, A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2011, pp. 4144–4147. doi:10.1109/ICASSP.2011.5947265. 260

- [14] M. A. Colominas, G. Schlotthauer, M. E. Torres, Improved complete ensemble emd: A suitable tool for biomedical signal processing, *Biomedical Signal Processing and Control* 14 (2014) 19–29. doi:10.1016/j.bspc.2014.06.009.
- 265
- [15] D. Labate, F. La Foresta, G. Occhiuto, F. C. Morabito, A. Lay-Ekuakille, P. Vergallo, Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ecg: A comparison, *IEEE Sensors Journal* 13 (7) (2013) 2666–2674. doi:10.1109/JSEN.2013.2257742.
- 270
- [16] M. A. Motin, C. K. Karmakar, M. Palaniswami, Selection of empirical mode decomposition techniques for extracting breathing rate from ppg, *IEEE Signal Processing Letters* 26 (4) (2019) 592–596. doi:10.1109/LSP.2019.2900923.
- [17] A. R. Hassan, A. Subasi, Automatic identification of epileptic seizures from eeg signals using linear programming boosting, *computer methods and programs in biomedicine* 136 (2016) 65–77. doi:10.1016/j.cmpb.2016.08.013.
- 275
- [18] L. Wang, X. Li, C. Ma, Y. Bai, Improving the prediction accuracy of monthly streamflow using a data-driven model based on a double-processing strategy, *Journal of Hydrology* 573 (2019) 733–745. doi:10.1016/j.jhydrol.2019.03.101.
- 280
- [19] D. Anguita, A. Ghio, L. Oneto, X. Parra, J. L. Reyes-Ortiz, A public domain dataset for human activity recognition using smartphones, in: *Proceedings of the 21th International European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2013, pp. 437–442.
- 285
- [20] M. A. Colominas, G. Schlotthauer, M. E. TORRES, P. Flandrin, Noise-assisted emd methods in action, *Advances in Adaptive Data Analysis* 4 (04) (2012) 1250025. doi:10.1142/S1793536912500252.
- 290

- [21] M. P. Murray, A. B. Drought, R. C. Kory, Walking patterns of normal men, *The Journal of Bone & Joint Surgery* 46 (2) (1964) 335–360.
- [22] C. A. Ronao, S.-B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, *Expert Systems with Applications* 59 (2016) 235–244.
- 295
- [23] W. Jiang, Z. Yin, Human activity recognition using wearable sensors by deep convolutional neural networks, in: *Proceedings of the 23rd ACM international conference on Multimedia*, ACM, 2015, pp. 1307–1310.
- [24] C. A. Ronao, S.-B. Cho, Human activity recognition using smartphone sensors with two-stage continuous hidden markov models, in: *Natural computation (ICNC), 2014 10th international conference on*, IEEE, 2014, pp. 681–686.
- 300
- [25] M. Inoue, S. Inoue, T. Nishida, Deep recurrent neural network for mobile human activity recognition with high throughput, *Artificial Life and Robotics* 23 (2) (2018) 173–185.
- 305
- [26] S. Yu, L. Qin, Human activity recognition with smartphone inertial sensors using bidir-lstm networks, in: *2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, IEEE, 2018, pp. 219–224.
- [27] F. Hernández, L. F. Suárez, J. Villamizar, M. Altuve, Human activity recognition on smartphones using a bidirectional lstm network, in: *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, IEEE, 2019, pp. 1–5.
- 310