



**UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH**

Departament d'Enginyeria Electrònica

Una contribución a la evaluación de la adherencia a hábitos de vida saludables basado en aplicaciones móviles

Tesis doctoral presentada como requisito parcial para la obtención del título de Doctor
por la Universitat Politècnica de Catalunya, dentro del Programa de Doctorado en
Ingeniería Electrónica

Por

Federico Guede Fernández

Codirectores:

Mireya Fernández Chimeno

Lluís Capdevila Ortís

Barcelona, junio 2018

Agradecimientos

En primer lugar querría agradecer a los directores Mireya Fernández y Lluís Capdevila por haber confiado en mí, haberme dado la oportunidad de trabajar con ellos y por sus consejos, sin su apoyo este trabajo no habría sido posible. También agradecer a Miguel Ángel García y Juan Ramos por su ayuda en diferentes fases de esta tesis. A mis compañeros Víctor Ferrer y Marc Mateu con los que he trabajado más estrechamente. A todos los con los que he trabajado en el laboratorio de instrumentación electrónica y biomédica: Noelia, Bernat, Hadis, Tomás, Antonio, Ángel, Albert, Sergio y tantos otros que han pasado por el laboratorio. Una mención especial para Alfonso que siempre está dispuesto a ayudar y nos anima a todos.

Querría agradecer a mis padres Carmina y Federico, a mi hermana Clara, a mi cuñado Pablo y al resto de mi familia por su cariño y afecto a lo largo de estos años.

Por últimos a mis amigos que siempre han estado ahí apoyándome a lo largo del desarrollo de la tesis.

Resumen

La adherencia a un estilo de vida saludable es un factor muy importante para alargar años de vida y aumentar su calidad. Los principales hábitos de vida saludable son: la actividad física, la dieta y la calidad del sueño. Hoy en día muchas personas utilizan un smartphone y lo llevan encima todo el día. El objetivo de esta tesis es demostrar la viabilidad de la evaluación de la adherencia a hábitos de vida saludables mediante aplicaciones móviles y sensores ya sean del propio smartphone o conectados externamente. Para ello, se utiliza el sensor de acelerometría para evaluar la actividad física y el gasto calórico asociado. En trabajos previos podemos encontrar clasificadores de actividad física a partir de los datos de estos sensores pero las medidas las realizan en un entorno de laboratorio o con el smartphone ubicado en una posición determinada. A partir de los datos de 26 sujetos recogidos durante una semana se ha alcanzado un 75.6 % de F1-score de la clasificación de actividades y un 3.18 % de error de estimación de gasto calórico.

Por otro lado, la variabilidad de la frecuencia cardíaca (VFC) puede servir de indicador de conductas relacionadas con la salud y la condición física. Se ha diseñado un sistema para evaluar la VFC utilizando como sensor la cámara trasera del smartphone. Para ello se ha utilizado la técnica de fotopleletismografía. En trabajos previos se ha utilizado esta técnica en smartphones para obtener el ritmo cardíaco pero no se ha comparado la variabilidad del ritmo cardíaco latido a latido. El sistema propuesto utiliza la GPU para procesar la imágenes en tiempo real. Los resultados obtenidos se han comparado con el electrocardiograma y con un dispositivo de fotopleletismografía de referencia. Para ello, se ha evaluado la desviación estándar del error cometido en la detección del latido cardíaco y el grado de acuerdo de los índices de VFC. Esta evaluación se ha realizado en 23 sujetos y se han comparado los resultados obtenidos con dos modelos de smartphone. La desviación estándar del error en la detección del latido cardíaco obtenida entre el smartphone y el electrocardiograma es de 5.4 ms, mientras que entre el dispositivo de referencia de fotopleletismografía y el electrocardiograma es de 4.9 ms.

Por otro lado, se ha desarrollado una aplicación para el análisis conjunto de la actividad física y el ritmo cardíaco. Se recogieron los datos de 11 personas utilizando esta aplicación, divididas en dos grupos de 5 y 6 personas durante 3 y 6 semanas respectivamente. A partir del análisis de los datos recogidos se ha encontrado que el nivel de la actividad física desciende a lo largo del tiempo y que existe alguna asociación entre la constancia en la práctica de la actividad física y los cambios en el estado de ánimo. Sin embargo, estas asociaciones se han de tomar con precaución debido al reducido número de sujetos que han participado en este estudio. Por lo tanto, el sistema desarrollado supone un punto de partida para evaluar la adherencia a un estilo de vida saludable de forma unificada en una única aplicación.

Finalmente, una de las consecuencias de llevar un estilo de vida poco saludable es el empobrecimiento de la calidad del sueño que puede provocar la somnolencia diurna.

Esto puede resultar un grave peligro para la salud, por ejemplo si se produce mientras se está al volante. Para prevenir esto, se ha propuesto y validado un sistema de detección de somnolencia temprana a partir del análisis de la señal respiratoria basado en la variabilidad del ritmo respiratorio. El algoritmo diseñado ha sido validado con 15 sujetos y se ha obtenido una especificidad del 96.6 % y una sensibilidad del 90.3 %.

Abstract

The adherence to a healthy lifestyle plays a key role for increasing life expectancy and living better. The main habits of healthy lifestyle are: physical activity, diet and sleep quality. Nowadays, many people use a smartphone and carry it all day. The objective of this thesis is to demonstrate the feasibility of the evaluation of the adherence to a healthy lifestyle by means of smartphone applications and sensors, whether internal or externally connected. On the one hand, the accelerometer sensor is used to evaluate the physical activity and the associated energy expenditure. In previous research, we can find classifiers of physical activity from data of this sensor but the measurements were performed in a laboratory environment or with smartphone fixed to a specific position. From the collected data during a week of 26 subjects, a 75.6% of F1-score of the classification of activities has been achieved and a 3.18% of error in the energy expenditure estimation.

On the other hand, the heart rate variability (HRV) can serve as indicator of behaviours related to health and physical condition. A system has been designed to evaluate the HRV using the rear camera of the smartphone as a sensor. For this purpose, the photoplethysmography technique has been used. In previous research, this technique has been used in smartphones in order to obtain the heart rate but it has not been assessed the beat-to-beat HRV. The proposed system uses the GPU for image processing in real time. The obtained results have been compared with the electrocardiogram and with a reference photoplethysmography device. For that, the standard deviation of error made for the beat detection and the level of agreement of HRV indices have been assessed. This assessment has been performed with 23 subjects and the results obtained for two different smartphone models have been compared. The standard deviation of error of heart rate detection between smartphone and electrocardiogram obtained was 5.4 ms, while between electrocardiogram and reference photoplethysmography device was 4.9 ms.

On the other hand, an application for the ensemble analysis of physical activity and heart rate has been developed. Using this application, the data of 11 people was collected, they have divided in two groups of 5 and 6 people during 3 and 6 weeks respectively. From the analysis of the collected data, it has been found that the level of physical activity decreases over the time and there is some association between the constancy of the practice of physical activity and changes in mood. However, these association should be taken with caution due to the reduced number of subjects which were involved in this study. Therefore, the developed system is a starting point in order to evaluate the adherence to a healthy lifestyle in a unified way with an single application.

Finally, one of the consequences of leading an unhealthy lifestyle is the decreasing of quality of sleep that can cause daytime sleepiness. This can be a serious health risk, for example if it occurs while driving. To prevent this, an early drowsiness detection

system based on the analysis of respiratory signal and respiratory rate variability has been proposed and validated. The designed algorithm has been assessed with 15 subjects and a specificity of 96.6% and a sensitivity of 90.3% has been obtained.

Índice general

Agradecimientos	iii
Abreviaturas	xvii
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	4
1.3 Esquema de la tesis	5
2 Estado del arte	7
2.1 Actividad física: acelerometría y smartphone	7
2.1.1 Métodos de medida del gasto calórico	7
2.1.2 Medida del gasto calórico con los sensores del smartphone	9
2.2 Variabilidad de la frecuencia cardíaca	11
2.3 Somnolencia durante la conducción	14
2.4 Sensores en los móviles	17
3 Estimación de la actividad física realizada a partir del acelerómetro del smartphone	21
3.1 Introducción	21
3.2 Medidas realizadas	22
3.3 Diseño de la aplicación móvil	24
3.4 Aprendizaje automático	25
3.4.1 Preprocesado	26
3.4.2 Generación y selección de características	26
3.4.3 Algoritmos de machine learning	28
3.4.4 Selección y evaluación del modelo	30
3.4.5 Métricas de evaluación	31
3.5 Resultados	32
3.5.1 Métricas de los diferentes modelos entrenados	32
3.5.2 Evaluación del modelo	33
3.6 Discusión	35
3.7 Conclusiones	36
4 Evaluación de la VFC con smartphone	37
4.1 Introducción	37
4.2 Aplicación móvil para extraer la señal RGB: evaluación de la frecuencia de muestreo	39

4.2.1	Diseño de la aplicación móvil	39
4.2.2	Pruebas experimentales	40
4.2.2.1	Montaje de las pruebas	41
4.2.2.2	Procesado de señal	43
4.2.2.3	Resultados	45
4.3	Evaluación de la influencia de la postura en la medida del ritmo cardíaco obtenido en el smartphone	46
4.3.1	Medidas realizadas	46
4.3.2	Extracción de la serie PP basado en smartphone	47
4.3.3	Índices de variabilidad de frecuencia cardíaca	47
4.3.4	Análisis de los datos	48
4.3.5	Rendimiento de la aplicación móvil diseñada	49
4.3.6	Resultados	49
4.3.7	Discusión	50
4.3.8	Conclusiones	51
4.4	Evaluación del error en la medida del pulso cardíaco con la cámara del móvil y de la VFC con respecto al ECG y PPG: influencias del modelo de smartphone y respiración	51
4.4.1	Introducción	51
4.4.2	Datos recogidos	52
4.4.3	Diseño de la carcasa	53
4.4.4	Procesado de la señal ECG, PPG y SPPG	53
4.4.5	Evaluación del error y análisis estadístico de los datos	55
4.4.6	Evaluación de diferentes fuentes de errores	56
4.4.7	Resultados	57
4.4.7.1	Comparación de las series PP y SPP con respecto al RR: influencia de los puntos fiduciales	57
4.4.7.2	Evaluación de las diferentes fuentes de errores	57
4.4.7.3	Desviación de estándar del error entre RR-SPP, RR-PP y PP-SPP: influencia del smartphone y de la respiración	58
4.4.7.4	Acuerdo entre SPP-RR, PP-RR y SPP-PP de los índices de VFC e influencia del modelo del smartphone	59
4.4.8	Discusión	60
4.4.9	Conclusiones	62
5	Análisis conjunto de la actividad física y ritmo cardíaco mediante smartphones	63
5.1	Introducción	63
5.2	Integración de ambos sistemas en una única aplicación	64
5.2.1	Cuestionarios psicológicos	64
5.2.2	Ritmo cardíaco	66
5.2.3	Actividad física	66
5.2.4	Interacción social	66
5.3	Implementación de la aplicación	67
5.3.1	Arquitectura del sistema	67
5.3.2	Índices utilizados	67
5.4	Medidas realizadas	68

5.5	Análisis de los datos	69
5.6	Resultados	70
5.7	Conclusiones	73
6	Detección temprana de somnolencia en conductores	75
6.1	Motivación	75
6.2	Métodos	76
6.2.1	Participantes	76
6.2.2	Sistema de medida	76
6.2.3	Protocolo de ensayo	78
6.2.4	Observadores externos	78
6.2.5	Región de tiempo de interés (TROI)	79
6.3	Algoritmo propuesto	81
6.3.1	Filtrado de la señal respiratoria	83
6.3.2	Caracterización de la variabilidad del ritmo respiratorio del sujeto	84
6.3.3	Evaluación de la variabilidad del ritmo respiratorio ciclo a ciclo e indicador de somnolencia	86
6.3.4	Algoritmo de evaluación de calidad de señal y combinación con el algoritmo de detección de somnolencia	89
6.4	Evaluación del rendimiento del algoritmo	90
6.4.1	Métricas de rendimiento	90
6.4.2	Optimización y validación del algoritmo	91
6.5	Resultados y discusión	92
6.5.1	Algoritmo de evaluación de calidad de señal: optimización y eva- luación del rendimiento	92
6.5.2	Algoritmo TEDD: resultados del algoritmo de clasificación	94
6.5.2.1	Ajuste de parámetros del algoritmo TEDD	94
6.5.2.2	Evaluación de los resultados del algoritmo TEDD	95
6.6	Conclusiones	97
7	Conclusiones	99
	Bibliografía	103

Índice de figuras

3.1	Capturas de pantalla de la aplicación diseñada	25
3.2	Diagrama de bloques de procesamiento del sistema de reconocimiento de actividad mediante algoritmos de aprendizaje automático	25
4.1	Diagrama de bloques del procesamiento fotograma a fotograma	40
4.2	Esquemático del circuito para emular la señal de PPG	41
4.3	Señal pulsátil captada por la cámara del smartphone ubicado sobre el diodo	42
4.4	Diagrama de conexiones del montaje de las pruebas con el circuito diseñado	42
4.5	Vista de las cámaras traseras de los smartphones utilizados	43
4.6	Gráfico de las distintas señales RR y del error, en este caso el móvil utilizado es el Samsung S5 y señal de entrada tiene frecuencia constante de 1 Hz	45
4.7	Diagrama de bloques del procesamiento de la componente verde de la señal RGB para obtener el pulso cardíaco	47
4.8	Diagrama de bloques del análisis de las series temporales RR y PP	48
4.9	Ejemplo del cronograma del protocolo de medida con diferentes modelos de smartphone y frecuencias respiratorias (FR). Para cada espacio de tiempo, se muestra la frecuencia respiratoria y el modelo de smartphone utilizado	52
4.10	Modelo 3D de las carcasas para cada uno de los smartphones	53
4.11	Ejemplo de la señal PPG con diferentes puntos fiduciales (FP): FP1 es el inicio de la señal PPG, FP2 es el máximo de la primera derivada de la señal y FP3 es el máximo de la segunda derivada	54
4.12	Ejemplo de las tres señales adquiridas: ECG, PPG y SPPG y los intervalos de tiempo obtenidos latido a latido: RR, PP y SPP	55
4.13	Diagrama de bloques de los métodos de acondicionamiento de señal anteriores a la detección del pulso. El filtrado paso banda y el método de interpolación son los detallados en la sección 4.4.4. El diezmado filtra paso bajo la entrada para evitar el aliasing y submuestra el resultado.	57
4.14	Sesgo y límites de acuerdo del Bland-Altman de los índices VFC. La primera columna muestra los resultados para SPP _{S5} -RR, la segunda para SPP _{MX} -RR y la tercera para PP-RR	61
5.1	Diseño final de las diferentes partes de la aplicación	65
5.2	Diagrama de la arquitectura del sistema	67
6.1	Setup de medida para detectar somnolencia del conductor en las instalaciones de IBV	77

6.2	Cronograma de los eventos registrados sobre la medida y la región de tiempo de interés (TROI) definida	79
6.3	Bloques principales del algoritmo propuesto de detección de somnolencia a partir del esfuerzo torácico (TEDD)	81
6.4	Diagrama de flujo detallado del algoritmo TEDD	83
6.5	Ejemplo de una señal respiratoria	84
6.6	Comparación de los índices $IndTedd$ y $qpTedd$ que se han obtenido para un sujeto	87
6.7	Un ejemplo de los resultados del algoritmo TEDD combinado con la evaluación de la calidad para un sujeto	88
6.8	Algoritmo de LOSOCV anidada para selección y evaluación del modelo	92
6.9	Resultados de optimización del algoritmo de calidad de señal	93
6.10	Resultados de clasificación de somnolencia del TEDD variando los valores $ThTedd$ y WLD para encontrar los mejores parámetros	94
6.11	Métricas de rendimiento de clasificación de somnolencia del algoritmo TEDD para $TeddDis$ y $TeddDisQ$. Se ha seguido el criterio de maximización de F1 para encontrar el mejor $ThTedd$. El número sobre el símbolo indica el valor de AT	96

Índice de tablas

1.1	Aplicaciones mHealth: fortalezas y debilidades	3
3.1	Datos de los sujetos medidos	23
3.2	Datos recogidos de acelerometría y cuestionarios	24
3.3	Características generadas	27
3.4	Algoritmos de clasificación e hiperparámetros ajustados	28
3.5	Ejemplo de matriz de confusión para la actividad A	31
3.6	Resultados de clasificación de actividades en validación	32
3.7	Resultados de evaluación del modelo de clasificación por sujetos	33
3.8	Resultados de evaluación del modelo de clasificación por MET asociados	34
3.9	Correlación de las variables con los resultados de clasificación	35
4.1	Error en la serie PP respecto a la Polar	44
4.2	Medida del error de la serie PP respecto a la Polar	49
4.3	Error relativo de la media de RR y los índices de VFC entre la banda Polar y los smartphones expresados en media \pm SD	50
4.4	Frecuencias de corte usadas en el procesamiento de la señal PPG y SPPG	54
4.5	Resultados para los distintos puntos fiduciales	58
4.6	SDE de las series utilizadas para evaluar las fuentes de errores	58
4.7	SDE de RR-SPP, RR-PP y PP-SPP por respiración y modelos de smartphone	59
4.8	Sesgo y límites de acuerdo Bland-Altman de los índices VFC entre SPP, RR y PP para cada modelo de smartphone	60
5.1	Resultados del índice de actividad	71
5.2	Resultados del índice de bienestar	71
5.3	Resultados del Perfil de Estado de Ánimo	72
5.4	Regresión del índice de actividad y el PEA	72
5.5	Resultados del cuestionario de satisfacción	73
6.1	Cantidad de datos disponibles para ser analizados según los criterios de inclusión	80
6.2	Métricas de rendimiento de clasificación utilizadas para el ajuste de parámetros y la evaluación del modelo	90
6.3	Métricas de rendimiento del algoritmo de clasificación de calidad de señal	93
6.4	Resultados de detección de somnolencia para diferente tiempo de avance (AT). Algoritmo TEDD (TeddDis) y TEDD combinado con evaluación de calidad de señal (TeddDisQ)	95

Abreviaturas

AC	Corriente Alterna
AF	Actividad Física
ANOVA	Análisis de varianzas
API	Interfaz de Programación de aplicaciones
AT	Tiempo de avance
AUC	Área Bajo la Curva
AWG	Generador de Forma de onda Arbitraria
BA	Bland-Altman
COPD	Enfermedad Pulmonar Obstructiva Crónica
CPU	Unidad Central de Procesamiento
CV	Validación Cruzada
DAC	Convertidor Digital Analógico
DC	Corriente Continua
DDS	Síntesis Digital Directa
DROI	Región de Interés de Somnolencia
ECG	Electrocardiograma
EEG	Electroencefalograma
EMA	Evaluación Ecológica Momentánea
ESD	Excesiva Somnolencia Diurna
ESP	Especificidad
ETA	Efecto Térmico de los Alimentos
E	Exactitud
F1	F1-score
FA	Fibrilación Auricular
FC	Frecuencia Cardíaca
FN	Falso Negativo
FP	Falso Positivo
G	G-media
GCT	Gasto Calórico Total
GPS	Sistema de Posicionamiento Global
GPU	Unidad de procesamiento gráfico
IBV	Instituto de Biomecánica de Valencia

IC	Intervalo de confianza
ICC	Coefficientes de la Correlación intraclase
IMC	Índice de Masa Corporal
K	Kappa de Cohen
kNN	k-vecinos más cercanos
LOSOCV	Validación cruzada dejando un sujeto fuera
LA	Límites de Acuerdo
LED	Luz emitida por diodo
MA	Macro-promediada
MI	Micro-promediada
MB	Metabolismo Basal
MET	Tasa Metabólica Equivalente
MX	Motorola Moto MX
OMS	Organización Mundial de la Salud
OSA	Apnea Obstrusiva del Sueño
PAC	Contracciones Auriculares Prematuras
PAT	Tiempo de llegada del pulso
PEA	Perfil de Estado de Ánimo
PP	Pulso a Pulso
PPG	Fotopletismograma
PSG	Polisomnografía
PVC	Contracciones Ventriculares Prematuras
QPD	Detector de cuasipico
RGB	Rojo Verde Azul
RIP	Pletismografía respiratoria inductiva
RL	Respiración Libre
ROC	Característica Operativa del Receptor
rpm	Respiraciones por minuto
RRV	Variabilidad del Ritmo Respiratorio
RSAF	Registro Semanal de Actividad Física
SDE	Desviación Estándar del Error
SEN	Sensibilidad
SPP	Smartphone Pulso a Pulso
SPPG	Smartphone Fotopletismografía
SVM	Máquinas vectores de soporte
TEDD	Detección de Somnolencia derivado del Esfuerzo Torácico
TROI	Región de Tiempo de Interés
S5	Samsung Galaxy S5
SD	Desviación estándar
SNA	Sistema Nervioso Autónomo
SNP	Sistema Nervioso Parasimpático

SNS	Sistema Nervioso Simpático
VFC	Variabilidad de la Frecuencia Cardíaca
VN	Verdadero Negativo
VP	Verdadero Positivo
YUV	Luminancia y Crominancia

Capítulo 1

Introducción

1.1 Motivación

Desde hace algunas décadas se está produciendo un envejecimiento de la población mundial como consecuencia del aumento de la esperanza de vida y la disminución de la tasa de fecundidad. Este fenómeno abre nuevos retos en el campo de la promoción de la salud. Según las predicciones realizadas por la Organización Mundial de la Salud (OMS), durante el período 2000-2050 la proporción mundial de personas con más de 60 años de edad se duplicará desde el 11 % hasta el 22 %, en números absolutos crecerá desde los 605 millones hasta alcanzar los 2.000 millones de personas [1]. Este hecho ha planteado el reto de mejorar la calidad de vida de las personas de edad avanzada y de retrasar la aparición de discapacidades propias de la edad, a este reto se le denomina como *Envejecimiento Saludable*. Por otra parte, según la definición de salud establecida por la OMS, una persona sana no es sólo aquella que no padece ninguna enfermedad, sino aquélla que además goza de un completo bienestar, tanto físico como psicológico y social. Por ello, la adherencia a un estilo de vida saludable resulta especialmente beneficiosa para la salud, porque no sólo contribuye a la prevención de enfermedades sino que aumenta el bienestar del individuo.

La OMS define la adherencia como “el grado en el que la conducta de una persona, en relación con la toma de medicación, el seguimiento de una dieta o la modificación de hábitos de vida se corresponde con las recomendaciones acordadas con el profesional sanitario” [2]. La adherencia a un estilo de vida saludable tiene manifestaciones y características diferentes según las edades. Además de proporcionar beneficios a las personas de edad avanzada, también tiene beneficios a corto plazo en cualquier etapa de la vida. Por un lado, la etapa infantil es una etapa de desarrollo físico y cognitivo de la persona, donde es de gran importancia cuidar las condiciones que potencian este desarrollo y minimizan el riesgo de que pueda entorpecerse, puesto que las alteraciones en el desarrollo tienen consecuencias negativas en el resto de las etapas de la vida. Además, en esta etapa

es cuando se modelan las conductas de comportamiento y actitudes [3, 4]. Por tanto, educar en hábitos duraderos que beneficien su bienestar y calidad de vida resulta una eficaz medida preventiva [5, 6].

En la etapa adulta, el trabajo desempeñado tiene un impacto sobre la salud, debido tanto a los riesgos asociados a las actividades que se llevan a cabo en el trabajo, como al ambiente físico y psicosocial del mismo. Además, en esta etapa, el individuo goza de autonomía sobre el estilo de vida que desea llevar. Por tanto, es importante mantener los hábitos adquiridos previamente y reorientar aquellos que puedan provocar mayor riesgo para la salud. Por último, en la etapa de la vejez se trata de fomentar aquellos hábitos que ayudan a disminuir el deterioro de las facultades físicas y cognitivas y por otra parte dotar al individuo de mecanismos de detección de anomalías en la salud y de alarma ante situaciones emergencia o cambios en la salud. Por lo tanto, la evaluación de la adherencia a un estilo de vida saludable debe estar orientada a atender las necesidades del individuo según su edad.

En estudios previos se afirma que llevar un estilo de vida no saludable durante un tiempo prolongado está involucrado en el inicio y en el desarrollo de enfermedades en las que se incluyen la diabetes, la hipertensión y la depresión [7–9]. En la literatura podemos encontrar estudios que indican que estas enfermedades pueden ser prevenidas mediante el mantenimiento de una dieta balanceada, práctica de actividad física moderada y abstenerse de fumar [10–12].

Las enfermedades relacionadas con el estilo de vida no son solo uno de los mayores factores que obstruyen una vida prolongada saludable, sino que también suponen un gran gasto al sistema sanitario. Uno de los factores de riesgo más importante está asociado a una dieta inadecuada y se estima que en España tiene un coste atribuible a casi 10.000 millones de euros. La baja actividad física es el segundo factor de riesgo atribuible por su importancia económica con un coste de más de 5.000 millones de euros, seguido del tabaco (1.572 millones) y el alcohol (523 millones) [13].

En concreto, la actividad física es uno de los pilares de la salud puesto que previene la obesidad [14], enfermedades tales como diabetes tipo 2 [15], cáncer [16, 17], cardiopatía coronaria [18] y enfermedades mentales como la depresión y ansiedad [19, 20]. Además, la actividad física está considerada como un comportamiento modificable.

Hoy en día, los smartphones que encontramos en el mercado llevan incorporados sensores tales como acelerómetro, barómetro y GPS. Además ofrecen la posibilidad de conectarse a otros sensores externos y su capacidad de almacenamiento de información es elevada. A esto habría que añadir la alta tasa de penetración de smartphones en los países desarrollados y que los usuarios que disponen de un smartphone suelen llevarlo encima todos los días y prácticamente durante todo el día. Esta combinación de elementos, permite utilizar los smartphones para recoger información acerca del usuario fuera de un laboratorio en ambientes representativos de la vida diaria. En el ámbito de la psicología,

Tabla 1.1: Aplicaciones mHealth: fortalezas y debilidades

Fortalezas	Debilidades
Smartphones: 334.3 millones de ventas (2018 Q1)	32500 aplicaciones mHealth (Nov 2017)
Evaluación Ecológica Momentánea	¿Situación de evaluación espontánea?
Sistemas no invasivos	¿Calidad de las aplicaciones? ¿Evidencias científicas?
Fácil de llevar siempre	Eficiencia dependiendo del modelo
Sensores propios: GPS, Bluetooth, acelerómetro. . .	Dificultad de entenderlos
Grabaciones sincronizadas: datos cognitivos, conductuales y fisiológicos	Complejidad de la aplicación. Demasiados clicks
Grabaciones automatizadas y en segundo plano	Baterías pobres, baja capacidad de procesamiento (dependiente del modelo)
Puede facilitar comportamientos relacionados con el estilo de vida saludables: motivación inicial	Rápida fatiga y aburrimiento. ¿Adherencia?

el concepto de evaluación ecológica momentánea (EMA) se refiere a las estrategias de monitorización para evaluar fenómenos en el momento que ocurren en condiciones naturales. El término ecológica significa que los datos capturados ocurren en el ambiente natural del sujeto. El término momentáneo quiere decir que las evaluaciones se focalizan en comportamiento actuales en lugar de concentrarse en el recuerdo o en el resumen sobre un largo período de tiempo. Por evaluación se entiende que múltiples evaluaciones se recogen a lo largo del tiempo para proporcionar un perfil sobre el comportamiento a través del tiempo. En nuestro caso, mediante el registro de datos en tiempo y entornos reales y naturalistas, se puede recopilar información sobre el comportamiento de la AF [21].

El uso de EMA en AF se ha incrementado debido a que puede proporcionar nuevos conocimientos sobre los determinantes del comportamiento en AF [22, 23]. Además, los teléfonos móviles son fáciles de usar y, por lo tanto, tienen la capacidad de recopilar datos rápidamente de un gran número de personas y transferir esta información a servidores de una manera discreta [24]. Por ello, los smartphones en combinación con comunicaciones basadas en Internet y algoritmos de *machine learning* permiten la evaluación de distintos comportamientos relacionados con la salud y permiten aportar información sobre su salud a los usuarios. Los sistemas basados en plataformas móviles que ofrecen soluciones en el ámbito de la salud se engloban bajo el término de *mHealth* [25]. En la tabla 1.1 podemos encontrar una serie de fortalezas y debilidades encontradas de las aplicaciones *mHealth*.

En la literatura podemos encontrar trabajos que evidencian que las aplicaciones móviles

pueden ayudar a los pacientes a alcanzar objetivos de salud y bienestar [26, 27]. Las tecnologías de *mHealth* están siendo cada vez más empleadas para evaluar actividad física en investigación clínica, epidemiológica e intervención [28]. Aunque existen muchas aplicaciones relacionadas con la actividad física, relativamente pocas han sido testadas en estudios de investigación para determinar su efectividad en la promoción de la salud [29].

La calidad de vida de una persona comporta tanto el bienestar fisiológico como el psicológico. De hecho, la salud mental es un componente integral y esencial de la salud del individuo que afecta al estado de ánimo. El estrés es un factor de riesgo de desórdenes mentales además de otras enfermedades no transmisibles [30]. El estrés psicológico genera a menudo una respuesta en el sistema nervioso simpático, de hecho reduce la variabilidad de la frecuencia cardíaca (VFC) [31]. Por un lado, algunos estudios encontraron que un descenso en la VFC está asociado al estrés [32, 33]. Además, un nivel bajo de VFC también está asociado a la depresión [34] y a la ansiedad [35]. Por otro lado, una baja VFC también está correlada con un aumento de riesgo cardíaco tales como la cardiopatía coronaria e infarto [36]. La VFC está regulada por múltiples mecanismos del organismo. La VFC está considerada como un marcador de bienestar psicológico y físico y como un predictor de morbilidad y mortalidad [34, 37]. En resumen, una VFC más baja acompaña a muchas enfermedades y una VFC más elevada acompaña a estados saludables, resiliencia y funcionamiento óptimo [38].

A partir del análisis del tiempo entre latidos cardíacos obtenido mediante el electrocardiograma, se puede obtener la variabilidad del ritmo cardíaco. Alternativamente a partir de la señal de la onda de pulso, se puede obtener el tiempo entre pulsos cardíacos y así evaluar la VFC.

Por otra parte, llevar un estilo de vida no saludable puede influir y favorecer el desajuste de los ritmos circadianos y favorece la somnolencia diurna puesto que este estilo de vida afecta negativamente a la calidad del sueño [39]. La falta de sueño o una mala calidad del sueño pueden provocar somnolencia cuando se conduce un vehículo, lo que supone un aumento del riesgo de accidente. De ahí el interés en desarrollar sistemas que permitan avisar al conductor cuando se está durmiendo al volante.

1.2 Objetivos

El objetivo de esta tesis es demostrar la viabilidad de la evaluación de la adherencia a hábitos de vida saludables mediante aplicaciones móviles y sensores ya sean propietarios del propio smartphone o conectados externamente. Este objetivo se puede desglosar en los siguientes subobjetivos:

- Diseñar un sistema para la evaluación del gasto calórico asociado a la actividad física a partir de los datos obtenidos del móvil en condiciones naturales
- Obtener los indicadores de VFC a través de la señal de fotopletimografía registrada utilizando la cámara del smartphone
- Integrar los sistemas en una única aplicación móvil para permitir el análisis combinado de la actividad física, VFC y estado de ánimo.
- Diseñar y validar un algoritmo para la detección de somnolencia al volante a partir del análisis de la señal respiratoria

Por un lado en esta tesis se pretende contribuir a la estimación del gasto calórico realizado durante la actividad física y el ejercicio físico mediante la información captada por un smartphone en condiciones naturales, utilizando principalmente el sensor de acelerometría y utilizando modelos basados en aprendizaje automático para el reconocimiento de las distintas actividades físicas.

Además, se pretende obtener indicadores de VFC a través de la señal de fotopletimografía capturada utilizando la cámara del smartphone. También se pretende analizar la validez de los indicadores obtenidos en distintas posturas, patrones respiratorios y modelos de smartphone.

Otro de los objetivos es integrar en una única aplicación la recogida de datos de acelerometría y la evaluación de la VFC mediante la medida del pulso. La finalidad de esta integración es poder realizar un análisis combinado de ambas variables utilizando un único dispositivo, en este caso el smartphone. El análisis de los datos obtenidos con esta aplicación permitirá, no solo evaluar las variables de gasto calórico y la información aportada por los indicadores de VFC, sino también la adherencia a estilos de vida saludables.

Por otro lado, un objetivo de la tesis es proponer y validar un algoritmo para la detección de episodios de lucha por no dormirse mientras se conduce a partir del análisis de la señal respiratoria. La señal respiratoria será adquirida mediante un sensor externo basado en una banda inductiva que mide el esfuerzo torácico.

1.3 Esquema de la tesis

El primer capítulo de la tesis está dedicado a hacer una introducción, centrar la tesis en el contexto de la evaluación de la adherencia a hábitos de vida saludable y fijar los objetivos de la misma. El Capítulo 2 contiene una revisión del estado del arte de la estimación del gasto calórico a partir del reconocimiento de la actividad física realizada y de la estimación de la variabilidad del ritmo cardíaco utilizando la cámara del smartphone.

En el Capítulo 3 se detalla el método propuesto para la evaluación del gasto calórico basado en el reconocimiento de la actividad física realizada a través de la información obtenida del acelerómetro del smartphone. Se describen los algoritmos de *machine learning* utilizados, las medidas realizadas a partir del cuestionario conocido como Registro Semanal de Actividad Física (RSAF) donde se anotan las conductas de actividad física en un entorno ecológico y momentáneo, con una resolución temporal de 15 minutos. Los resultados se clasifican tanto nivel de actividades como de nivel de gasto calórico asociado.

En el Capítulo 4 se describe el sistema para la evaluación de la VFC basado en fotopleletismografía utilizando la cámara trasera del smartphone. Se describe la aplicación desarrollada, un primer conjunto de ensayos de validación del método de medida en diferentes posturas y un segundo conjunto de ensayos con 23 sujetos en los que se evaluó el grado de acuerdo del sistema propuesto basado en una aplicación móvil con respecto al electrocardiograma y un dispositivo de fotopleletismografía de referencia.

En el Capítulo 5 se propone un sistema también basado en una aplicación móvil para la evaluación conjunta de la actividad física, el ritmo cardíaco y el estado de ánimo. La actividad física se estima a partir de los datos del acelerómetro, el ritmo cardíaco a partir de la cámara trasera del smartphone y el estado de ánimo se registran en pantalla mediante cuestionarios psicológicos implementados en la propia aplicación desarrollada. Se describen las pruebas y los resultados obtenidos con un grupo de 11 sujetos.

En el Capítulo 6 se explica cómo se ha desarrollado un sistema para la detección de episodios de lucha por no dormirse al volante. Se explica el algoritmo propuesto que está basado en el análisis de la variabilidad del ritmo respiratorio a partir de la señal del esfuerzo torácico medido con una banda inductiva. También se detallan las medidas realizadas para validar el algoritmo propuesto y los resultados obtenidos.

Por último, en el Capítulo 7 se resumen las conclusiones a las que se ha llegado y se proponen algunas líneas de trabajo futuro.

Capítulo 2

Estado del arte

2.1 Actividad física: acelerometría y smartphone

En este apartado mostraremos los conceptos relacionados con el gasto calórico y los distintos métodos de medida y por otro lado los sistemas basados en acelerometría para medir el gasto calórico asociado a la actividad física.

2.1.1 Métodos de medida del gasto calórico

El gasto calórico total (GCT) diario de una persona se puede calcular a partir del metabolismo basal (MB), el gasto por actividad física (AF) y el efecto térmico de los alimentos (ETA). Los principales métodos de medida son la calorimetría indirecta, la calorimetría directa, el agua doblemente marcada, la monitorización del ritmo cardíaco, los autoinformes del propio sujeto y los sensores de movimiento [40].

El metabolismo basal es el gasto calórico diario que necesita el propio organismo para llevar a cabo las funciones vitales del cuerpo. Representa entre el 60 % o 70 % del gasto energético diario para la mayoría de personas sedentarias y cerca del 50 % para aquellas que son activas. La AF es el gasto producido al realizar las tareas diarias que implican movimiento incluido el ejercicio físico. El ejercicio físico es aquel movimiento estructurado, planeado y repetitivo realizado con la intención de mejorar o mantener el estado de forma. El gasto calórico asociado a la AF es el componente más variable y puede variar día a día [41]. En concreto, en individuos activos el gasto calórico asociado a la actividad física puede corresponder a una o dos veces el basal mientras que en personas sedentarias puede representar menos de la mitad del basal [42].

El efecto térmico de los alimentos es el gasto calórico que realiza el cuerpo para la digestión, absorción, uso y almacenamiento de los nutrientes después de ingerir los alimentos.

Cuando la persona está en balance energético, el gasto es igual al consumo, se estima que el ETA es un 10 % del GCT diario [43].

El método que se conoce como agua doblemente marcada está considerado como el método de referencia o gold estándar para evaluar el gasto calórico total. Este método permite estimar de forma precisa el GCT en condiciones libres y en un período de 7 a 14 días pero no permite distinguir específicamente el tipo de actividad física, duración, frecuencia o intensidad. Está basado en la ingestión de isótopos estables, deuterium (^2H) y oxígeno-18 (^{18}O) mediante una bebida de agua y la monitorización de la eliminación de estos isótopos por parte del organismo. La diferencia entre los ratios de eliminación de ^2H y de ^{18}O es equivalente a la producción de dióxido de carbono, que puede convertirse en gasto calórico total. El hecho de que el individuo haya de ingerir isótopos radioactivos y el posterior análisis y la recogida de orina hace que este método sea costoso y pueda ser rechazado por la personas.

La calorimetría directa mide el calor producido por una persona encerrada en un habitáculo cerrado en condiciones de control. Aunque es una medida que puede considerarse muy precisa como medida del GCT, tiene el inconveniente de que requiere el confinamiento de la persona en una cámara, esto resulta costoso y poco viable para la vida diaria. La calorimetría indirecta permite medir el consumo de oxígeno a través de la respiración mediante una máscara y un analizador de gases. Es un método preciso y que permite la estimación del gasto calórico de la AF, pero es un sistema costoso, obstrusivo y la instrumentación puede alterar los patrones de AF, por lo que resulta poco viable para evaluar el gasto calórico en condiciones naturales.

Otro método de estimación del gasto calórico es a partir de la monitorización del ritmo cardíaco. Este método está basado en la relación lineal entre el ritmo cardíaco y el consumo de oxígeno [44]. Existen diferencias en esta relación para los distintos sujetos. Este método sirve para medir la intensidad del ejercicio físico y el gasto asociado al ejercicio físico pero no a la actividad física ligera [45–47]. Se puede medir con un sencillo sensor que consiste en una banda que se coloca en el pecho y que transmite la actividad cardíaca. Es ampliamente utilizado en la práctica deportiva pero el ritmo cardíaco también está relacionado con un amplio rango de factores no relacionados con la AF; por ello no sirve para medir el gasto calórico de las actividades diarias.

El gasto calórico asociado a la actividad física también se puede estimar utilizando acelerómetros, que se basan en medir la aceleración de alguna parte del cuerpo [48]. La acelerometría permite estimar la intensidad y duración del movimiento. Una de las principales ventajas es la practicidad de su uso. La desventaja es que los estudios indican que necesita calibrarse para cada usuario.

Otro método de medida es la estimación del gasto calórico mediante autoinformes del propio sujeto. En estos cuestionarios se deben detallar las actividades realizadas durante el día. Las actividades están codificadas según el tipo y el Compendio de Actividades

Físicas ofrece una codificación de las mismas junto con el gasto calórico asociado expresado en tasa metabólica equivalente (MET) [49]. 1 MET es el gasto en media de una persona tumbada, es equivalente a 1 kcal/kg/hora. Cabe destacar que aunque esta conversión puede no ser exacta para todas las personas sí que es ampliamente utilizada para distinguir entre intensidades de actividad física: de 1.0 a 2.9 actividad ligera, de 3.0 a 5.9 moderada y mayor o iguales que 6 se considera vigorosa [50]. El principal inconveniente de este método es que es retrospectivo y requiere del recuerdo de la persona y el rellenar cuestionarios puede ser viable para una estimación para un período corto pero puede resultar engorroso para el largo plazo.

Aunque el gasto calórico asociado a la actividad no es el mayor componente del GCT, cabe destacar que es en el que más variaciones pueden producirse día a día y tiene una dependencia conductual que puede ser modificable, y también afecta al metabolismo basal en tanto que una mayor AF puede ayudar a bajar el peso y hacer descender el MB. Sin embargo, el MB debe ser medido mediante sofisticadas técnicas basadas en calorimetría o estimado a partir de ecuaciones como la de Harris-Benedic que cuenta tanto con elementos modificables a medio y largo plazo como el peso y otros no modificables como la edad, altura o sexo [51]. Por eso, esta tesis está centrada en la evaluación del gasto calórico asociado a la actividad física a través de acelerómetros. A continuación, haremos una revisión de los métodos existentes basados en acelerometría.

Además, se pretende poder evaluar el cumplimiento de las recomendaciones de la OMS en cuanto a la realización de actividad y ejercicio físico para la promoción de una vida activa y saludable. Estas recomendaciones están basadas en la duración y el tipo de actividad realizada.

2.1.2 Medida del gasto calórico con los sensores del smartphone

Hoy en día, la mayoría de smartphones llevan incorporado un acelerómetro. Por ello, han aparecido numerosos trabajos acerca de la estimación del gasto calórico y detección de actividades usando los propios smartphones de los usuarios.

En una gran cantidad de trabajos el conjunto de actividades estimadas es muy reducido y han sido medidas en un entorno de laboratorio. Además, las tasas de detección de actividad estaban calculadas para una ubicación concreta del smartphone: en la cadera [52], en el bolsillo del pantalón [53, 54], en el pecho [55] o en la mano [56]. Además en muchos trabajos el tiempo de medida es corto siendo como mucho de 30 min en condiciones naturales [57, 58].

Por otro lado, podemos encontrar varios trabajos de detección de actividad con sensores de acelerometría externos ubicados en distintas partes del cuerpo [59–62]. Los datos procedentes de cada sensor de acelerometría se procesan para clasificar los distintos tipos de actividad utilizando algoritmos de *machine learning*. Los resultados de la precisión en

términos de detección de actividad según la ubicación del sensor varían según el tipo de actividad, por ello no concluyen que exista una ubicación óptima del sensor para todas las actividades.

También podemos encontrar trabajos basados en el análisis del número de cuentas. Una cuenta es la señal de aceleración rectificada $|x|$ integrada sobre un cierta cantidad de tiempo. El inconveniente es que al usar simplemente las cuentas se pierde información del patrón de movimiento que permite obtener otras características. Además, cuando el resultado del acelerómetro se expresa en cuentas, aparece el problema que la relación entre cuentas y gasto energético es distinta dependiendo del tipo de actividades que se lleva a cabo. Por ejemplo, en el caso de ir en bicicleta y andar. Cuando el acelerómetro se coloca en la cadera o en el pecho, las cuentas aumentarán cuando aumente la intensidad del ejercicio y de ahí aumentará el gasto calórico. Cuando la persona está andando, las cuentas aumentarán más con el aumento de la intensidad que con el ciclismo. El reconocimiento del tipo de actividad puede solucionar este problema, asignando distinto gasto calórico para la intensidad de cada tipo de actividad [63].

Existe un trabajo en el que se compara el reconocimiento de intensidades de actividad de smartphones que usan Android con el acelerómetro de ActiGraph. Este acelerómetro es uno de los más usados en este tipo de investigación. Esta comparación está realizada a partir de la conversión en cuentas de los datos procedentes del smartphone. Encontraron que existe correlación entre los datos del ActiGraph y del Android, que están fuertemente correlados entre ellos en los distintos smartphones. Además, que el hecho de llevar el ActiGraph en la cadera y el Android en el bolsillo no tiene un impacto negativo en la detección de intensidades de actividad [64].

Podemos encontrar algunos trabajos que utilizan el acelerómetro del smartphone para reconocimiento de actividad física. Se obtienen unos resultados muy buenos con tasas de acierto entorno al 90%. Sin embargo cabe destacar que sólo están validados para un conjunto pequeño de actividades diarias: subir y bajar escaleras, caminar, sentado, de pie y tumbado [57, 65]. En otro trabajo se analizan otras actividades más complejas como escribir o comer, aunque para ello utilizan dos sensores, uno ubicado en la muñeca y otro en el bolsillo. Los datos de las actividades están recogidas durante sólo 5 minutos en un entorno controlado [66].

Entre los distintos métodos de medida del gasto calórico, la acelerometría es el método menos intrusivo de todos y el que parece más adecuado para evaluar el gasto calórico en condiciones naturales. Algunos trabajos utilizan métodos basados en *machine learning* para identificar distintos tipos de actividades físicas [53, 59]. En esta tesis se pretenden evaluar los métodos combinados que son unos métodos más avanzados que los algoritmos básicos tipo máquinas vectores de soporte, árboles de decisión, etc. Podemos encontrar trabajos basados en *machine learning* pero una limitación es que utilizan base de datos con sólo actividades básicas [16, 67].

En entornos naturales no se debe asumir que la persona va a llevar el smartphone en una posición fija y determinada. Al contrario, la persona puede llevar el móvil en distintas partes del cuerpo: cadera, bolsillo y en la mano o puede llevarlo en un bolso o mochila. En algunos trabajos que hablan de condiciones naturales se refieren a fuera del laboratorio pero el smartphone va siempre ubicado en el bolsillo [58].

En esta tesis se pretende identificar los distintos tipos de actividad y utilizarlo para realizar una estimación del gasto calórico. De esta manera, se puede realizar las medidas de evaluación del gasto calórico utilizando la acelerometría. La identificación de actividades y la estimación del gasto calórico asociado se pretenden realizar a partir de medidas en condiciones naturales, independientemente de donde lleve la persona puesto el smartphone. Se recogerán datos durante una semana, el periodo propuesto al utilizar el Registro Semanal de Actividad Física (RSAF) y que parece suficiente para recoger datos de los días típicos y el fin de semana.

2.2 Variabilidad de la frecuencia cardíaca

Un corazón saludable no se comporta como un metrónomo, de ahí el interés en evaluar la variabilidad de la frecuencia cardíaca. El término VFC, se refiere al fenómeno de los cambios en los intervalos de tiempo entre latidos cardíacos consecutivos [68]. La VFC está generada por las interacciones entre el sistema nervioso autónomo (SNA) y el corazón según la dinámica de los subsistemas simpático y parasimpático. Las oscilaciones de un corazón sano son complejas. En un estado de reposo predomina el sistema parasimpático (SNP) y en uno de estrés, ansiedad o ejercicio físico el sistema simpático (SNS). El SNP gestiona los cambios reflejos de la frecuencia cardíaca (FC) debidos a señales procedentes de los barorreceptores arteriales y del sistema respiratorio. El SNS es el responsable de los cambios en la FC debidos a estrés físico y mental [69].

La *Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology* establece diferentes períodos de tiempo para evaluar la VFC: 24h, 5 minutos o inferior a 5 minutos [37]. La evaluación del VFC se puede realizar usando medidas en el dominio del tiempo, dominio de la frecuencia y no lineales [70].

La VFC se determina mediante el procesado de la señal del electrocardiograma (ECG), los picos R del complejo QRS de la señal ECG se utilizan para obtener la duración de un ciclo cardíaco. Por ello, la VFC se obtiene a partir del análisis de la serie temporal de intervalos RR a partir del ECG. La manera convencional de obtener el ritmo cardíaco es utilizar el ECG para capturar la actividad eléctrica del corazón mediante electrodos pegados a la superficie de la piel. La señal que se obtiene es precisa y limpia. Sin embargo, para medidas ambulatorias adquirir la señal del ECG implica utilizar electrodos y conectarlos a un dispositivo específico tipo Holter. Aunque es un sistema bastante preciso, es un sistema sofisticado y su uso requiere de especialistas y puede resultar complejo para el

público en general. Además, durante el ejercicio físico se utilizan aparatos inalámbricos como la banda Polar para medir la VFC de forma válida y fiable mediante el registro de los intervalos RR [71].

La señal que se conoce como fotopleletismograma (PPG), se captura mediante una técnica óptica simple y de bajo coste que se utiliza para detectar cambios en el volumen sanguíneo. Se utiliza a menudo para realizar medidas en la superficie de la piel de forma no invasiva. La forma de onda de PPG tiene una componente pulsátil que está atribuida a los cambios sincrónicos del corazón en el volumen sanguíneo a cada latido cardíaco y tiene superpuesta una línea base con variaciones lentas con varios componentes de baja frecuencia atribuidos a la respiración, actividad del sistema nervioso autónomo y la termoregulación [72]. Esta técnica óptica consiste en que se emite una luz que es enviada al tejido y su reflexión es recogida por un fotoreceptor. El bajo coste y la simplicidad de esta tecnología pueden ofrecer beneficios significativos para la atención sanitaria por ejemplo en la atención primaria donde son deseables métodos no invasivos, precisos y cuya técnica de diagnóstico sea simple de usar [73].

Con la finalidad de evitar el uso de electrodos se puede utilizar la señal de la onda de pulso y a partir de la detección del pulso cardíaco se obtiene la serie pulso a pulso que se utiliza para evaluar la VFC. Este sistema es fiable, más barato y simple, aunque también requiere usar un dispositivo externo para la recogida de datos.

Alternativamente, existen una serie de trabajos en los que se evalúa el pulso cardíaco a partir de la medida de cambios sutiles de coloración en la piel de la cara. Estos cambios son registrados a través de grabaciones de vídeo de la cara de la persona utilizando la cámara frontal del smartphone [74–78]. En la literatura podemos encontrar estudios en los que estos dos métodos basados en fotopleletismografía son comparados con un sensor que adquiere ECG y con un pulsioxímetro [79–81]. Estos estudios muestran que se obtienen peores estimaciones del ritmo cardíaco a partir de los cambios en la piel usando la cámara frontal que a partir de los cambios en el dedo usando la cámara trasera.

En esta tesis es de interés el uso de la cámara trasera del smartphone para realizar medidas de pulso cardíaco. En la literatura podemos encontrar algunos trabajos en los que se utiliza un smartphone para evaluar el pulso cardíaco. En ellos, la yema del dedo se coloca sobre la cámara, se utiliza el flash de luz emitida por diodo (LED) como emisor y la cámara como fotodetector. Algunos trabajos previos se limitan a analizar el ritmo cardíaco medio y no la VFC [82]. Algunos que sí miden la VFC, no utilizan el smartphone como sensor si no que sólo utilizan el smartphone para conectarle un dispositivo que mide el pulso [83].

Podemos encontrar algún trabajo que evalúa la variabilidad del ritmo cardíaco pero a partir del análisis de grabaciones de vídeo. Este método no permite realizar un procesado

en tiempo real, porque requiere de la grabación de toda la medida y después procesarla [84].

Sobre los parámetros de VFC sólo vamos a tratar aquellos que están más relacionados con un estilo de vida saludable y que se utilizan en medidas de 5 minutos: SDNN, RMSSD y LF/HF. El SDNN es la desviación estándar de los intervalos entre latidos. RMSSD es la raíz cuadrada de la media de la suma de las diferencias al cuadrado de los intervalos RR sucesivos. Este parámetro nos informa de las variaciones a corto plazo de los intervalos RR y puede ser utilizado para observar la influencia del SNS sobre el sistema cardiovascular. La medida del espectro de frecuencias de la VFC se obtiene a partir de una transformación matemática, habitualmente la Transformada de Fourier, que permite descomponer la energía de la señal RR en diferentes componentes espectrales. La proporción LF/HF es la relación entre baja frecuencia (LF) situada entre 0.04 Hz y 0.15 Hz y la alta frecuencia (HF) situada entre 0.15 Hz y 0.4 Hz. De esta proporción podemos estimar la influencia vagal relacionada con la relajación y la HF y, la influencia del SNS que está relacionada con el estrés y el LF. Según algunos estudios si predomina la influencia del SNS de manera permanente puede ser causa de trastornos de salud, depresiones o sobreentrenamiento y puede afectar al equilibrio biofísico de la persona. Debido a la controversia en la interpretación del LF de forma aislada, se utiliza la proporción de LF/HF para estimar de manera más efectiva la actividad del SNS [69].

En resumen, el sistema nervioso autónomo juega un papel importante en las funciones fisiológicas y en la patogénesis de muchos trastornos médicos. La medida de la VFC proporciona un método no invasivo fácilmente aplicable de evaluación de las funciones autónomas [85]. Una alta VFC caracteriza a una persona sana con mecanismos autónomos eficientes y una buena capacidad adaptación del SNA [86, 87]. En la mayoría de estudios que evaluaron la asociación de la VFC con el estrés, mostraron que está asociado con una baja VFC [32, 88].

En el mercado de las aplicaciones móviles podemos encontrar numerosas aplicaciones para medir el ritmo cardíaco, concretamente en 2013 se podían encontrar un total de 94 para los sistemas operativos de Android y iOS [89]. La mayoría de ellas no están respaldadas por ningún estudio de validación y los algoritmos utilizados no son públicos. Podemos encontrar la aplicación Cardio para iPhone (<https://www.cardiio.com>), hay algunos trabajos que la respaldan aunque solamente está validada la estimación del nivel de ritmo cardíaco medio [90, 91]. La aplicación FibriCheck (https://play.google.com/store/apps/details?id=com.qompium.fibricheck&hl=en_US) que está disponible para Android también está validada pero sólo a nivel de ritmo medio [92]. Además, existe un estudio en el que se valida la aplicación Instant Heart Rate de Azumio (<https://www.azumio.com/s/instantheartrate/index.html>) tanto para Android como para iPhone con respecto a un monitor de ritmo cardíaco a partir del ECG tipo Polar FT7. En este estudio se obtiene una alta correlación del ritmo medio entre la banda Polar y la aplicación para iPhone y Android. Sin embargo, no se hace ningún tipo de

evaluación de los índices de VFC [93]. La aplicación Heart Rate Plus también ha sido analizada comparándola con un pulsioxímetro. En este estudio han obtenido un alto índice de correlación y una repetibilidad elevada del ritmo medio pero tampoco han analizado la VFC [94].

En un trabajo de metaanálisis recientemente publicado [95], encontraron una serie de estudios [90, 94, 96–101] que evalúan la PPG del smartphone con respecto al ECG o un dispositivo de referencia de PPG utilizando el método más apropiado para comparar el acuerdo entre dos métodos, el método Bland-Altman. Todos estos trabajos evalúan el grado de acuerdo del ritmo medio, pero sólo en dos de ellos se analiza el grado de acuerdo de la VFC [96, 98]. Acerca del estudio de Bolkhovsky et al. [96], se analizan el RMSSD, SDNN y LF/HF en un dispositivo iPhone y otro Android, sin mostrar los resultados de los gráficos del método Bland-Altman y los resultados no parecen muy fiables porque las unidades de los valores mostrados de los índices no concuerdan con el rango normal de valores. En otro estudio se utiliza un iPhone, se analizan el SDNN y RMSSD, aunque se ha hecho el análisis Bland-Altman y se muestra el gráfico del SDNN, en los resultados sólo se muestran los datos correlación y no el grado de acuerdo [98].

Adicionalmente, en un trabajo más reciente se comparan smartphone, PPG y ECG a partir del ritmo medio y el RMSSD, obteniendo una correlación casi perfecta del ritmo medio y un grado de acuerdo aceptable para el RMSSD [102]. En la literatura podemos encontrar un trabajo en el que se analiza la variabilidad del pulso cardíaco utilizando un smartphone de tipo iPhone que permite adquirir a 240 Hz [103]. En este análisis sí estudian el grado de acuerdo y obtienen buenos resultados. En nuestro caso nos interesa utilizar dispositivos tipo Android cuya frecuencia de muestreo es menor, alrededor de 30 Hz, y evaluar si esto supone una limitación importante. En resumen, existe una ausencia de validación robusta de la adquisición de los sistemas basados en PPG con smartphones [104, 105].

En esta tesis lo que nos interesa es diseñar un sistema de adquisición de la señal producida por el pulso arterial periférico en el dedo mediante fotopleletismografía utilizando como sensor la cámara trasera de smartphones con el sistema operativo Android. Además también nos interesa evaluar el grado de acuerdo de la estimación de la VFC entre el smartphone y el ECG. También evaluar si este acuerdo es similar al obtenido con dispositivos de PPG comerciales y evaluar los resultados en distintas condiciones de respiración y distintas marcas comerciales de smartphones.

2.3 Somnolencia durante la conducción

La somnolencia es un estado intermedio entre la vigilia y el sueño, que puede ser definido como la progresiva pérdida de la eficiencia de procesado cortical. Está también asociado a un deseo o inclinación a dormir [106]. La somnolencia durante la conducción

puede ser causada por una combinación de falta de sueño, conducir cuando los ritmos circadianos están bajos (primera hora de la mañana o media tarde) o conducir durante largos períodos de tiempo. La somnolencia afecta a elementos del rendimiento humano que son críticos para la seguridad al volante tales como: tiempo de reacción, alerta y procesamiento de información [107].

La somnolencia al volante es un serio problema de salud pública y preponderante que merece atención. La *AAA Foundation for Traffic Safety* en su informe sobre somnolencia durante la conducción publicado en 2017, afirma que el porcentaje estimado de conductores que admiten haberse dormido al volante (en los 30 días previos) es de alrededor de un 30%. En el mismo informe, podemos encontrar que casi todos los conductores (95.9%) consideran que es algo inaceptable conducir cuando estás tan dormido que tienes problemas para mantener los ojos abiertos. Sin embargo, uno de cada cinco (19.6%) admitió haber conducido así más de una vez y un 2.9% dijo haberlo hecho a menudo o regularmente [108].

Los daños debidos a accidentes de tráfico aumentarán hasta convertirse en séptima causa de muerte en el mundo en 2030 de acuerdo con las predicciones de la OMS [109]. Estudios recientes estiman que alrededor del 20% de los accidentes de coche han sido causados por conductores somnolientos [110]. Además, se han investigado los efectos de la privación de sueño en la tasa de accidentes. Se encontró que los conductores que generalmente duermen menos de 5 horas diariamente y el conductor que ha dormido durante menos de una hora o más que su cantidad habitual de sueño en las 24 horas anteriores, tiene una tasa significativamente elevada de accidente. La tasa de accidentes estimada es similar al riesgo asociado con conducir con una concentración de alcohol en sangre igual o ligeramente superior al límite legal [111].

Sentirse dormido es un proceso gradual y los conductores no son conscientes de su falta de atención durante este tiempo. Por eso, un sistema de evaluación en tiempo real de la somnolencia del conductor para avisar al conductor cuando aparecen los primeros síntomas de fatiga puede evitar accidentes mediante la prevención de episodios de sueño tanto en conductores profesionales como particulares.

Hoy en día, uno de los principales objetivos en el desarrollo de nuevos sistemas de asistencia al conductor avanzada es la detección de somnolencia de manera fiable. Los métodos de detección de somnolencia más extendidos se pueden dividir en tres categorías, según la información que analizan: el comportamiento del conductor, las características faciales y las características fisiológicas. Los métodos basados en el análisis del comportamiento del conductor analizan información como la posición del coche dentro del carril, la velocidad, uso del volante, frenos y cambio de marchas [112]. El principal inconveniente de este método es la variación en su exactitud según las características del coche y del estilo de conducción.

Por otro lado, otros grupos se han focalizado en métodos de visión por computador para analizar la cara del conductor. Estos sistemas procesan la información del estado del ojo calculando el porcentaje de cierre del ojo [113], duración del cierre del ojo y la frecuencia de cierre del ojo [114, 115]. Otros sistemas combinan esta información con los movimientos de la cabeza, bostezos y expresión facial [116]. El principal inconveniente de estos sistemas es recoger la información del estado del ojo adecuadamente cuando el conductor lleva gafas de sol oscuras. Esto ocurre con frecuencia durante la conducción. Aunque, este es un sistema de monitorización no invasivo, hay algunas dudas sobre su comercialización puesto que a la gente generalmente le disgusta ser grabada por una cámara enfocando su cuerpo continuamente durante un largo período de tiempo.

Por otro lado, en cuanto a la monitorización del sueño, la polisomnografía (PSG) es el sistema de referencia para el diagnóstico de trastornos de sueño. Este método está basado en el análisis y procesado de múltiples señales biomédicas, tales como electroencefalograma (EEG), electrocardiograma (ECG), respuesta galvánica de la piel y la respiración. PSG es el método más utilizado para el propósito clínico. El rendimiento de los métodos basados en EEG ha sido evaluado por múltiples autores [117–119], pero este método sufre de serias limitaciones relacionadas con la ergonomía del sensor de EEG: requiere sensores y cables en el cuerpo que perturban al conductor. Por un lado, el EEG es una señal de baja amplitud y el coche es electromagnéticamente ruidoso lo que puede dificultar la detección de señal. Por otro lado, las vibraciones mecánicas del coche pueden introducir artefactos en la interfaz electrodo-piel lo que también dificulta la detección de señal.

En cuanto al análisis de las características fisiológicas, la variabilidad de la frecuencia cardíaca (VFC) ha sido generalmente empleada como indicador de fatiga para atletas [120]. Además, en PSG, la VFC extraída ha sido usada para determinar las fases de sueño [121]. Sin embargo, uno de los principales inconvenientes de los métodos basados en VFC es que no parece que sea un indicador viable de somnolencia [122]. Finalmente, el análisis de la variabilidad del ritmo respiratorio (RRV) puede ser útil para recoger información precisa de los ciclos de somnolencia así que los métodos basados en respiración pueden anticipar situaciones de riesgo durante la conducción [123, 124].

En esta tesis se propone un método de detección de somnolencia basado en cambios en la señal respiratoria. La señal respiratoria se ha obtenido usando una banda de pletismografía inductiva y que se procesa en tiempo real para clasificar el estado de alerta del conductor como somnoliento o despierto. Además se validan los resultados obtenidos por el método de detección somnolencia propuesto cuando aparecen los primeros síntomas de somnolencia.

2.4 Sensores en los móviles

Los smartphones que podemos encontrar hoy en día en el mercado están equipados con una gran cantidad de sensores tales como: acelerómetro, giroscopio, brújula, sensor de luz, sensor de proximidad y barómetro. También vienen equipados con otros elementos que pueden ser usados como sensores: micrófono, cámara trasera y frontal y GPS [125]. Muchos de estos sensores están incorporados en el smartphone para apoyar a la interfaz de usuario y enriquecerla como por ejemplo el acelerómetro o para aumentar los servicios basados en localización como por ejemplo la brújula. Los sensores de proximidad y de luz permiten al smartphone llevar a cabo un reconocimiento contextual simple asociado con la interfaz de usuario. El sensor de proximidad detecta cuando el usuario sostiene el smartphone junto a la cara para hablar. En ese caso, para ahorrar batería se apaga la pantalla. El sensor de luz se utiliza para ajustar el brillo de la pantalla. En el ámbito de la localización, el GPS permite localizar el smartphone, que se utiliza en aplicaciones de navegación, búsquedas localizadas y redes sociales. La brújula y el giroscopio representan una extensión de la localización, proporcionando un mayor conocimiento de su posición con respecto al mundo físico (dirección y orientación) realizando las aplicaciones de localización [126].

Estos sensores son útiles no sólo para la interfaz de usuario o aplicaciones de localización sino que los datos recogidos por estos sensores pueden ser utilizados para aplicaciones relacionadas con la salud [127, 128].

Podemos agrupar el acelerómetro, brújula, giroscopio y barómetro como sensores de movimiento. Estos sensores se utilizan en el ámbito de la salud para aplicaciones tales como la monitorización de actividad y la detección de caídas. La señal de aceleración adquirida con los sensores está asociada a la aceleración del propio cuerpo. Cuando se está en posición tumbada, el valor de la aceleración eje vertical del smartphone es cercano a cero, mientras que en posición de pie y sentado está en un valor máximo. Las posiciones sentado y de pie se distinguen analizando las variaciones de la señal de aceleración del eje vertical. Estas son altas cuando la persona está de pie debido al movimiento continuo del cuerpo para mantener la estabilidad y son más bajas cuando la persona permanece sentada. La actividad de caminar se detecta mediante el filtrado paso banda de la señal de aceleración entre el rango de frecuencias que generalmente corresponde al del caminar humano [129].

También se pueden utilizar los datos de estos sensores de movimiento para ser analizados mediante algoritmos de *machine learning* para distinguir un mayor número de actividades. Los datos del barómetro se pueden utilizar para detectar actividades con desplazamientos verticales, tales como subir o bajar escaleras [130, 131].

Los datos del acelerómetro también se pueden utilizar para detectar caídas [132]. Durante la caída, se produce un rápido decrecimiento de la aceleración hasta un mínimo cuando

la persona cae, el impacto con el suelo causa un rápido aumento de la aceleración hasta un valor máximo. Las aceleraciones de las actividades diarias no alcanzan estos niveles y por tanto mediante unos umbrales adecuados se pueden detectar las caídas. En un estudio con 18 personas se obtuvieron unos resultados de detección de caída de una especificidad de 81 % y una sensibilidad del 77 % [133]. Estos detectores de caída son esenciales para proporcionar una rápida asistencia [134] y para prevenir el miedo a caerse especialmente en gente mayor [135].

Además, a partir de los datos del GPS se pueden analizar las distancias y rutas recorridas, esta información se utiliza en aplicaciones relacionadas con la salud mental como la detección de la depresión [136]. A partir de modelos de clasificación no lineales entrenados con características extraídas de sensores del smartphone como Wifi, acelerómetro y GPS, se ha demostrado una prueba de concepto para la detección de la depresión mejor que la clasificación aleatoria en un grupo de 36 sujetos [137]. En otro estudio con 28 sujetos, se analizan los datos extraídos del GPS y datos del uso del smartphone. Para ello, se utiliza un clasificador para distinguir los participantes que presentan síntomas de depresión de los que no los presentan y los resultados obtenidos muestran una exactitud del 86.5 % [138].

Por otro lado, podemos encontrar un prototipo de un sistema basado en smartphones para detectar una conducta deambulatoria en personas con demencia. Este sistema utiliza los datos del GPS y compara la estimación de la ubicación actual con ubicaciones conocidas para determinar si el paciente está deambulando. Este sistema fue evaluado con 11 pacientes que sufren demencia durante 84 semanas y se encontró que era fiable y robusto [139].

Los smartphones actuales llevan incorporadas cámaras de alta resolución y flashes LED. Podemos encontrar en el mercado numerosas aplicaciones que utilizan la cámara y su flash para detectar el ritmo cardíaco. Están basadas en la técnica de fotopletismografía y su uso está bastante popularizado. Una vez se obtiene la señal de pulso cardíaco, esta puede ser utilizada para diversas finalidades como: la detección de fibrilación auricular (FA), evaluar la variabilidad del ritmo cardíaco [103] o únicamente obtener el ritmo medio.

La FA es una irregularidad de ritmo cardíaco peligrosa que puede aumentar el riesgo de sufrir complicaciones relacionadas con el corazón. La aplicación “PULSE-SMART” analiza el pulso cardíaco adquirido con la cámara del móvil para detectar la FA y que pueda discriminarlo del ritmo sinusal, contracciones auriculares prematuras (PACs) y contracciones ventriculares prematuras (PVCs). Esta aplicación ha sido testeada con 121 personas y se ha obtenido una exactitud del 95.1 % de identificación de pulso irregular durante AF y la discriminación del PAC ha sido del 95.5 % y para el PVC del 96.0 % [140].

Podemos encontrar algunos estudios acerca de la utilización del micrófono para la evaluación de aspectos relacionados con la salud. Por un lado, la aplicación llamada “Stress-Sense” detecta el estrés en diversos escenarios utilizando el micrófono del smartphone. Realizaron un estudio con 14 personas en un periodo de 4 días e incluyeron habla bajo estrés (entrevista de trabajo y empleado en tareas de marketing), así como habla neutral (leyendo cuentos). Mediante el uso como referencia de un sensor de respuesta galvánica de la piel para recoger los datos de temperatura de la piel, los investigadores demostraron que el clasificador era capaz de detectar estrés con una exactitud del 81 % y el 76 % para ambientes interiores y exteriores respectivamente [141].

Por otro lado, la aplicación “SpiroSmart” utiliza el micrófono para inferir la tasa de flujo respiratorio durante un test de espirometría. La espirometría es la medida objetiva de la función pulmonar más empleada y es importante para diagnosticar y gestionar enfermedades pulmonares crónicas como el asma o la enfermedad pulmonar obstructiva crónica (COPD) [142]. La espirometría mide el flujo de aire y sirve para detectar la capacidad del sujeto para exhalar aire de los pulmones. Por lo tanto, permite detectar si las vías respiratorias están obstruidas y la predisposición para desarrollar enfermedades pulmonares obstructivas crónicas. La aplicación ha sido evaluada con 52 sujetos y muestra un error medio de un 5.1 % con respecto a un espirómetro clínico. Aunque no sirve para sustituir los equipos clínicos, sí puede ser utilizada para realizar medidas a nivel doméstico [143].

Capítulo 3

Estimación de la actividad física realizada a partir del acelerómetro del smartphone

3.1 Introducción

Como hemos comentado anteriormente numerosos estudios han demostrado que la práctica de actividad física de forma habitual genera beneficios en la salud. La práctica de la actividad física está considerada cómo uno de los pilares de la vida saludable. Por el contrario, un estilo de vida sedentario está asociado a distintas enfermedades cardiovasculares, diabetes y enfermedades psicológicas como la depresión [144]. La insuficiente práctica de actividad física es uno de los mayores factores de riesgo de mortalidad global y va en aumento en muchos países. Las personas que son insuficientemente activas tienen entre un 20 % y un 30 % de aumento de riesgo de muerte comparados con aquellas que son suficientemente activas. Globalmente, alrededor del 23 % de las personas mayores de 18 años no eran lo suficiente activas en 2010 [145]. Por eso existe un interés en el diseño de sistemas que permitan evaluar el nivel de actividad física de la persona [146].

En la literatura podemos encontrar trabajos acerca del reconocimiento de actividad física a partir de sensores de acelerometría. En los últimos años a medida que se ha ido extendiendo el uso de smartphone, han ido apareciendo trabajos que utilizan el sensor de acelerometría del smartphone, de forma que se puede evaluar la actividad física de manera menos intrusiva. La zona de ubicación del móvil es otro aspecto tratado en distintos trabajos, los dispositivos o smartphone se han ubicado en distintas partes del cuerpo como: cadera, bolsillo y brazos. Sin embargo, en un escenario más natural y menos intrusivo, la ubicación del smartphone debe ser indiferente porque no todas las personas llevan el smartphone en la misma ubicación.

Existen trabajos previos de reconocimiento de actividad física a partir de la señal de acelerometría de un smartphone pero están enfocados a un conjunto pequeño de actividades. En ocasiones, estas actividades se reducen a correr, caminar, caminar más rápido de 4 km/h, subir y bajar escaleras. Por otro lado, la mayoría de trabajos se han realizado en entornos de laboratorio y controlados.

Por ello, en esta tesis se ha propuesto un sistema de reconocimiento de actividad recogiendo los datos de acelerometría del smartphone en condiciones naturales durante el día a día de la persona y en el cual, el smartphone puede ser llevado donde la persona desee. El reconocimiento de actividades está basado en algoritmos de aprendizaje automático. Los datos se recogen de forma no intrusiva mediante una aplicación móvil diseñada para este fin sin que sea necesario llevar puesto ningún sensor adicional.

El RSAF es un instrumento de evaluación conductual adaptado a las actividades físicas cotidianas de una población adulta en países desarrollados. El RSAF es un cuestionario basado en papel, que propone un compendio de actividades reducido que comprende las actividades cotidianas, adaptadas a las actividades de nuestra sociedad basado en compendios ya existentes [147]. Además, como la finalidad es evaluar el gasto calórico se ha utilizado el RSAF como modelo del conjunto de actividades a evaluar. Los datos se recogen con una periodicidad de 15 minutos como establece el RSAF. La estimación del gasto calórico se realiza a partir de la traducción de las actividades detectadas en MET. La relación entre actividades y MET está establecida en el RSAF.

3.2 Medidas realizadas

Se ha construido una base de datos a partir de los datos de acelerometría de la actividad física realizada por personas recogidos por la aplicación móvil desarrollada. Además, los sujetos indicaron la actividad física realizada en cada momento a través de la aplicación móvil seleccionando una actividad dentro de un conjunto de actividades recogidas en el RSAF. El RSAF está diseñado para registrar la actividad física de forma semanal. Por ello, las pruebas realizadas por los voluntarios tienen la duración de una semana completa.

Se recogieron datos de 26 sujetos sanos, 10 mujeres y 16 hombres. En la tabla 3.1 podemos ver un resumen de las variables de edad, peso, altura e IMC de los sujetos. Los modelos de los smartphones utilizados fueron variados porque la aplicación se instaló en el propio smartphone de cada sujeto.

En cuanto a la cantidad de datos recogidos, la tabla 3.2 muestra un resumen de los mismos. Los datos se han promediado entre todos los sujetos. Se muestran los datos en unidades de fracciones de 15 minutos porque ésta es la periodicidad del RSAF, además,

Tabla 3.1: Datos de los sujetos medidos

	Media \pm SD
Edad (años)	23.42 \pm 7.15
Altura (cm)	173.96 \pm 9.54
Peso (kg)	69.50 \pm 11.55
IMC (kg/m ²)	22.89 \pm 2.81

se muestra el porcentaje para tener una idea de la cantidad de datos durante una semana. La primera fila muestra la cantidad de cuestionarios de actividad física que fueron recogidos, estos datos pueden ser superiores a 7 días porque la aplicación no restringe las fechas de los datos introducidos. Además como el registro de datos empieza a la hora que se registra el usuario por primera vez, a los usuarios se les pedía que indicaran los datos de actividad durante 8 días completos para asegurar que se recogían datos de acelerometría y cuestionarios durante 7 días.

Por otra parte, la segunda fila nos indica la cantidad de datos de acelerometría captados por la aplicación. Los valores encontrados nos indican que algunos sujetos tuvieron apagado el smartphone durante algún tiempo dentro de la semana. En la tercera fila tenemos la cantidad de datos del acelerómetro adquiridos que tienen un dato de actividad física asignado.

De los datos recogidos, se han descartado los datos de aquellas actividades de las que había 6 o menos ocurrencias durante todo el registro. Este límite representa aproximadamente un 1.5 % sobre la media de actividades totales de un sujeto con aceleración. Se ha escogido este límite porque consideramos que estas actividades están lo suficientemente infrarepresentadas para ser tenidas en cuenta. Finalmente, debido a la gran cantidad de datos de la actividad de dormir, se han reducido los datos de esta actividad hasta que tengan la misma cantidad de datos que la actividad inmediatamente inferior en cantidad de datos.

El número de actividades diferentes que realiza cada persona puede variar para cada una de ellas según la actividad realizada. En concreto, los datos después de haber limitado la actividad de dormir muestran que la mediana de las diferentes actividades registradas por los sujetos es 9 con rango intercuartilíco 8–10.

En un principio nos habíamos planteado utilizar los datos proporcionados por el sensor de presión atmosférica, puesto que a partir de las variaciones de la presión atmosférica se pueden analizar los cambios de altura durante la realización de actividad física. Sin embargo, hay que tener en cuenta que al contrario que los acelerómetros no todos los móviles llevan incorporado el barómetro. De los datos recogidos, solamente los smartphones de 3 sujetos de un total de 26 tenían esta información, por tanto se descartó utilizar este tipo de datos en este estudio.

Tabla 3.2: Datos recogidos de acelerometría y cuestionarios

	Fracciones de 15 min.		Porcentaje sobre una semana	
	Media	SD	Media	SD
Cuestionarios actividad	689.0	182.9	102.5	27.2
Datos acelerómetro	517.6	179.5	77.0	26.7
Datos acelerómetro y cuestionarios	424.7	154.9	63.2	23.1
Datos después de eliminar menor que 7	415.3	152.3	61.8	22.7
Datos después de limitar dormir	364.4	132.0	54.2	19.6

3.3 Diseño de la aplicación móvil

Se ha diseñado e implementado una aplicación móvil que sirva como herramienta para obtener una base de datos de acelerometría con los datos asociados de actividad física realizada. Además todos los datos recogidos deben ser enviados a un servidor para poder ser procesados.

Por un lado la aplicación adquiere los datos del sensor de acelerometría y presión atmosférica del propio smartphone. Esta recogida de datos se ha programado para ser realizada durante 7 días, que es el plazo que establece el RSAF. Los datos de acelerometría de los ejes x, y, z están adquiridos a 50 Hz, que es una frecuencia suficiente para el reconocimiento de actividad puesto que para el análisis de la actividad diaria es deseable frecuencias mayores de 20 Hz [148]. En los 26 sujetos analizados encontramos que el margen dinámico del sensor de acelerometría es diferente para los distintos smartphones. En concreto, tenemos 6 a $\pm 2g$, 10 a $\pm 4g$, 4 a $\pm 16g$ y de 6 no tenemos los datos disponibles. El rango de amplitud recomendado para registrar movimientos del cuerpo usando un acelerómetro a la altura de la cintura es de al menos $\pm 6g$ [148]. La utilización de acelerómetros con un rango menor al mencionado puede provocar que la señal de acelerometría se sature en las actividades de mayor intensidad. A medida que se van adquiriendo los datos, se guardan en un fichero para cada día y la relación de ficheros se guarda en una base de datos para poder ser enviados a un servidor.

Por otro lado, como podemos ver en la figura 3.1(b), se ha diseñado una interfaz de usuario (UI) que permite introducir las actividades físicas del RSAF para cada intervalo de tiempo. Además, en la figura 3.1(a) se observa que la aplicación permite consultar las actividades introducidas y editarlas para poder llevar un control de las mismas. Estos registros de actividades se guardan en una base de datos de forma estructurada para poder consultarlos, editarlos y luego poder enviarlos a un servidor. La figura 3.1(c) muestra que la aplicación también permite registrar las variables antropométricas del usuario. Se ha desarrollado un servicio web tipo REST y se ha instalado en un servidor



(a) Lista de actividades (b) Nueva actividad (c) Registro variables antropométricas

Figura 3.1: Capturas de pantalla de la aplicación diseñada



Figura 3.2: Diagrama de bloques de procesamiento del sistema de reconocimiento de actividad mediante algoritmos de aprendizaje automático

Apache Tomcat para así poder recibir los datos a través de la comunicación con la aplicación. En la aplicación se ha desarrollado el correspondiente cliente del servicio web para poder enviar los datos. Además, se ha desarrollado en la aplicación un módulo que detecte cuando hay una conexión a una red de Internet disponible y en caso de que tenga datos por enviar los envíe al servidor. De esta forma el envío de datos se realiza de manera automática sin necesidad de que el usuario intervenga.

3.4 Aprendizaje automático

El procesamiento de los datos se realiza utilizando algoritmos de machine learning. Para ello, se han seguido los pasos que indica la figura 3.2, se han generado las características, se han seleccionado las características más representativas, se han construido los modelos, se ha seleccionado el más preciso y finalmente se ha evaluado. Todo el procesamiento de los datos se ha realizado en MATLAB. Todo ello se ha realizado para cada sujeto de forma individual.

3.4.1 Preprocesado

En algunos casos hemos detectado que la señal de acelerometría no se adquiría correctamente porque el sensor daba valores erróneos para los distintos ejes y para el timestamp. Por ello, para el análisis posterior se han eliminado los tramos de señal en los que se ha detectado que no se adquiría la información correctamente. La señal es adquirida por la aplicación a una frecuencia no constante entorno a los 50 Hz, por ello se ha remuestreado la señal a una frecuencia constante de 50 Hz. La señal se ha filtrado paso alto con una frecuencia de corte de 0.25 Hz mediante un filtro Butterworth de orden 2 para eliminar la componente gravitatoria. Además, se ha filtrado paso bajo a una frecuencia de corte de 20 Hz con un filtro Butterworth de orden 2 para eliminar ruido, esta frecuencia es suficiente teniendo en cuenta el ancho de banda del movimiento del cuerpo humano.

3.4.2 Generación y selección de características

Los datos obtenidos son datos continuos de los tres ejes del acelerómetro, así que para generar las características hemos dividido los datos en fragmentos de un minuto de duración. Esta duración es apropiada para comportamientos de actividad física y ha sido utilizada en estudios anteriores [149]. Los datos fueron agregados en muestras solapadas de un 50 % usando una ventana deslizante.

La tabla 3.3 contiene las características generadas. Se han utilizado las características típicamente utilizadas en los sistemas de reconocimiento de actividad física a partir de los datos de acelerometría. Se han generado las características tanto en el dominio temporal como en el frecuencial. Estas características se han obtenido para los tres ejes de acelerometría x, y y z. Además, se ha generado una variable categórica llamada “parte del día” que divide el momento en que la actividad ha sido realizada en tres partes: 6 a. m. a 2 p. m. , 2 p. m. a 10 p. m. y de 10 p. m. a 6 a. m. Las siguientes características han sido analizadas en ventanas de 0.5 segundos para ver las transiciones de actividad: μT , μD , σT , σD , d_x y dA . Además, se ha calculado el tiempo entre picos y el tiempo entre cruces por cero para cada uno de los ejes para analizar la periodicidad de la señal.

Una vez se han obtenido las características, el siguiente paso es su selección. La selección de características es un método comúnmente empleado en machine learning porque permite seleccionar un subconjunto de las características extraídas relevante que serán usadas para construir el modelo. Por un lado se trata de disminuir la presencia de atributos redundantes que hacen más lento el aprendizaje y reducir el ruido que introducen y que pueden confundir a algunos clasificadores. Por otro lado, el hecho de tener una mayor dimensión puede hacer que sea necesario tener más datos para obtener buenos resultados.

Tabla 3.3: Características generadas

Característica	Ecuación
Media	$\mu = \frac{1}{N} \sum_{i=1}^N x_i$
Varianza	$\sigma^2 = \sum_{i=1}^N (x_i - \mu)^2$
Correlación entre ejes	$r_{xy} = \frac{\sum_{i=1}^N (x_i - \mu_x) * (y_i - \mu_y)}{\sqrt{\sum_{i=1}^N (x_i - \mu_x)^2 * \sum_{i=1}^N (y_i - \mu_y)^2}}$
Módulo del área de la señal	$SMA = \sum_{i=1}^N a_{x_i} + \sum_{i=1}^N a_{y_i} + \sum_{i=1}^N a_{z_i} $
Ángulo de inclinación	$\arccos(\mu_x)$
Media de la tendencia	$\mu T = \sum_{i=2}^{N_s} \mu_i - \mu_{i-1} $
Media de la diferencia	$\mu D = \sum_{i=1}^{N_s} \mu - \mu_i $
Varianza de la tendencia	$\sigma T = \sum_{i=2}^{N_s} \sigma_i - \sigma_{i-1} $
Varianza de la diferencia	$\sigma D = \sum_{i=1}^{N_s} \sigma - \sigma_i $
Diferencia máxima	$d_x = \max(x) - \min(x)$
Diferencia máxima total	$dA = \sqrt{d_x^2 + d_y^2 + d_z^2}$
Coefficientes de autoregresión de Burg	$a = \text{arburg}(x, 3)$
Energía espectral de la FFT	$E = \frac{1}{N} \sum_{k=1}^N FFT(x[k]) ^2$
Energía no correlada entre dos ejes	$EU_{xy} = 2 * (E_x + E_y) - r_{xy} * (E_x + E_y)$
Entropía en el dominio frecuencial	$H = \frac{1}{\log_2 N} \sum_{i=1}^N P_i * \log_2(P_i); P_i = FFT(x[i]) ^2$

El objetivo principal de la selección de características es elegir el mejor subconjunto de las características disponibles que alcanza menor error de clasificación. Hemos utilizado selección secuencial en conjunto creciente para establecer que características son importantes. Principalmente, este algoritmo sigue un procedimiento de abajo a arriba que empieza con el conjunto vacío. Secuencialmente, añade las características seleccionadas por la función objetivo que minimiza el error cuadrático medio hasta que no hay mejora en la predicción.

Tabla 3.4: Algoritmos de clasificación e hiperparámetros ajustados

Clasificador	Hiperparámetros optimizados
Árboles de decisión	Mínimo número de observaciones, criterio de poda, criterio de partición
Naive Bayes	Tipo de kernel y ancho de la ventana del kernel
k-vecinos más cercanos	Número de vecinos, función del cálculo de la distancia
Máquinas vectores de soporte	Box constraint y escala del kernel
Bagging	Mínimo tamaño de hoja y número de predictores para muestrear
AdaBoost	Número de árboles, máximo número de particiones y número de árboles
Rusboost	Número de árboles, máximo número de particiones y número de árboles
Random subspace	Número de vecinos, cálculo de la distancia, número de predictores para muestrear y número de aprendices

3.4.3 Algoritmos de machine learning

Hemos utilizado varios algoritmos que son a menudo implementados para reconocer actividades humanas en la literatura y están disponibles en el paquete de MATLAB de estadística y aprendizaje automático. Hemos utilizado los modelos de aprendizaje automático de tipo supervisado porque nuestros datos están etiquetados con la actividad realizada. La tabla 3.4 muestra los clasificadores utilizados y los hiperparámetros optimizados que son descritos a continuación.

Los árboles de decisión son un modelo jerárquico en los que los atributos están mapeados en ramas y nodos que representan los posibles valores de los atributos. Cada rama desde la raíz hasta un nodo hoja es una regla de decisión de clasificación. Se han usado los árboles de decisión que utilizan el algoritmo Clasificación and Regession Trees (CART) [150]. CART se caracteriza por seleccionar el predictor de partición que maximiza la ganancia del criterio de partición sobre todas las posibles particiones de todos los predictores. Las particiones se hacen recursivamente hasta que se alcanza un criterio de parada. Los parámetros que han sido ajustados son: el mínimo número de observaciones en un nodo hoja, criterio de poda y criterio de partición [151, 152]. Los árboles de decisión son fáciles de entender pero les falta precisión. Su rendimiento puede ser realizado introduciendo un elemento aleatorio en la construcción y en la agregación resultante de muchos árboles. Los bagged trees y boosted trees son métodos que implementan esto.

Naive Bayes es un clasificador probabilístico. Éste asume que los predictores son unos independientes de los otros dentro de cada clase. El método calcula la probabilidad

posterior de esta muestra perteneciendo a cada clase. Esto está basado en la estimación de la densidad de probabilidad del predictor para cada clase. Los parámetros que han sido ajustados son: el tipo de kernel y el ancho de la ventana del kernel.

El método k-vecinos más cercanos (kNN) clasifica ejemplos no etiquetados basados en su similaridad a ejemplos en el conjunto de entrenamiento. Para dado un ejemplo no etiquetado, encuentra las k muestras más cercanas etiquetadas y asigna la muestra no etiquetada a la clase que aparece más frecuentemente dentro del k-subconjunto. Este algoritmo requiere de un gran almacenamiento y tiene un coste computacional elevado en la fase de evaluación debido a que cada nueva instancia debe ser comparada con el conjunto de entrenamiento entero para ser clasificada. El número de vecinos y la función de cálculo de la distancia son los parámetros que han sido optimizados.

Máquinas vectores de soporte (SVM) está basado en el concepto de planos de decisión que definen las fronteras de decisión. SVM utiliza un mapeado no lineal mediante la transformación de los datos originales de entrada en una dimensión mayor. SVM busca el hiperplano que da la distancia mínima a los ejemplos de entrenamiento más larga. SVM utiliza un margen blando, que significa que el hiperplano separa a muchos pero no a todos los puntos de los datos. Para este problema de clasificación de múltiples clases, se ha utilizado el modelo de códigos de corrección de errores el cual reduce el modelo de clasificación con más de dos clases a un conjunto de clasificadores binarios. Se ha usado una función de kernel gaussiana y un diseño uno contra todos. Los parámetros box constraint y escala de kernel han sido optimizados.

Los métodos combinados de aprendizaje están basados en la combinación de clasificadores para mejorar la precisión de los modelos. Se han usado distintos modelos basados en diferente métodos combinados: bagging, boosting y random subspace.

Bagging está basado en bootstrap y agregación. Bootstrap significa muestreo aleatorio con reemplazamiento y agregación quiere decir combinar clasificadores. Cada clasificador base está entrenado con una réplica bootstrap del conjunto de entrenamiento, así que esto puede a veces conseguir menos objetos engañosos. Entonces, el clasificador construido puede tener un mejor rendimiento. Los clasificadores se combinan mediante votación. La combinación de clasificadores puede dar mejores resultados que los clasificadores individuales debido a que las ventajas de los clasificadores individuales son combinadas en la solución final. Puede resultar de utilidad para construir un clasificador mejor si hay elementos engañosos en el conjunto de entrenamiento y funciona adecuadamente con clasificadores inestables tales como árboles de decisión [153]. Han sido optimizadas el mínimo tamaño de la hoja y el número de predictores para muestrear parámetros del método.

En boosting, los clasificadores y el conjunto de entrenamiento se obtienen de manera determinista. Se obtienen secuencialmente en el algoritmo. Cada clasificador base está entrenado con los datos que está ponderado basado en el rendimiento del clasificador

previo. Cada clasificador vota para obtener la salida final. Se ha utilizado la implementación de Adaboost que es una de las más populares. El algoritmo empieza con los datos igualmente ponderados. AdaBoost pondera los datos de entrenamiento dándole mayor peso a aquellos que están incorrectamente clasificados. Se han ajustado tres parámetros: tasa de aprendizaje, máximo número de particiones y número de árboles. Se ha usado también la implementación RUSBoost que combina muestreo de datos y boosting, está diseñada para mejorar el rendimiento de clasificación cuando los datos no están balanceados [154]. Se han optimizado los mismos parámetros que para AdaBoost.

En random subspace cada clasificador está entrenado con todo el conjunto de entrenamiento pero con un subconjunto de predictores. Los predictores son seleccionados aleatoriamente. El clasificador base recomendado para este algoritmo es kNN que es el que se ha usado. La decisión final combinada es realizada por votación de mayoría. Los parámetros que han sido optimizados son: el número de vecinos, el cálculo de la distancia, número de predictores para muestrear y número de aprendices [155].

3.4.4 Selección y evaluación del modelo

Para realizar la selección y la evaluación del modelo de forma adecuada se han dividido los datos en tres partes. La asignación de los datos a cada una de las partes se ha realizado de forma aleatoria, los conjuntos de datos generados reciben los siguientes nombres: entrenamiento, validación y test. El conjunto de entrenamiento es utilizado para construir el modelo con los diferentes parámetros del clasificador, el modelo ajustado se utiliza para predecir las respuestas para las observaciones en el conjunto de validación. El error obtenido en el conjunto de validación se utiliza para seleccionar entre los distintos clasificadores entrenados. El modelo seleccionado es usado para predecir las respuestas para las observaciones en el conjunto de test. Por tanto el conjunto de test se utiliza sólo para evaluar el rendimiento del clasificador completamente entrenado. Los datos han sido divididos: 60 % para entrenamiento, 20 % para la selección del modelo y 20 % para la evaluación del modelo. Por lo tanto la primera parte (60 %) es usado para entrenar cada modelo y seleccionar los hiperparámetros de cada modelo. La segunda parte (20 %) es utilizada para seleccionar entre los diferentes tipos de modelos entrenados y el otro 20 % es utilizado para evaluar el modelo final.

La técnica de validación cruzada (CV) de k-iteraciones, comúnmente utilizada en este ámbito, ha sido utilizada para estimar los errores sin perder datos. En la CV de k-iteraciones, los datos de la muestra se dividen en k-subconjuntos. Uno de los subconjuntos se utiliza como test y el resto como datos de entrenamiento. Este proceso de CV se repite durante k iteraciones, con cada uno de los posibles subconjuntos de datos de test. Para obtener el resultado final se calcula la media aritmética de los resultados de cada iteración. Este método es preciso porque evaluamos sobre k combinaciones de datos entrenamiento y test, pero tiene la desventaja de que tiene un alto coste computacional

Tabla 3.5: Ejemplo de matriz de confusión para la actividad A

Observadas \ Clasificadas	Actividad A	Actividad B	Actividad C
Actividad A	VP_a	FP_a	FP_a
Actividad B	FN_a	–	–
Actividad C	FN_a	–	–

[156, 157]. En concreto, 4-iteraciones CV ha sido utilizada en la selección del modelos y 5-iteraciones CV en la evaluación del modelo para conseguir las mencionadas proporciones en la partición de los datos. En resumen, el método de selección y evaluación consiste en 2 bucles de validación cruzada anidados: un bucle interior para escoger el modelo de clasificación y los parámetros ajustados y un bucle exterior para evaluar el modelo. Por otro lado, la selección de características ha sido realizada en el bucle interior para evitar una subestimación del error de generalización [158].

Además, la regla de una desviación estándar ha sido utilizada para seleccionar los parámetros de manera que se seleccionen algoritmos menos complejos para los árboles de decisión y el número de aprendices en los algoritmos combinados de boosting y random subsapce [159].

3.4.5 Métricas de evaluación

Se han obtenido distintas métricas a partir de la matriz de confusión del algoritmo de clasificación de actividad. Un ejemplo de matriz de confusión para una actividad se muestra en la tabla 3.5. El rendimiento de clasificación ha sido evaluado con las siguientes métricas: exactitud, sensibilidad, precisión y f-score. Para una actividad, la precisión mide la proporción de muestras clasificadas de una actividad que han sido correctamente clasificadas. Se calcula como $P = \frac{VP}{VP+FP}$, donde VP es número de verdaderos positivos y FP el número de falsos positivos. La sensibilidad se mide como la proporción de muestras observadas de una actividad que han sido correctamente clasificadas. Se calcula como $S = \frac{VP}{VP+FN}$, donde FN son los falsos negativos. F-score es la media armónica de la precisión y sensibilidad, $f\text{-score} = \frac{2*P*S}{P+S}$. La exactitud se calcula de forma que engloba todas las actividades y es el cociente de la suma de todos los VP entre todos los elementos de la matriz de confusión.

En el caso de los problemas de clasificación multiclase, como es este caso donde se clasifican distintos tipos de actividades y cada actividad es una clase distinta, los datos se promedian para todas las clases. En este caso las clases pueden no estar balanceadas. Por eso, nosotros hemos promediado los resultados de las diferentes clases de dos maneras: macro y ponderada. Macro consiste en promediar dando el mismo peso a todas las clases y ponderada es darle un peso a cada clase según el número de elementos de cada clase.

Tabla 3.6: Resultados de clasificación de actividades en validación

Algoritmo	Exactitud	Macro F-score
Árboles de decisión	0.638	0.543
Naïve Bayes	0.485	0.37
Máquinas vectores de soporte	0.612	0.532
K-vecinos más cercanos	0.630	0.546
Bagging	0.757	0.702
AdaBoost	0.740	0.680
RUSBoost	0.656	0.627
Random Subspace	0.681	0.606

Durante cada ejecución de la k-iteración de CV, el rendimiento de cada clasificador entrenado fue evaluado en términos de sensibilidad, precisión y f-score totales. Las métricas fueron promediadas para todas las iteraciones para obtener la media y desviación estándar de cada métrica para todo el conjunto de datos.

Cada actividad está asociada a un valor de GC medido en MET. Por eso, se ha calculado el error relativo de clasificación asociado al GC en MET. Los MET asociados a cada actividad están establecidos en el RSAF [147].

Adicionalmente, se ha evaluado el modelo teniendo en cuenta el gasto calórico asociado a cada actividad. Para ello, a la hora de evaluar el modelo se han agrupado en una misma clase de la matriz de confusión aquellas actividades que tienen el mismo número de MET. Una vez, obtenida la nueva matriz de confusión también se han obtenido las correspondientes métricas.

3.5 Resultados

3.5.1 Métricas de los diferentes modelos entrenados

En un total de 26 sujetos por 5 iteraciones, el modelo seleccionado como más preciso ha sido Bagging 126 veces y AdaBoost ha sido seleccionado 4 veces. La tabla 3.6 muestra los resultados del rendimiento de clasificación de los diferentes modelos en el conjunto de validación. Los datos han sido macro-promediados para todas las actividades y promediados para todos los sujetos. La tabla 3.6 muestra que se obtienen mejores resultados en media usando métodos combinados que usando los clasificadores simples. Además Bagging es el mejor clasificador y los resultados son parecidos a los obtenidos con el algoritmo AdaBoost.

Tabla 3.7: Resultados de evaluación del modelo de clasificación por sujetos

	Macro Precisión	Sensibilidad	F-score	Exactitud	Ponderado F-score	eRelativoMET Media	std
Sujeto 1	0.717	0.662	0.688	0.718	0.719	-2.288	0.820
Sujeto 2	0.757	0.741	0.749	0.793	0.791	-1.544	0.985
Sujeto 3	0.762	0.683	0.720	0.760	0.760	-1.947	0.704
Sujeto 4	0.676	0.520	0.588	0.702	0.698	-4.123	0.824
Sujeto 5	0.664	0.571	0.614	0.729	0.721	-5.027	1.584
Sujeto 6	0.866	0.852	0.859	0.862	0.863	4.811	1.286
Sujeto 7	0.757	0.745	0.751	0.771	0.770	-0.861	1.160
Sujeto 8	0.756	0.697	0.725	0.766	0.766	0.294	0.436
Sujeto 9	0.768	0.725	0.746	0.781	0.780	-1.438	0.412
Sujeto 10	0.758	0.626	0.686	0.802	0.799	0.230	0.483
Sujeto 11	0.871	0.845	0.858	0.878	0.878	-0.122	0.064
Sujeto 12	0.792	0.730	0.760	0.795	0.795	-4.574	1.462
Sujeto 13	0.844	0.745	0.791	0.873	0.870	-0.591	0.452
Sujeto 14	0.765	0.713	0.738	0.764	0.763	-3.754	0.571
Sujeto 15	0.696	0.598	0.643	0.722	0.719	1.252	1.446
Sujeto 16	0.678	0.564	0.616	0.683	0.683	-14.255	0.535
Sujeto 17	0.901	0.902	0.901	0.902	0.902	0.822	1.732
Sujeto 18	0.688	0.584	0.632	0.751	0.744	-9.459	2.172
Sujeto 19	0.739	0.619	0.673	0.774	0.770	-4.723	1.195
Sujeto 20	0.675	0.588	0.628	0.726	0.720	-5.344	0.946
Sujeto 21	0.696	0.515	0.592	0.699	0.703	2.505	1.091
Sujeto 22	0.667	0.610	0.637	0.649	0.649	-0.736	0.620
Sujeto 23	0.758	0.725	0.741	0.760	0.762	1.324	0.766
Sujeto 24	0.628	0.589	0.608	0.637	0.636	-3.702	1.149
Sujeto 25	0.674	0.634	0.654	0.684	0.680	-4.901	1.326
Sujeto 26	0.695	0.591	0.639	0.677	0.681	3.275	0.510
Total	0.741	0.669	0.702	0.756	0.755	3.184	3.182

3.5.2 Evaluación del modelo

En la tabla 3.7 se muestran las diferentes métricas de evaluación de la clasificación de actividades por parte del modelo para todos los sujetos. Además, teniendo en cuenta que cada actividad tiene asociado un valor en MET, las actividades agrupadas en MET se han agrupado como pertenecientes a la misma clase. La tabla 3.8 muestra los valores de las métricas de evaluación con las actividades agrupadas por MET. Los valores totales se han calculado promediando los resultados de todos los sujetos y el total del error relativo en MET se ha calculado promediando el valor absoluto del error relativo de todos los sujetos.

Debido a la dispersión de los resultados para cada sujeto se han buscado asociaciones entre los resultados para cada sujeto y las variables descriptivas de los datos de cada sujeto tales como: número de actividades distintas de cada sujeto, número de observaciones, homogeneidad de la distribución de las observaciones en cada actividad. También

Tabla 3.8: Resultados de evaluación del modelo de clasificación por MET asociados

	Macro P	Macro S	Macro F-score	Exactitud	F-score ponderado
Sujeto 1	0.767	0.674	0.717	0.901	0.897
Sujeto 2	0.821	0.815	0.818	0.876	0.875
Sujeto 3	0.856	0.816	0.835	0.931	0.929
Sujeto 4	0.801	0.792	0.796	0.868	0.866
Sujeto 5	0.723	0.630	0.673	0.829	0.824
Sujeto 6	0.883	0.896	0.889	0.887	0.888
Sujeto 7	0.799	0.810	0.804	0.878	0.878
Sujeto 8	0.814	0.759	0.785	0.860	0.859
Sujeto 9	0.870	0.819	0.844	0.955	0.954
Sujeto 10	0.782	0.691	0.734	0.814	0.810
Sujeto 11	0.987	0.995	0.991	0.995	0.995
Sujeto 12	0.823	0.802	0.812	0.839	0.838
Sujeto 13	0.963	0.967	0.965	0.966	0.966
Sujeto 14	0.824	0.752	0.786	0.882	0.880
Sujeto 15	0.743	0.632	0.683	0.777	0.777
Sujeto 16	0.692	0.601	0.644	0.745	0.745
Sujeto 17	0.913	0.922	0.918	0.949	0.949
Sujeto 18	0.718	0.634	0.674	0.798	0.795
Sujeto 19	0.767	0.627	0.689	0.876	0.874
Sujeto 20	0.755	0.676	0.714	0.836	0.832
Sujeto 21	0.724	0.637	0.678	0.742	0.744
Sujeto 22	0.765	0.684	0.722	0.870	0.871
Sujeto 23	0.803	0.829	0.816	0.840	0.842
Sujeto 24	0.729	0.733	0.731	0.815	0.812
Sujeto 25	0.792	0.791	0.791	0.854	0.853
Sujeto 26	0.771	0.790	0.780	0.767	0.771
Total	0.803	0.760	0.780	0.860	0.859

se han buscado asociaciones con el número de características seleccionadas para cada sujeto. Se han explorado estas asociaciones para los resultados de F-score tanto para la clasificación actividades como para las actividades agrupadas en MET. La tabla 3.9 nos muestra las asociaciones encontradas para cada variable. Para comparar la homogeneidad hemos utilizado el coeficiente de variación que se define como el cociente entre la desviación estándar y la media de las observaciones en cada actividad.

Se ha encontrado que existen asociaciones significativas y la correlación es negativa, del número de actividades distintas, del número de observaciones y del coeficiente de variación tanto con el macro F-score como con el F-score ponderado para la clasificación de actividades y para la clasificación de actividades agrupadas en MET. También se han encontrado que el número de observaciones, el número de actividades y el coeficiente de variación están significativamente asociados positivamente entre sí.

Tabla 3.9: Correlación de las variables con los resultados de clasificación

	Macro F-score		Ponderado F-score		Error relativo
	Actividades	MET	Actividades	MET	
Número de observaciones	-0.65**	-0.61**	-0.71**	-0.50*	0.22
Número de actividades	-0.47*	-0.42*	-0.59*	-0.38	0.20
Coefficiente de variación de las actividades	-0.80**	-0.70**	-0.59*	-0.56*	0.25
Número de características seleccionadas	-0.11	-0.01	0.13	0.16	-0.29

* denota que el p-valor de la correlación es < 0.05 y ** denota que el p-valor es < 0.001

3.6 Discusión

Se han obtenido los resultados diferentes según el tipo de clasificador utilizado en la fase de selección del modelo. Como era esperado los algoritmos combinados obtienen mejores resultados de evaluación que los clasificadores simples. Además, los resultados obtenidos con bagging y boosting son mejores que con random subspace porque después de la selección de características el número de características es pequeño.

Las métricas de rendimiento de clasificación obtenidas son inferiores a las de estudios previos, esto puede ser debido a que las actividades son etiquetadas cada 15 minutos, que es menos preciso que en los estudios anteriores y que las actividades han sido realizadas en condiciones naturales. En términos del MET asociado al reconocimiento de actividad los resultados obtenidos son mejores.

El estudio de la asociación de los datos del sujeto con los resultados facilita la interpretación de los resultados. Encontramos que en los resultados de clasificación del modelo entrenado influyen las diferencias entre la cantidad de datos de las actividades, cabría por lo tanto intentar homogeneizar la cantidad de datos de cada actividad para mejorar los resultados. Por otro lado, el número de características seleccionadas no está asociado a los resultados, esto nos hace descartar que los resultados pudieran estar afectados por un menor o mayor número de características del modelo.

Podemos observar que existen diferencias entre los resultados macro F-score y F-score ponderado. En concreto, para la clasificación de actividades el F-score ponderado en media es de un 5% mayor que el macro F-score y para la clasificación de las actividades agrupadas por MET el F-score ponderado es alrededor de un 8% mayor. Esto nos indica que aquellas actividades que tienen menor representación empobrecen los resultados.

También podría ser que las actividades menos representadas sean más complejas y que sea necesario añadir los datos de otros sensores para clasificarlas correctamente.

3.7 Conclusiones

Hemos encontrado que el sistema es capaz de clasificar las distintas actividades con un F-score ponderado del 75.5% para un total de 26 sujetos. Además, el gasto calórico se cuantifica con un error del $3.18\% \pm 3.18\%$. Se ha obtenido un F-score ponderado de 85.9% para la clasificación de actividades en MET. Hay que destacar que los datos se han obtenido en condiciones naturales con un conjunto de actividades basado en el RSAF y que ello ha podido afectar a la precisión de la clasificación de actividades y estimación del gasto calórico. Sin embargo, estas condiciones son más realistas que si se hubieran realizado en un entorno de laboratorio de manera controlada o con un conjunto de actividades más reducido y acotado.

Como trabajo futuro sería interesante utilizar los sensores del smartwatch para realizar un análisis combinado con los datos que procedentes de los sensores del smartphone. Se realizó una aplicación para recoger los datos de los sensores, pero como el uso de smartwatch no está tan extendido, en estas pruebas no fue posible utilizar un smartwatch.

También se podría usar alguna característica más y la información procedente de otros sensores como vector de rotación, giróscopo y presión atmosférica. A nivel de algoritmos se podrían explorar soluciones basadas en algoritmos de *deep learning*.

Capítulo 4

Evaluación de la variabilidad de la frecuencia cardíaca utilizando la cámara del smartphone en tiempo real

4.1 Introducción

En la literatura podemos encontrar estudios acerca de la relación que existe entre la variabilidad de la frecuencia cardíaca (VFC) y aspectos fisiológicos y psicológicos de la vida saludable. La hipertensión, depresión y ansiedad son algunos ejemplos [160]. Generalmente, los diferentes índices de VFC se obtienen a partir del análisis de la serie temporal del tiempo entre latidos (RR o NN) detectados en la señal de ECG. Las definiciones para el cálculo de los índices de VFC están estandarizadas [37]. Por otro lado, la actividad cardíaca también se puede evaluar mediante el análisis de la señal de la onda de pulso. Además, a partir de la serie temporal pulso a pulso (PP) se pueden calcular los índices de VFC derivados. Algunos estudios han comparado los índices de VFC obtenidos a partir de la serie RR obtenida del ECG con los índices obtenidos a partir de la serie PP [161, 162].

La fotoplethismografía se basa en la medición óptica de los cambios en el volumen de la sangre en los vasos sanguíneos para medir el ritmo cardíaco [163]. Este método se basa en el hecho de que el volumen de sangre transportada en las arterias cambia durante el ciclo de bombeo de la sangre. Cuando la luz se emite sobre la piel, una proporción de luz se atenúa por el tejido de la piel y también por la sangre en los capilares. La atenuación debido al tejido de la piel es pequeña y constante [164]. Por ello la detección de la luz atenuada que atraviesa la piel da una medida del contenido de la sangre. Los procesos

de atenuación de la luz en la sangre son complicados e incluyen absorción, dispersión y reflexión [165]. En los glóbulos rojos de la sangre se encuentra una hemoproteína llamada hemoglobina que se encarga del transporte de oxígeno. Este compuesto absorbe la luz, con un diferente coeficiente de absorción de luz según la longitud de onda [163]. Por ello, el resultado es que una mayor presencia de sangre en la piel está asociada a una mayor absorción de luz con la consiguiente menor cantidad de luz captada por el fotodetector [166]. Teniendo en cuenta que el corazón rítmicamente expulsa sangre (sístole) y la atrae (diástole) en un ciclo constante, tenemos un mayor flujo de sangre a través de las arterias en la fase sistólica y una menor durante la diastólica. Por tanto, midiendo el cambio de flujo de sangre en el dedo se puede obtener una señal de ritmo cardíaco a través de la periodicidad de la señal medida. El sensor consiste en una fuente de luz y un fotodetector colocado al lado, que se colocan directamente sobre la piel. Este modo de adquisición se denomina modo de reflexión.

La cámara del *smartphone* se puede utilizar para la adquisición de la señal de fotopleletismografía (SPPG). La cámara actúa como fotosensor y el flash LED actúa como emisor. Ubicando la yema del dedo índice sobre la cámara se puede adquirir la señal de la onda de pulso de forma no invasiva utilizando únicamente el *smartphone*. La señal obtenida la denominamos como SPPG. Esta señal consta de dos componentes: una llamada corriente alterna (AC) y otra llamada corriente continua (DC). La componente DC varía lentamente debido a la respiración, la actividad vasomotora y las onda vasoconstrictoras. La componente AC es la muestra los cambios en el volumen de la sangre asociados a las fases diastólicas y sistólicas del ciclo cardíaco. La frecuencia fundamental de la componente AC depende del ritmo cardíaco y está superimpuesta a la componente DC [72].

En la literatura podemos encontrar estudios previos en los que los índices de VFC obtenidos con el *smartphone* se comparan con los obtenidos a partir del ECG. Sin embargo, estos estudios se limitan a mostrar los errores en media y la correlación pero no se evalúa el grado de acuerdo entre los resultados obtenidos con ambos métodos así que la validez clínica de estos sistemas es limitada [167]. En este capítulo se evalúa el grado de acuerdo entre el sistema desarrollado en *smartphone* y el ECG. Además, se realizan medidas simultáneas con un fotopleletismógrafo comercial para evaluar el error respecto al ECG y así poder compararlo con el *smartphone*.

En esta tesis nos hemos centrado en la evaluación de los índices de VFC a partir de la señal de pulso adquirida con el *smartphone* porque estos índices son los que están relacionados con la adherencia a hábitos de vida saludables. Los estudios que podemos encontrar en la literatura acerca de la precisión de los índices de VFC adquiridos con la señal obtenida a partir de la cámara del *smartphone* son limitados. En esta tesis hemos comparado los índices de VFC que hemos obtenido para distintos modelos de *smartphone*.

En este capítulo primero explicaremos el diseño de la aplicación móvil, luego presentaremos los resultados obtenidos en medidas con personas en distintas posturas y distintos smartphones. Finalmente, analizaremos las medidas de VFC con smartphone en un grupo más amplio de personas y compararemos distintos ritmos respiratorios y modelos de smartphone. Para ello, se ha diseñado una carcasa para smartphone con el objetivo de facilitar la colocación del dedo sobre la cámara, esta carcasa se ha utilizado en la última tanda de medidas.

4.2 Aplicación móvil para extraer la señal RGB: evaluación de la frecuencia de muestreo

Para poder extraer la señal de pulso a partir de la cámara del móvil se ha desarrollado una aplicación específica. La aplicación se encarga de: configurar la cámara, adquirir los fotogramas, procesarlos para obtener las componentes rojo, verde y azul (RGB) y almacenar la serie RGB.

A la hora de adquirir la imagen fotograma a fotograma, uno de los elementos clave es la frecuencia de muestreo y su estabilidad. De entrada, no está garantizado que la frecuencia de muestreo de adquisición de los fotogramas sea constante. A partir del análisis del tiempo de llegada de los fotogramas, se puede comprobar que cada cierto tiempo se pierden algunas muestras. Para estudiar el comportamiento de la frecuencia de muestreo y para comprobar que la adquisición de la señal es adecuada se han realizado pruebas experimentales que se detallan en la sección 4.2.2.

4.2.1 Diseño de la aplicación móvil

Se ha desarrollado una aplicación para adquirir la señal de pulso mediante la cámara del smartphone. Para ello, se utiliza el flash de la cámara como emisor y la cámara como sensor. La configuración de la cámara consiste en establecer la resolución a 640 por 480 píxeles, que es suficiente para adquirir el pulso y la carga computacional es soportable por el smartphone, fijar la exposición para reducir las derivas en la frecuencia de muestreo y encender el flash que es la fuente de luz de nuestra aplicación. Además, hay que fijar el formato de grabación de la imagen a NV21 que ofrece una buena compresión y establecer la frecuencia de muestreo a la máxima disponible que es 30 Hz para tener la mayor resolución temporal posible.

Uno de los principales retos que nos encontramos, es evitar que la frecuencia de muestreo de la cámara baje o que se pierdan muestras debido al tiempo de procesado de cada muestra. Se probó a almacenar en memoria cada fotograma de la cámara y procesarlo con la unidad central de procesamiento (CPU), pero el rendimiento era bajo y

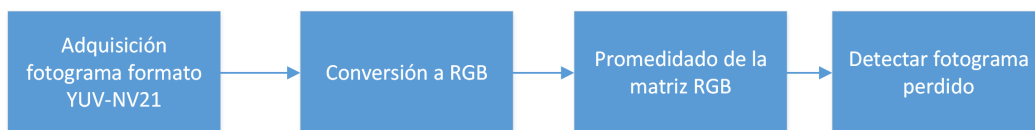


Figura 4.1: Diagrama de bloques del procesado fotograma a fotograma

no se lograba procesar a los 30 Hz que permite adquirir la cámara. Por ello, se realiza el almacenamiento de cada fotograma en la memoria de la unidad de procesamiento gráfico (GPU), de esta manera podemos procesar los píxeles del fotograma en paralelo obteniendo un mejor rendimiento. Con este método se consigue adquirir los fotogramas a 30 Hz.

Se ha utilizado la interfaz de programación de aplicaciones (API) de Renderscript para el procesado fotograma a fotograma de la imagen de la cámara usando la GPU. Renderscript API es un conjunto de funciones de programación de Android para ejecutar tareas computacionalmente intensivas. Esta API permite escribir código de script en un lenguaje basado en C99, este código es llamado de forma nativa en tiempo de ejecución y se comunica con la máquina virtual de Android. Renderscript proporciona un alto rendimiento computacional porque permite paralelizar y planificar el trabajo de forma eficiente a través de los procesadores disponibles en el procesador tales como la GPU.

Los pasos seguidos de procesado de imagen para obtener la señal RGB se describen a continuación y se muestran en la figura 4.1. Primero, la aplicación móvil adquiere un fotograma en formato luminancia-crominancia (YUV) y codificado con NV21 que proporciona una compresión fácil y eficiente. Entonces, se ejecuta un script para convertir el búfer de YUV a una matriz de píxeles de RGB. El renderscript se encarga de guardar en memoria los fotogramas y de la paralelización de la conversión. Segundo, se promedian las componentes RGB de array de píxeles para obtener la señal RGB. Esta parte también se paraleliza para que sea realizada más rápidamente. Tercero, se procesan los timestamp para detectar las tramas perdidas. Cada vez que se detecta una trama perdida, se interpola esa muestra linealmente.

La aplicación procesa los timestamp de llegada de los fotogramas para detectar fotogramas perdidos. Pero no se utiliza para la interpolación puesto que es un timestamp del sistema y no el de adquisición del fotograma. A continuación se detallan las pruebas experimentales para evaluar la precisión del método de adquisición de la señal RGB y en concreto de la componente verde que es la de interés.

4.2.2 Pruebas experimentales

Se ha diseñado un circuito electrónico para emular la señal de PPG a partir de una señal cardíaca. El esquemático del circuito se muestra en la figura 4.2. Este sistema de emulación se ha diseñado para comprobar que la señal RGB se adquiere correctamente

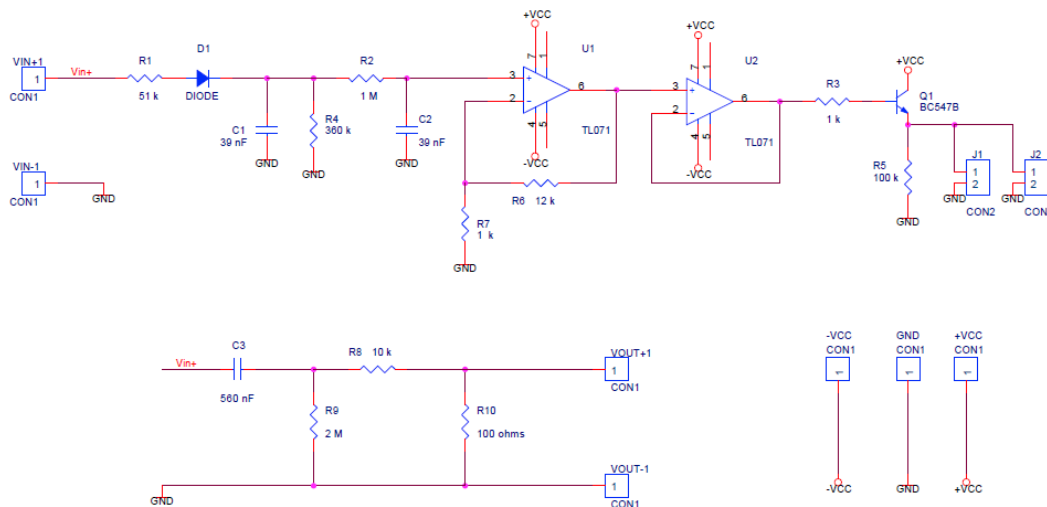


Figura 4.2: Esquemático del circuito para emular la señal de PPG

a partir de señales conocidas en un entorno controlado. Además, se ha utilizado para comparar las diferencias de procesar la señal RGB fotograma a fotograma con la alternativa de grabar un fichero de vídeo y luego procesarlo offline. Esta alternativa tiene una frecuencia de muestreo más estable. En la figura 4.3 podemos ver la componente verde de la señal RGB captada por la cámara del smartphone que está ubicada sobre el diodo.

4.2.2.1 Montaje de las pruebas

Tal como se muestra en la figura 4.4, se ha conectado el generador de funciones de onda arbitraria (modelo HP33120A) al circuito para que genere una señal cardíaca. El generador de forma de onda arbitraria (AWG) es un equipo de instrumentación electrónica que se utiliza para generar señales con formas de onda almacenadas. El AWG utiliza la técnica denominada síntesis digital directa (DDS) para generar la señal. El principio de funcionamiento es que el AWG tiene una descripción numérica de la forma de onda almacenada en memoria. Las muestras de la forma de onda se van leyendo secuencialmente y se envían al convertidor digital-analógico (DAC) y entonces con un apropiado filtrado pasabajo anti-aliasing y acondicionamiento de señal, generan a la salida la forma de onda analógica.

A la salida circuito se ha conectado una banda Polar y dos diodos, en cada diodo se ha colocado adecuadamente la cámara para que se capte la luz emitida por el diodo correctamente. Un móvil graba el vídeo a un fichero y el otro utiliza la aplicación diseñada para guardar la señal RGB fotograma a fotograma. Como utilizando la cámara del móvil la señal de la onda de pulso se obtiene a partir de la componente verde de la señal RGB, en estas pruebas los diodos utilizados son de color verde. Se han realizado las pruebas con

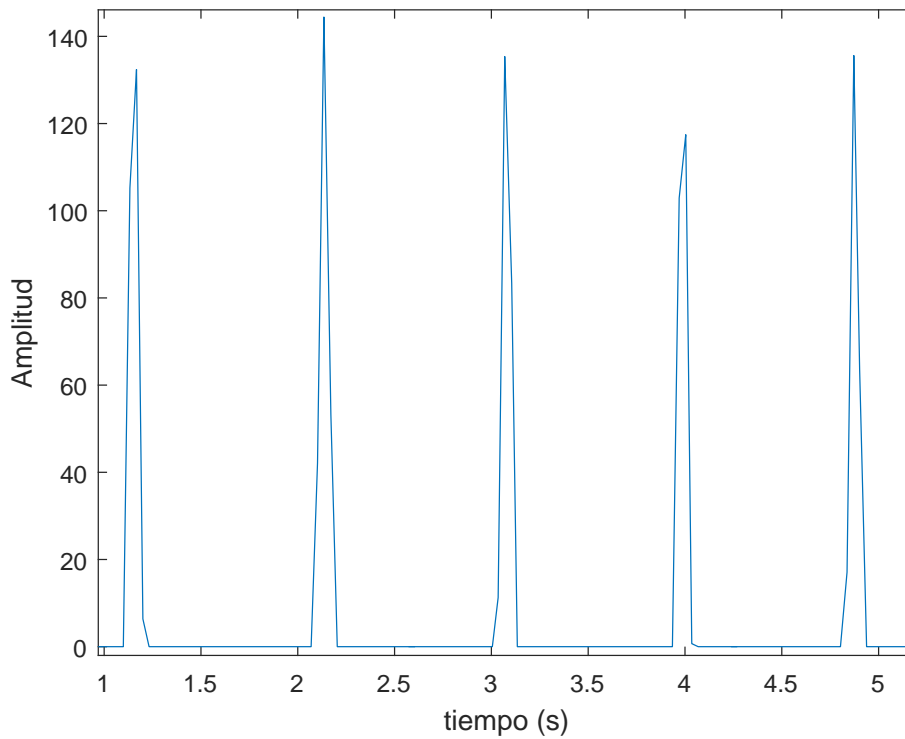


Figura 4.3: Señal pulsátil captada por la cámara del smartphone ubicado sobre el diodo

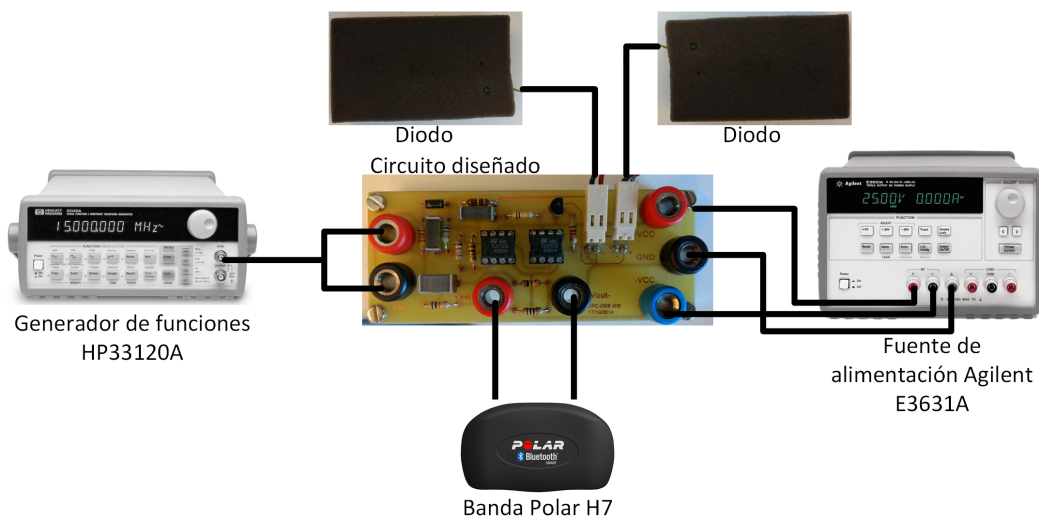


Figura 4.4: Diagrama de conexiones del montaje de las pruebas con el circuito diseñado

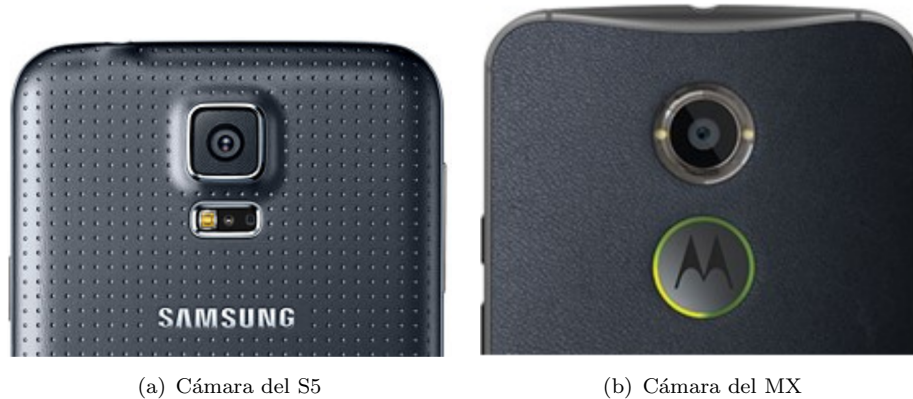


Figura 4.5: Vista de las cámaras traseras de los smartphones utilizados

dos señales diferentes: una señal a frecuencia constante de 1 Hz y una señal de frecuencia modulada. Como se pretende evaluar el rendimiento de dos modelos de smartphone se han hecho dos pruebas para cada señal. En una prueba un smartphone con la aplicación diseñada para adquirir la señal RGB y el otro en modo vídeo y en la siguiente al revés. De este modo, tenemos para cada smartphone y cada señal de entrada una grabación en modo archivo de vídeo y otra grabación fotograma a fotograma. Cada grabación tiene una duración total de 100 segundos.

Los dos smartphones utilizados son el Motorola Moto X (MX) y el Samsung Galaxy S5 (S5), se han escogido estos modelos porque la disposición del LED y la cámara son distintas pero el resto de características son similares. La cámara trasera del MX es de 13 Mpx y tiene dos flashes LED en forma de anillo y la cámara del S5 es de 16 Mpx y tiene un único flash LED. Las principales características son: CPU 4 núcleos Snapdragon-801 2.5 GHz, GPU Adreno-330 578 MHz, 2GB de RAM y versión de Android 4.4. En la figura 4.5(a) podemos ver la disposición de la cámara y el flash LED para el S5 y en la figura 4.5(b) el MX.

4.2.2.2 Procesado de señal

La componente verde de la señal RGB que se obtiene en el móvil es la que se utiliza para detectar los pulsos generados porque el diodo es de color verde. Los diodos se han escogido de color verde porque el canal verde es el que se utilizará para detectar el pulso cardíaco. Esto es debido a que tanto en luz ambiente como para luz blanca la señal de fotopleletismografía es más fuerte para el canal verde que para el azul y el rojo, porque el coeficiente de absorción de luz de la hemoglobina es mayor para esa longitud de onda [168–170].

Por un lado se detectan los fotogramas que se pueden haber perdido por no muestrear a una frecuencia constante y los corregimos interpolándolos linealmente entre las muestras vecinas. Por otro lado luego se interpola la señal de 30 Hz a 1 kHz tomando como base

Tabla 4.1: Error en la serie PP respecto a la Polar

Móvil	Señal procesada	1 Hz		FM	
		Error Media (ms)	Error SD (ms)	Error Media (ms)	Error SD (ms)
MX	Vídeo	-0,011	0,382	-0,0002634	3,9973
	Gcint	-1,2697	5,0153	-0,4289	1,6366
	Git	0,0112	6,3969	0,0093	6,452
	Gint	-0,5169	0,5025	-0,4289	1,6366
S5	Vídeo	0	0,2596	0,043	0,7793
	Gcint	1,0778	3,5386	1,0338	3,7363
	Git	0,044	4,303	-0,0684	4,9429
	Gint	10,333	16,366	8,997	20,6386

de tiempo que las muestras nos han llegado a frecuencia constante. Para comprobar la eficiencia del método de detectar fotogramas perdidos y la interpolación tomando frecuencia constante, hemos generado las siguientes tres señales:

- Gcint: se detecta y corrige los fotogramas y luego se interpola usando como base de tiempos frecuencia constante
- Git: no se detecta ni corrige los fotogramas y luego se interpola usando como base tiempos el timestamp del móvil
- Gint: no se detecta ni corrige los fotogramas y luego se interpola usando como base de tiempos frecuencia constante

Después se filtra la señal para reducir el ruido introducido por la interpolación, con un filtro paso bajo Butterworth bilineal de orden 4 a una frecuencia de corte 10 Hz. Como lo que capta el móvil es la luz emitida por el diodo, se puede utilizar la función de MATLAB para detectar los picos de señal, que básicamente busca los picos locales en la señal de entrada, donde un pico se define como aquella muestra que es mayor que sus muestras vecinas. A partir de los picos detectados se calcula la serie PP como la diferencia entre dos picos. Esta serie se ha alineado manualmente a la serie RR obtenida de la banda Polar para poder compararlas y así obtener el error cometido por el móvil con respecto a la banda Polar.

Cuando se procesa el vídeo se procesa directamente la componente verde, sin detectar pérdidas e interpolando tomando como base de tiempo que los fotogramas están muestreados a frecuencia constante. Se usa el mismo filtro y el mismo método de detección de pulsos que en el procesado fotograma a fotograma.

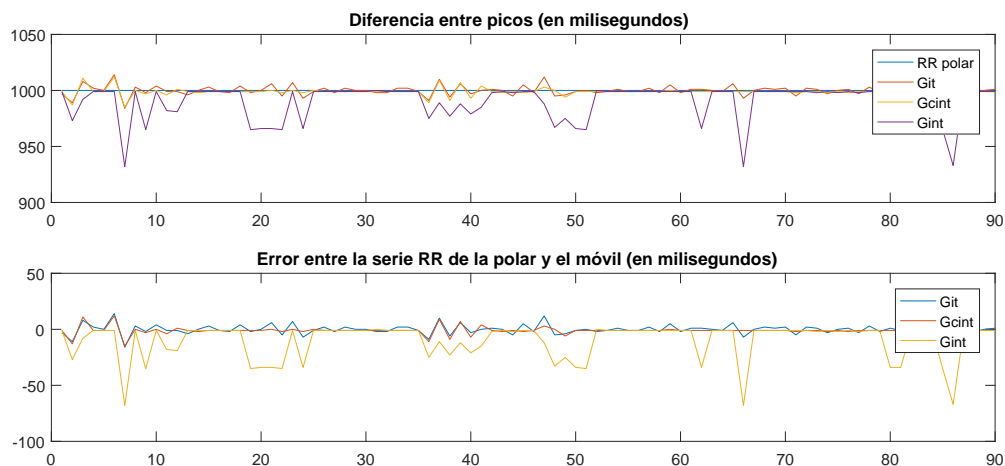


Figura 4.6: Gráfico de las distintas señales RR y del error, en este caso el móvil utilizado es el Samsung S5 y señal de entrada tiene frecuencia constante de 1 Hz

4.2.2.3 Resultados

En la figura 4.6 se muestra la serie que se obtiene al procesar las distintas señales que se han definido anteriormente: Gcint, Gint y Git.

La tabla 4.1 muestra los resultados para los distintos móviles, distinto procesado de señal y para la señal a 1 Hz y la FM, mostrando el error en media y desviación estándar (SD). Podemos ver que la SD del error de la señal Git es mayor que Gcint tanto para MX como S5 y para ambas señales. También para el S5 el error en Gcint es menor que en Gint para ambas señales. Por el contrario, para el MX en FM son iguales los errores de Gint y Gcint y en 1 Hz Gint tiene menos error que Gcint. Esto puede ser debido a que para el MX no se han perdido muestras aunque haya variaciones en los valores del timestamp. Como consecuencia queda demostrado que para interpolar es más preciso utilizar una base de tiempos que asume que los timestamp son constantes que utilizar el timestamp de llegada del fotograma en la aplicación del móvil. Para el Samsung S5 también podemos ver como la detección e interpolación del fotograma perdido hace que sea más preciso. Acerca de la señal de vídeo, la SD de los errores es menor que el resto puesto que en la grabación a fichero la frecuencia de muestreo es más constante.

Posteriormente, a través de la visualización de la imagen con el dedo puesto en la cámara se ha fijado la exposición que maximiza la sensibilidad de la componente verde. Este parámetro podría variar según el smartphone utilizado, en este caso se ha comprobado para los móviles Samsung S5 y Motorola Moto X.

4.3 Evaluación de la influencia de la postura en la medida del ritmo cardíaco obtenido en el smartphone

En la literatura podemos encontrar estudios acerca de la influencia de la postura en el ritmo cardíaco y en la VFC obtenido a partir de la señal del ECG. Estos estudios encontraron que el ritmo cardíaco aumenta y la VFC disminuye al pasar de la postura supina a sentado y de sentado a de pie [171–173]. Esto se debe a que cuando se producen estos cambios de postura se produce una activación del sistema nervioso autónomo simpático y un decrecimiento de la actividad parasimpática. Además al pasar de la posición supina a de pie hay un decremento de la actividad vagal [174].

En esta sección se describe la evaluación del método de adquisición del pulso cardíaco medido en personas utilizando la cámara del smartphone. Para ello, se ha realizado una tanda de medidas para evaluar la influencia de la postura en el error de la medida del pulso con el móvil con respecto a un dispositivo comercial que obtiene el ritmo cardíaco a partir del ECG. El dispositivo comercial utilizado es la banda Polar, cuyo uso está muy extendido en el ámbito del deporte y se ha evaluado que es precisa [175]. Se han hecho medidas con dos modelos de smartphones diferentes para evaluar las diferencias en el error según el modelo utilizado. Como se ha comentado anteriormente en la sección 4.2.2.1, los smartphones utilizados son el Motorola Moto X (MX) y el Samsung Galaxy S5 (S5).

4.3.1 Medidas realizadas

11 personas sanas de entre 21 y 45 años fueron reclutadas para participar en estas medidas. Se utilizó la banda Polar H7 para adquirir la serie RR a partir del ECG. A los sujetos se les pidió que llevaran la banda Polar alrededor del pecho. Simultáneamente, ellos sostenían el smartphone en su mano y ponían el dedo índice sobre la lente y el flash de la cámara. Las pruebas fueron llevadas a cabo bajo supervisión para asegurar que el smartphone se mantenía bien colocado durante la medida. A los participantes se les pidió que llevaran un smartphone en cada mano.

Además, a los participantes se les pidió que permanecieran lo más quietos posible en tres posturas diferentes para poder evaluar la influencia de la postura en el error: sentado con los brazos estirados hacia abajo, tumbado en posición supina con los brazos estirados hacia abajo y sentado con los brazos levantados al nivel del corazón. En total cada participante fue medido cuatro veces: dos para la medida supina y una vez para cada una de los restantes posturas. La duración de la medida para cada una de las posturas fue de 100 segundos. Fue elegida esta duración porque parecía suficiente para comparar los resultados teniendo en cuenta que se realizaban cuatro medidas por persona.

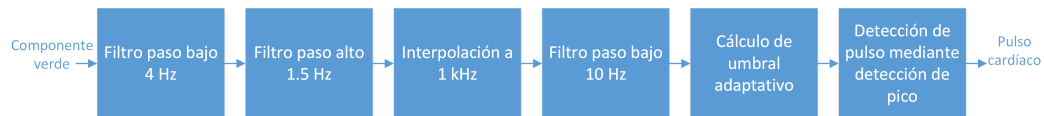


Figura 4.7: Diagrama de bloques del procesamiento de la componente verde de la señal RGB para obtener el pulso cardíaco

4.3.2 Extracción de la serie PP basado en smartphone

La componente verde de la señal RGB de la cámara del smartphone es la que aporta más información para la medida del pulso [176]. Por eso, una vez hemos obtenido la RGB tal como se ha explicado en secciones anteriores, nos quedamos con la componente verde para obtener la señal de pulso. Una vez están corregidas las pérdidas de fotogramas detectadas, tratamos la señal con la finalidad de obtener el pulso cardíaco siguiendo los pasos mostrados en la figura 4.7. En concreto, filtramos la señal para reducir el ruido con un filtro Butterworth de segundo orden paso bajo con una frecuencia de corte de 4 Hz. Luego, filtramos la señal con un filtro Butterworth de segundo orden paso alto con frecuencia de corte 1.5 Hz para filtrar la línea base y analizar la derivada de la señal de SPPG. Después interpolamos la señal a 1 kHz mediante una interpolación cúbica de 3 puntos para aumentar la resolución temporal. Finalmente, la señal interpolada se filtra paso bajo a 10 Hz con un filtro de Butterworth de orden 2 para reducir el ruido que puede ser introducido por la interpolación.

Una vez que la señal está interpolada y filtrada, consideramos que la señal ya está acondicionada para detectar el pulso cardíaco. En este caso hemos utilizado un algoritmo para detectar el pulso cardíaco más sofisticado que la función de MATLAB para buscar picos.

Los picos de la onda SPPG se obtienen utilizando un umbral adaptativo. El valor inicial del umbral se obtiene a partir del 70% de la desviación estándar de los primeros 5 segundos de señal. Se calcula un umbral adaptativo a partir de 0.5 veces la media de los últimos cinco picos detectados. Además se le aplica un coeficiente de corrección cuando el pico es mayor que 2.5 veces el pico actual para evitar cambios abruptos. La siguiente fase es la propia detección de los picos a partir del umbral adaptativo. Los picos son detectados a partir de los cruces de la señal del umbral de abajo hacia arriba. El pico es detectado buscando el máximo durante los siguientes 400 ms, $12 * N$ ($N=33$ factor de interpolación) muestras. A partir de las diferencias entre los picos obtenemos la serie temporal pulso a pulso (PP).

4.3.3 Índices de variabilidad de frecuencia cardíaca

A partir de la serie PP del smartphone además de calcular la media de estos intervalos (media NN), se han calculado los siguientes índices de VFC para evaluar el error cometido

respecto al ECG: SDNN, RMSSD y LF/HF. Estos índices son los más utilizados y se han calculado siguiendo el estándar [37] que los describe como:

- SDNN: desviación estándar de todos los intervalos NN
- RMSSD: raíz cuadrada de la media de la suma de los cuadrados de las diferencias entre sucesivos intervalos NN
- LF/HF: ratio entre potencia en baja frecuencia (0.04-0.15 Hz) y alta potencia (0.15-0.4 Hz)

4.3.4 Análisis de los datos

La figura 4.8 muestra los diferentes pasos realizados para analizar los datos. Las series PP y RR se han alineado automáticamente mediante la maximización de los coeficientes de la correlación intraclass (ICC). Para evaluar el error se ha utilizado la desviación estándar del error (SDE) y se ha calculado la correlación entre las series expresada por los coeficientes de ICC. Se ha utilizado la correlación ICC en lugar de la Pearson, porque la ICC no ignora el sesgo sistemático así que no solo evalúa la correspondencia sino también la concordancia. Además se han calculado tres índices de VFC para registros cortos en ambas series para analizar las diferencias entre ellos: SDNN, RMSSD y LF/HF.

Finalmente, se ha empleado un análisis de varianzas (ANOVA) de tercer orden para evaluar si el SDE, ICC y los errores relativos de los índices VFC obtenidos estaban influenciados por el modelo de smartphone, la postura y el sujeto. Adicionalmente, se

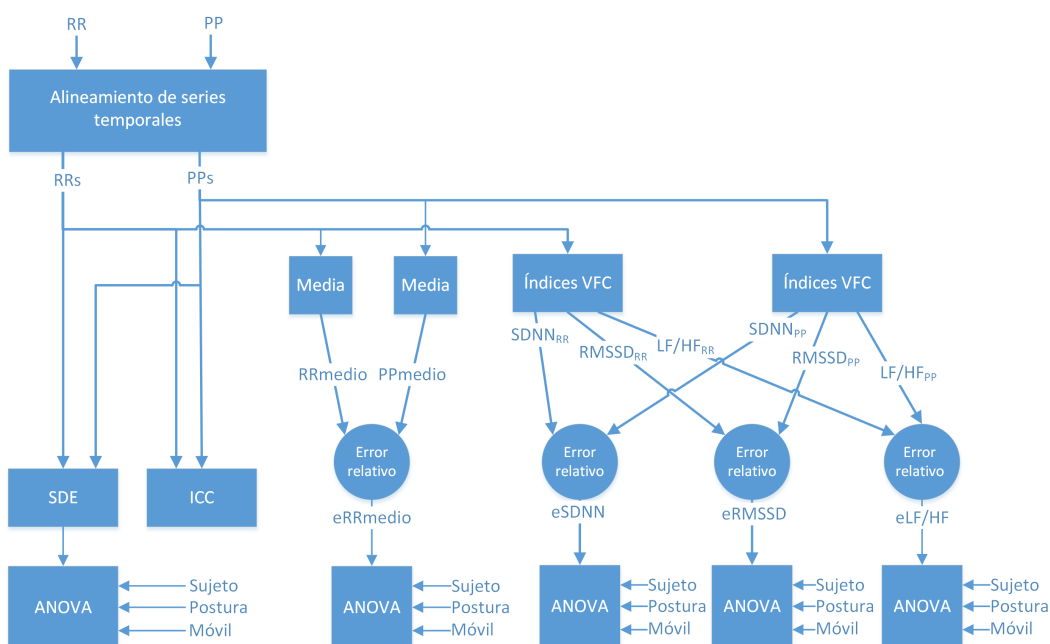


Figura 4.8: Diagrama de bloques del análisis de las series temporales RR y PP

ha realizado un t-test de Student pareado para comparar las diferencias en el SDE entre los dispositivos MX y S5 por cada postura.

4.3.5 Rendimiento de la aplicación móvil diseñada

Se ha evaluado el rendimiento de la aplicación a nivel de consumo de CPU y de GPU. Para ello se ha utilizado la aplicación Trepro Profiler (https://play.google.com/store/apps/details?id=com.quicinc.trepro&hl=en_US), esta aplicación es una herramienta de diagnóstico del rendimiento diseñada por Qualcomm. Específicamente se ha utilizado para medir el uso de la CPU y GPU de la aplicación. Hemos obtenido que la carga de la CPU en el S5 fue de $43.8\% \pm 11.6$ que es mayor que para el MX $34.4\% \pm 6.3$. Por otro lado la carga de la GPU para el S5 fue de $15.3\% \pm 0.61$ que es también mayor que para el MX $6.21\% \pm 1.21$. A pesar de la diferencia en el rendimiento de la CPU y GPU para ambos móviles, estos resultados nos muestran que ambos modelos de smartphone tienen la suficiente capacidad computacional para realizar las tareas requeridas.

4.3.6 Resultados

En la tabla 4.2 se muestra la media \pm SD del SDE y el ICC para todas las medidas realizadas en cada postura y para cada modelo de smartphone. De esta tabla se desprende que existe una alta correlación ICC ($r > 0.95$) entre la serie PP y la serie RR en todas las posturas para ambos modelos de smartphone. Para todas las posturas el SDE es menor en el S5 que en el MX. El análisis ANOVA de tercer orden muestra que el SDE depende del sujeto de medida y del modelo de smartphone pero no hay diferencias significativas asociadas a la postura del sujeto. Analizándolo más en profundidad, el t-test de Student pareado no muestra diferencias significativas del SDE entre el S5 y el MX en ambas posturas sentadas. Sin embargo, se han encontrado diferencias significativas ($p < 0.05$) para la posición supina donde el error del S5 es menor que el MX.

Los resultados del error relativo entre la media del RR y los índices VFC obtenidos por la banda Polar y los obtenidos por el smartphone se muestran en la tabla 4.3. El ANOVA de tercer orden muestra que para la media del RR hay diferencias significativas según el

Tabla 4.2: Medida del error de la serie PP respecto a la Polar

Postura	SDE (ms) media \pm SD		ICC (.r) media \pm SD	
	MX	S5	MX	S5
Sentado a nivel del corazón	7.81 ± 3.81	5.67 ± 2.50	0.98 ± 0.011	0.99 ± 0.0093
Sentado brazos bajados	7.88 ± 2.48	7.08 ± 2.64	0.97 ± 0.024	0.96 ± 0.026
Supina	6.83 ± 2.32	5.20 ± 2.19	0.98 ± 0.026	0.99 ± 0.0093

Tabla 4.3: Error relativo de la media de RR y los índices de VFC entre la banda Polar y los *smartphones* expresados en media \pm SD

Postura	eRRmedio (%)		eSDNN (%)		eRMSSD (%)		eLF/HF (%)	
	MX	S5	MX	S5	MX	S5	MX	S5
Sentado a nivel del corazón	0.12 \pm 0.10	0.18 \pm 0.006	4.25 \pm 1.28	2.68 \pm 2.01	13.6 \pm 8.02	11.4 \pm 8.26	16.6 \pm 14.0	11.9 \pm 11.4
Sentado brazos bajados	0.094 \pm 0.04	0.15 \pm 0.03	5.00 \pm 2.73	5.24 \pm 3.14	18.0 \pm 20.2	39.3 \pm 35.6	16.5 \pm 9.38	27.2 \pm 12.1
Supina	0.092 \pm 0.06	0.18 \pm 0.02	5.02 \pm 2.99	3.31 \pm 2.04	14.6 \pm 10.9	9.4 \pm 5.62	15.3 \pm 11.6	16.07 \pm 9.74

modelo del *smartphone* ($p < 0.001$) pero no según el sujeto ni la postura. El error relativo de la media del RR es pequeño para ambos *smartphones*.

Se han encontrado diferencias significativas en el error relativo del SDNN en el sujeto ($p < 0.001$) y en el modelo del *smartphone* pero no hay diferencias en la postura. El error relativo del SDNN para el S5 es menor que en el MX. Además, se han encontrado diferencias significativas en el error relativo del RMSSD en el sujeto ($p < 0.001$) y la postura ($p < 0.05$) pero no en modelo de *smartphone*. El error relativo del RMSSD es más elevado para la postura sentado con los brazos bajados. Finalmente, el sujeto tiene un efecto significativo en el error de LF/HF ($p < 0.001$) pero no hay diferencias significativas asociadas a la postura del sujeto ni al modelo del *smartphone*.

4.3.7 Discusión

En este estudio se han encontrado diferencias asociadas al modelo de *smartphone* en el SDE y la media del SDE es menor en el S5 que en el MX. Esto puede deberse a la distinta disposición de la iluminación de la cámara. Aunque el MX tiene dos emisores de luz y el S5 uno, la luz emitida por el S5 es más intensa que la del MX. Además otros factores relacionados con las especificaciones de la cámara tales como la sensibilidad o un post-procesado embebido pueden afectar.

Por otro lado, se han encontrado diferencias significativas del error relativo del RMSSD asociado a la postura del sujeto. Sin embargo, las diferencias del error absoluto de RMSSD asociado a la postura del sujeto no eran estadísticamente significativas. Además, hay diferencias significativas del RMSSD asociadas a la postura. Más sujetos deben ser medidos para explicar la influencias de la postura en el error del RMSSD.

4.3.8 Conclusiones

Se ha desarrollado una aplicación de smartphones para la adquisición en tiempo real de los intervalos PP mediante el procesado de imagen fotograma a fotograma. Además, los resultados indican que el error relativo de la media del RR es muy bajo y la SDE es 7.81 ± 3.81 ms en el peor caso. Se ha encontrado que el modelo del smartphone tiene una influencia significativa en el error relativo del SDNN, el cual es menor usando el S5 que el MX. Entonces, esto sugiere que el S5 puede ser más adecuado que el MX para esta aplicación.

Acerca de los errores relativos de los índices de VFC obtenidos: el error relativo de SDNN es aceptable pero el del RMSSD y LF/HF puede ser demasiado elevado según el escenario de uso.

Estos resultados tienen la principal limitación de que el gold estándar es la banda polar y no la señal del ECG adquirida con electrodos. Por otro lado, la postura con brazos bajados resultaba poco natural lo que dificultaba la medida.

El hecho de tener los móviles en distintas manos puede introducir variaciones debido a la diferencias en el tiempo de llegada del pulso. Como trabajo futuro, las medidas se deberán realizar siempre en la misma mano para evaluar la influencia de los móviles.

Puede haber diferencias entre móviles pero se deben realizar un mayor número de medidas. La realización de algunas medidas ha costado bastante debido a que el dedo debe estar correctamente puesto. Como trabajo futuro, se debería facilitar a los sujetos la correcta ubicación el dedo.

Para aumentar la fiabilidad de las medidas de VFC está recomendado que duren 5 minutos. Por ello como trabajo futuro queda evaluar las diferencias del VFC en pruebas de duración de 5 minutos para que sean más fiables. Por otro lado, es necesario realizar pruebas con un mayor número de sujetos para sacar conclusiones más firmes.

4.4 Evaluación del error en la medida del pulso cardíaco con la cámara del móvil y de la VFC con respecto al ECG y PPG: influencias del modelo de smartphone y respiración

4.4.1 Introducción

Se han realizado una serie de medidas con el objetivo de comparar la adquisición de la serie PP obtenida con el smartphone (SPP) y la serie RR adquirida con el ECG utilizado como referencia. Además, en estas medidas también se ha utilizado un fotopleletismógrafo

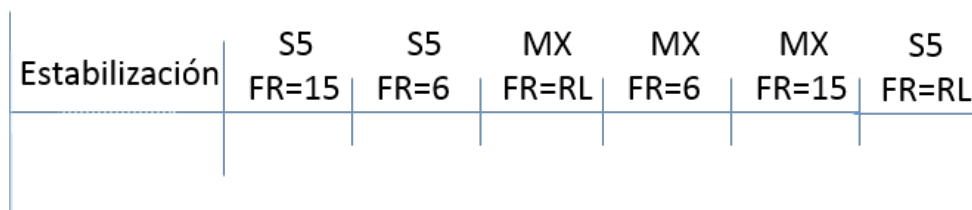


Figura 4.9: Ejemplo del cronograma del protocolo de medida con diferentes modelos de *smartphone* y frecuencias respiratorias (FR). Para cada espacio de tiempo, se muestra la frecuencia respiratoria y el modelo de *smartphone* utilizado

comercial para comparar los errores con respecto al ECG del fotopleletismógrafo y el *smartphone*. Para mejorar la ergonomía a la hora de utilizar la aplicación se ha diseñado una carcasa que ayuda a colocar el dedo correctamente en la cámara y se han realizado las medidas usando esta carcasa. Estas medidas se han utilizado para buscar el punto fiducial donde el error entre la serie RR y SPP es menor. En estas medidas también se ha evaluado la influencia en el error de la respiración y del modelo de *smartphone* empleado y se han comparado las distintas fuentes posibles de error.

4.4.2 Datos recogidos

Se reclutaron 23 sujetos sanos para estas pruebas de edades comprendidas entre los 22 a 56 años, los cuales 18 fueron hombres y 5 mujeres. Todos los participantes dieron su consentimiento informado y las pruebas fueron realizadas siguiendo los principios de la Declaración de Helsinki [177]. El ECG fue medido con 3 electrodos (3M Red Dot 2560) pegados a la piel colocados en la derivación estándar I y la señal de pulso fue medida con el fotopleletismógrafo (SS4LA) colocado en el dedo medio de la mano derecha. El ECG y el PPG fueron adquiridos a 5 kHz con el dispositivo de adquisición Biopac MP36. La señal de fotopleletismografía adquirida con el *smartphone* (SPPG) fue adquirida con dos modelos distintos el Samsung Galaxy S5 (S5) y el Motorola Moto X (MX), ambos a una frecuencia de muestreo de 30 Hz tal y como se describió en la sección 4.2. Se usaron estos dos *smartphones* porque como se comentó en la sección 4.2.2.1 la disposición del flash LED y la cámara son distintas. En las medidas, el *smartphone* se coloca siempre en el dedo índice de la mano derecha por ello estas medidas fueron realizadas dos veces, una para cada modelo de *smartphone* empleado. El orden del modelo de *smartphone* utilizado en cada medida fue aleatorizado. Después de un período de 5 minutos de estabilización del ritmo cardíaco, seis medidas consecutivas de 5 minutos fueron realizadas bajo tres patrones respiratorios diferentes: respiración libre (RL), controlada a 6 respiraciones por minuto (rpm) y controlada a 15 rpm. Las secuencias de patrones respiratorios han sido también aleatorizadas. La duración de cada medida fue de 5 minutos porque se pretenden medir índices VFC sobre segmentos cortos los cuales tienen una duración de 5 minutos [37]. Los sujetos permanecieron sentados durante toda la prueba. Un ejemplo

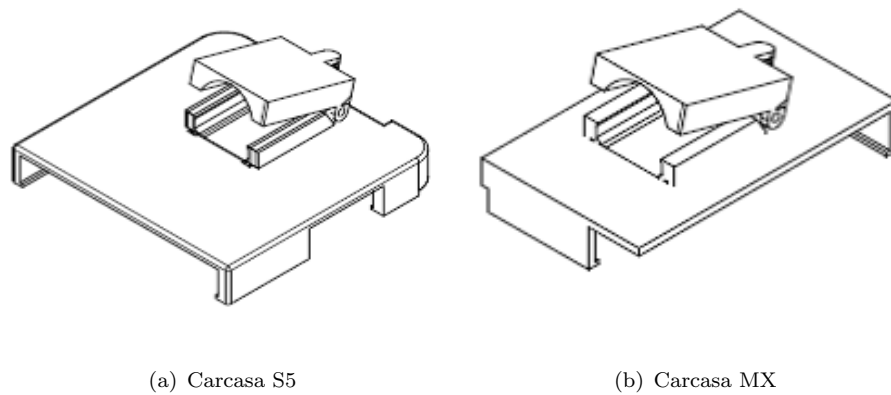


Figura 4.10: Modelo 3D de las carcasas para cada uno de los smartphones

del cronograma del protocolo de medida se muestra en la figura 4.9. Con estas medidas, el ECG, PPG, SPPG se miden simultáneamente para evaluar las diferencias entre ellos relacionadas con el modelo del smartphone y los patrones respiratorios y para comparar el grado de acuerdo de la VFC entre ellos.

4.4.3 Diseño de la carcasa

Para que la señal de pulso pueda ser adquirida de forma precisa, el dedo debe estar sobre el flash y el objetivo de la cámara en una manera adecuada y reproducible. Por ello, se ha diseñado una carcasa para el smartphone para reducir la variabilidad causada por cambios en la colocación del dedo. En investigaciones previas, se ha encontrado que el posicionamiento del dedo tiene un papel importante en la adquisición del SPPG. En concreto, pueden aparecer algunos artefactos en la señal PPG debido a ligeros movimientos del dedo.

La carcasa diseñada tiene dos partes: una parte rígida sobre el smartphone con un agujero sobre la cámara y la otra parte que tiene la forma de la mitad de un dedal y forma una pinza. Por tanto, la carcasa diseñada puede ayudar a obtener medidas más fiables evitando posicionamientos inadecuados del dedo y movimientos del dedo no deseados. El diseño de la carcasa para el S5 se muestra en la figura 4.10(a) y para el MX en la figura 4.10(b).

4.4.4 Procesado de la señal ECG, PPG y SPPG

Las señales ECG, SPPG, PPG fueron procesadas offline para obtener las series temporales latido a latido. Se ha utilizado el detector del complejo QRS Pan-Tompkins [178] para obtener la serie RR a partir del ECG y después de la detección, la localización del complejo QRS ha sido refinada usando corrección por plantilla.

Tabla 4.4: Frecuencias de corte usadas en el procesado de la señal PPG y SPPG

Punto Fiducial	Frecuencia paso alto (Hz)	Frecuencia paso bajo (Hz)
FP 1	0.846	5.303
FP 2	0.107	2.726
FP 3	0.752	3.507

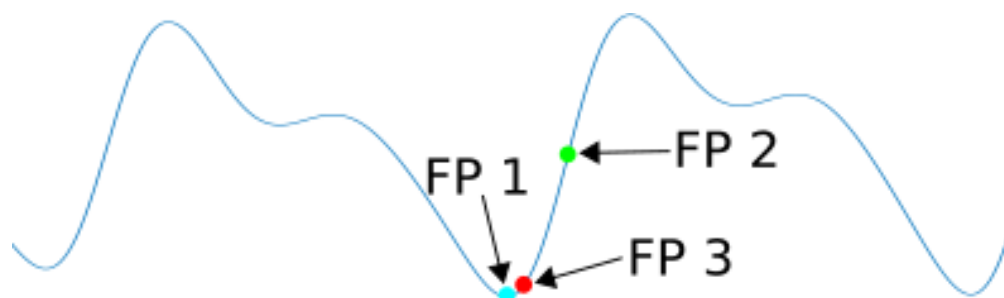


Figura 4.11: Ejemplo de la señal PPG con diferentes puntos fiduciales (FP): FP1 es el inicio de la señal PPG, FP2 es el máximo de la primera derivada de la señal y FP3 es el máximo de la segunda derivada

La llegada de la onda de pulso a partir del SPPG se ha detectado en los tres puntos fiduciales. Primero, se filtra la señal paso banda usando un filtro Butterworth de orden 4 bidireccional para eliminar la derivas de línea base y reducir el ruido. Para cada punto fiducial, se han utilizado las diferentes frecuencias del filtro que minimizan la desviación estándar error entre las series RR y PP. Estas frecuencias fueron encontradas en un trabajo previo [179] y son mostradas en la tabla 4.4. Segundo, la señal filtrada es interpolada de 30 Hz a 5 kHz mediante una interpolación cúbica de 3 puntos para obtener la misma resolución temporal en SPPG que en PPG y ECG. Tercero, para reducir el ruido, la señal interpolada es filtrada paso bajo a 10 Hz con un filtro de Butterworth bidireccional de orden 4. Entonces, los picos de pulso se detectan con un umbral adaptativo para cada punto fiducial. Los puntos fiduciales se muestran en la figura 4.11 y se definen como: mínimo de la señal (FP1), máximo de la primera derivada (FP2) y máximo de la segunda derivada (FP3). Estos puntos fiduciales han sido ampliamente utilizados en la literatura existente [84, 180, 181]. Por tanto, la serie pulso a pulso obtenida por el smartphone (SPP) ha sido obtenida para cada punto fiducial.

Los métodos de filtrado y detección de picos del pulso mencionados anteriormente, también se han utilizado para obtener la serie temporal PP obtenida a partir de la señal PPG del Biopac. La presencia de valores aberrantes en las series PP y SPP han sido detectados mediante inspección visual para detectar la presencia de artefactos en la señal debido a movimientos. Se ha considerado como valor aberrante un intervalo PP o SPP si difiere más de 30 ms del intervalo RR. Finalmente, aquellas medidas en las que más del 10% de sus picos PP son valores aberrantes o 5 picos consecutivos toman valores

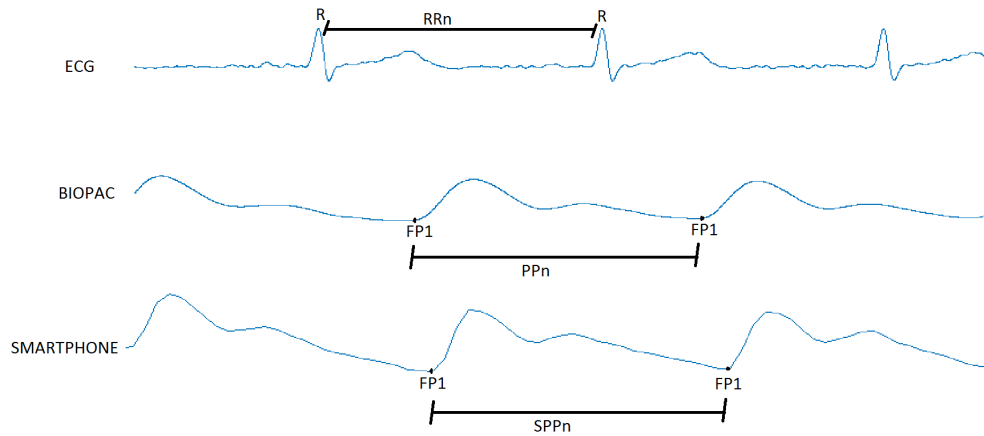


Figura 4.12: Ejemplo de las tres señales adquiridas: ECG, PPG y SPPG y los intervalos de tiempo obtenidos latido a latido: RR, PP y SPP

aberrantes fueron rechazadas para el análisis. Un total de 138 medidas fueron registradas (23 sujetos por 3 frecuencias de respiración y 2 smartphones). 13 medidas del MX y 1 del S5 han sido descartadas debido a la presencia de artefactos. Por lo tanto, 124 medidas han sido consideradas en este estudio.

La figura 4.12 muestra un ejemplo de las señales adquiridas mencionadas anteriormente: ECG, PPG y SPPG. Esta figura muestra los intervalos de tiempo latido a latido que se obtienen para cada señal: RR, PP y SPP respectivamente.

4.4.5 Evaluación del error y análisis estadístico de los datos

Las series temporales RR, PP y SPP deben ser alineadas para poder evaluar el error. Por eso, estas series se han sincronizado para obtener las series alineadas con el mismo número de muestras. Este proceso consiste en que una serie temporal es retrasada o avanzada respecto a la otra un número finito de muestras hasta que se maximiza el coeficiente de correlación intraclase (ICC) de Fisher. Debido a que las series temporales son medidas casi en el mismo momento, se ha buscado el máximo de ICC mediante el avance o retraso de no más de 10 latidos [182].

La desviación estándar del error (SDE) se ha utilizado para cuantificar el error entre los distintos métodos: SPP, PP y RR. Además, se ha utilizado para encontrar el mejor punto fiducial, aquel que minimiza el SDE. El test de Mack-Skilling [183] es una extensión del test de Friedman no paramétrico ANOVA de segundo orden que se utiliza para diseños de bloque no balanceados e incompletos, cuando el número de observaciones para cada tratamiento y bloque es uno o mayor. Este método se ha utilizado para testear las diferencias del SDE entre puntos fiduciales para todos los sujetos. Además para analizar las diferencias significativas entre pares de puntos fiduciales, se ha realizado el test de Mack-Skilling para todos los pares de puntos fiduciales. Después, se ha utilizado la

corrección de Holm para ajustar los p-valores del test en comparaciones múltiples [184]. Además, la correlación entre las series SPP, PP y RR se ha evaluado con el ICC.

La SDE también se ha utilizado para estudiar las influencias del modelo de smartphone y de la respiración en el error. Las influencias en el SDE se han analizado a partir del test ANOVA de segundo orden no paramétrico (test de Friedman). Se utilizaron métodos no-paramétricos porque los residuos de la regresión no estaban normalmente distribuidos. Para cada uno de los tipos de respiración, se ha utilizado un test de Friedman para analizar las diferencias entre smartphones. Del mismo modo, para cada smartphone, un test de Friedman se ha utilizado para analizar las diferencias entre respiraciones. El test post-hoc de Nemeyi se ha utilizado para comparar los tipos de respiración entre ellos [185].

El grado de acuerdo de los índices de VFC entre los métodos se ha obtenido usando el método Bland-Altman (BA) y también se ha evaluado la influencia del modelo del smartphone en este grado de acuerdo. Este método se ha usado para evaluar no sólo la exactitud sino también la precisión de los índices VFC. El test Mack-Skilling se ha utilizado para testear las diferencias de los índices de VFC entre los modelos de smartphone para todos los sujetos. La media y la SD de las diferencias y el 95 % límites de acuerdo (LA) ($\pm 1.96 * SD$) se han calculado. Las series temporales RR se han tomado como método de referencia y se han calculado la media del NN y tres índices de VFC para registros cortos: SDNN, RMSSD y LF/HF. Para cuantificar el grado de acuerdo, también se ha utilizado el ratio BA, que se calcula como la mitad del rango entre los LA dividido por la media de las medias pareadas de las medidas. Se define como buen acuerdo si el ratio $BA \leq 10$, como moderado si el ratio BA está ente 10 y 20 e insuficiente si el ratio está por encima de 20 [186]. Los datos han sido procesados y analizados usando R 3.3.1 y MATLAB 2016b.

4.4.6 Evaluación de diferentes fuentes de errores

Las desviaciones de la serie pulso a pulso obtenidas con el fotopleletismógrafo y el smartphone con respecto a las series latido a latido derivadas del ECG, pueden ser atribuidas a diferentes razones tales como una inadecuada detección del pulso causadas por limitaciones del sistema, artefactos, ruido o variabilidad en el tiempo de llegada del pulso (PAT). La figura 4.13 muestra un diagrama de bloques de los métodos de procesado utilizados para obtener diferentes señales que puedan ser analizadas para evaluar las diferentes fuentes de error asociadas a las diferentes partes del método propuesto. Cabe remarcar que la señal obtenida SPPG se adquiere con una frecuencia de muestreo de 30 Hz pero el ECG y el PPG se adquieren a 5 kHz. Por eso, la señal de SPPG ha sido interpolada a 5 kHz. La SDE ha sido utilizada para evaluar las principales fuentes de errores. Las siguientes fuentes de error pueden ser distinguidas:

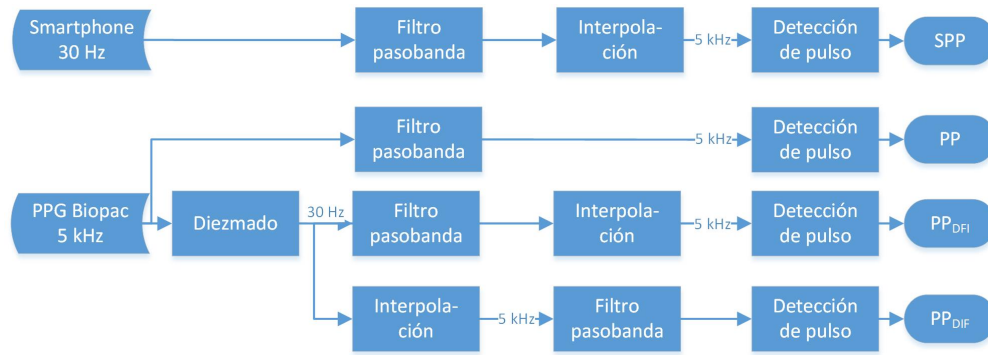


Figura 4.13: Diagrama de bloques de los métodos de acondicionamiento de señal anteriores a la detección del pulso. El filtrado paso banda y el método de interpolación son los detallados en la sección 4.4.4. El diezmado filtra paso bajo la entrada para evitar el aliasing y submuestra el resultado.

- Errores asociados al método de adquisición de señal incluyendo las limitaciones hardware (potencia del flash, sensor de la cámara) y software (Android SO, aplicación diseñada) del smartphone usado, comparando el SDE del SPP y PP_{DIF}
- Errores asociados a la pérdida de información del SPP debido a utilizar una frecuencia de muestreo baja, evaluando las diferencias del SDE entre PP y PP_{DIF}
- Errores asociados a la influencia en el algoritmo de la frecuencia de muestreo, comparando las diferencias del SDE entre PP y PP_{DIF}

4.4.7 Resultados

4.4.7.1 Comparación de las series PP y SPP con respecto al RR: influencia de los puntos fiduciales

La tabla 4.5 muestra los resultados del SDE de la serie de intervalos RR entre ECG, smartphone (SPP) y fotopleletismógrafo (PP). Se han encontrado diferencias significativas entre los puntos fiduciales FP2 y FP3 y entre FP1 y FP2 para RR-SPP, RR-PP y PP-SPP y entre FP1 y FP3 para RR-PP y PP-SPP. Los valores de ICC obtenidos son similares y cercanos a 1 para todos los puntos fiduciales. Además, la mediana del SDE del FP2 es menor que del FP1 y FP3 para RR-SPP y RR-PP. Entonces, el punto fiducial FP2 ha sido utilizado para el resto del análisis. Tenemos que para este punto fiducial, la mediana del SDE para RR-PP (4.9 ms) es menor que RR-SPP (5.4 ms) pero es mayor que el PP-SPP (3.95 ms).

4.4.7.2 Evaluación de las diferentes fuentes de errores

La tabla 4.6 muestra los resultados de la desviación estándar de los errores que se han obtenido para evaluar las fuentes de errores. Estos resultados se han calculado

Tabla 4.5: Resultados para los distintos puntos fiduciales

Punto fiducial	RR-SPP	RR-PP	PP-SPP
SDE ms			
FP1 (n=124)	6.54 (5.00–8.48)*	5.42 (4.01–7.99)*	4.75 (3.93–5.91)
FP2 (n=124)	5.4 (4.06–7.35)†	4.90 (3.30–6.99) †	3.95 (2.98–5.09) †
FP3 (n=124)	6.2 (4.34–7.89)	5.24 (3.6–7.86) ‡	3.34 (2.59–4.6) ‡
ICC			
FP1	0.995 (0.988–0.998)	0.997 (0.993–0.999)	0.998 (0.991–0.999)
FP2	0.997 (0.992–1)	0.998 (0.994–1)	0.999 (0.995–1)
FP3	0.996 (0.99–0.999)	0.997 (0.994–0.999)	1 (0.997–1)

Los datos se presentan como mediana y rango intercuartil (Q25 %–Q75 %). P-valores aplicando el ajuste de Holm: entre FP1 y FP2 $p < 0.05^*$, entre FP2 y FP3 $p < 0.05^\dagger$, entre FP1 y FP3 $p < 0.05^\ddagger$

Tabla 4.6: SDE de las series utilizadas para evaluar las fuentes de errores

SDE mediana (Q25 %–Q75 %) ms			
SPP-PP _{DFI}	PP-PP _{DFI}	PP-PP _{DIF}	RR-PP _{DFI}
4.01 (3.00–5.01)	0.28 (0.23–0.36)	0 (0–0)*	4.93 (3.31–6.89)

* Aunque la mediana el rango intercuartil es 0, el rango del error es (0; 0.24)

con el punto fiducial 2. La mediana del SDE de SPP-PP_{DFI} es ligeramente superior al PP-SPP que ha sido mostrado previamente en la tabla 4.5. Además, el rendimiento del algoritmo para distintas frecuencias de muestreo (PP-PP_{DFI}) es bastante similar. Además, el error asociado a la pérdida de información debido a una baja frecuencia de muestreo (PP-PP_{DIF}) puede ser despreciado. El error de RR-PP_{DFI} simula la situación de un fotopleximógrafo validado con la misma frecuencia que el smartphone (RR-SPP) y aísla las diferencias asociadas a ambos sensores.

4.4.7.3 Desviación de estándar del error entre RR-SPP, RR-PP y PP-SPP: influencia del smartphone y de la respiración

La tabla 4.7 muestra la SDE de las series temporales en diferentes condiciones respiratorias y para ambos modelos de smartphone. Se han encontrado diferencias significativas en el SDE del RR-PP para el S5 ente RL y FR=15 ($p < 0.05$). Por otro lado, las diferencias del SDE SPP-PP son significativas para el S5 entre RL y FR=6 ($p < 0.05$). En cuanto a las influencias del smartphone, se han encontrado diferencias significativas para FR=15 entre el S5 y el MX ($p < 0.05$). Aunque estas diferencias son significativas,

Tabla 4.7: SDE de RR-SPP, RR-PP y PP-SPP por respiración y modelos de smartphone

	S5			MX		
	RL (n=22)	FR=6 (n=23)	FR=15 (n=23)	RL (n=18)	FR=6 (n=19)	FR=15 (n=19)
RR-SPP	5.82 (4.27-7)	5.16 (3.98- 6.11)	5.42 (3.87- 7.54)	5.26 (3.55- 7.84)	5.82 (4.21-7)	6.02 (4.25- 8.73)
RR-PP	4.94 (3.4- 6.55)*	5.21 (3.77- 6.58)	5.25 (3.74- 7.71)	4.41 (3.18- 7.55)	4.73 (2.8- 6.84)	4.9 (3.33- 7.14)
PP-SPP	3.15 (2.75- 4.62)	4.17 (3.01- 4.88)	3.35 (2.82- 4.35)	3.68 (3.12- 4.37)	4.3 (3.55- 6.93)	4.72 (3.48- 5.74)
	†		‡			

FR es frecuencia de respiración en respiraciones por minuto y RL significa respiración libre. Los resultados se expresan en mediana (Q25 %-Q75 % IQR) en ms. RL-FR15 $p < 0.05^*$; RL-FR6 $p < 0.05^\dagger$ y MX-S5 $p < 0.05^\ddagger$

son pequeñas y van desde 0.5 ms hasta 1 ms aproximadamente. Por tanto, podemos decir que los resultados en términos del SDE son muy parecidos para ambos modelos de smartphone en las tres condiciones de respiración.

4.4.7.4 Acuerdo entre SPP-RR, PP-RR y SPP-PP de los índices de VFC e influencia del modelo del smartphone

La tabla 4.8 muestra los resultados del análisis de los LA de los índices de VFC entre las series SPP, PP y RR. Para evaluar la influencia del smartphone en los índices de VFC, se han seleccionado las medidas que están en ambos modelos de smartphone para todas las combinaciones de sujeto y tipo de respiración, así que hemos descartado 14 muestras (1 MX y 13 S5), quedando un total de 110 medidas. Se encontraron diferencias significativas asociadas al smartphone en el SDNN, RMSSD y LF/HF. Los LA del 95 % del NN y SDNN para el S5 son menores que para el MX pero los del RMSSD para el S5 son mayores que para el MX. Por otro lado, se debe resaltar que se encontró buen acuerdo para el NN, SDNN y RMSSD del SPP-RR pero un acuerdo insuficiente para el LF/HF.

En la figura 4.14 podemos ver los gráficos de dispersión de la media de la diferencia y los LA de los índices de VFC para SPP-RR y PP-RR. Esta figura muestra que los datos están dispersos uniformemente a largo del eje horizontal para NN, SDNN y RMSSD

Tabla 4.8: Sesgo y límites de acuerdo Bland-Altman de los índices VFC entre SPP, RR y PP para cada modelo de smartphone

	SPP-RR		PP-RR		SPP-PP	
	S5 (n=55)	MX (n=55)	S5	MX	S5	MX
NN (ms)	0.06 ± 1.64	0.3 ± 2.72	-0.37 ± 1.83	0.07 ± 3.11	0.43 ± 1.88	0.23 ± 3.44
SDNN (ms)	1.19 ± 2.75	1.68 ± 2.79‡	0.97 ± 3.07	0.74 ± 4.42	0.22 ± 1.79	0.93 ± 4.11
RMSSD (ms)	1.43 ± 4.07	2.91 ± 3.56‡	0.81 ± 5.87	1.2 ± 5.01	0.62 ± 4.37	1.7 ± 4.33‡
LF/HF (n.u)	-0.7 ± 4.1	-1.3 ± 5.62‡‡	-0.49 ± 4.53	-0.45 ± 3.21	-0.2 ± 2.14	-0.85 ± 4.29‡‡
BA ratio						
NN (%)	0.18	0.29	0.2	0.34	0.2	0.37
SDNN (%)	3.87	3.86	4.33	6.16	2.5	5.66
RMSSD (%)	7.15	6.15	10.36	8.79	7.62	7.42
LF/HF (%)	52.35	66.43	57.16	36.07	28.25	52

Sesgo ± 95 % intervalo de confianza de los límites de acuerdo. P-valores de las diferencias estadísticas de las medias: p<0.05‡ and p<0.001‡‡ Mack-Skilling

pero hay una tendencia de las diferencias negativa y proporcional a la magnitud de la medida en LF/HF tanto para el S5 y como para el MX.

4.4.8 Discusión

La desviación estándar del error de la series SPP y PP han sido obtenidas para cada punto fiducial. El SDE más bajo ha sido obtenido para el FP2 pero las diferencias son pequeñas, alrededor de 1 ms. Para este punto fiducial, el error cometido en la estimación de la serie temporal RR usando el smartphone (RR-SPP=5.4 ms) es sólo un 10 % mayor que usando un fotopleletismógrafo (RR-PP=4.9 ms). Como era de esperar, debido a que PP y SPP se obtienen a partir de la señal de la onda de pulso, la mediana del SDE PP-SPP (3.95 ms) es menor que la del RR-PP que es menor que la del SPP-RR.

Las fuentes posibles de mayor error en el smartphone que en el fotopleletismógrafo son las diferentes frecuencias de muestreo y los dispositivos empleados. En cuanto a lo primero, los resultados obtenidos cuando la señal de pulso se muestrea a baja frecuencia (30 Hz) y está libre de ruido (PPG), muestran que si se le aplica una interpolación cúbica, la pérdida de información debido a esta frecuencia es despreciable. Por lo tanto, la señal SPPG (30 Hz) debe ser filtrada para reducir el ruido antes de ser interpolada a 5 kHz, Además, se ha evaluado el error cometido cuando el algoritmo empleado es muestreado a baja frecuencia (30 Hz) y la señal es libre de ruido (PPG) y es bajo (0.28 ms).

En cuanto al dispositivo empleado, la diferencia entre ellos (SPP-PP) es alrededor de 4 ms. Este no está solo asociado al smartphone puesto que de otro modo el incremento en el error ente el RR-SPP y RR-PP sería mayor. Entonces, esta diferencia puede estar relacionada por la manera en que se realiza la medida. De ahí, que no todo el error

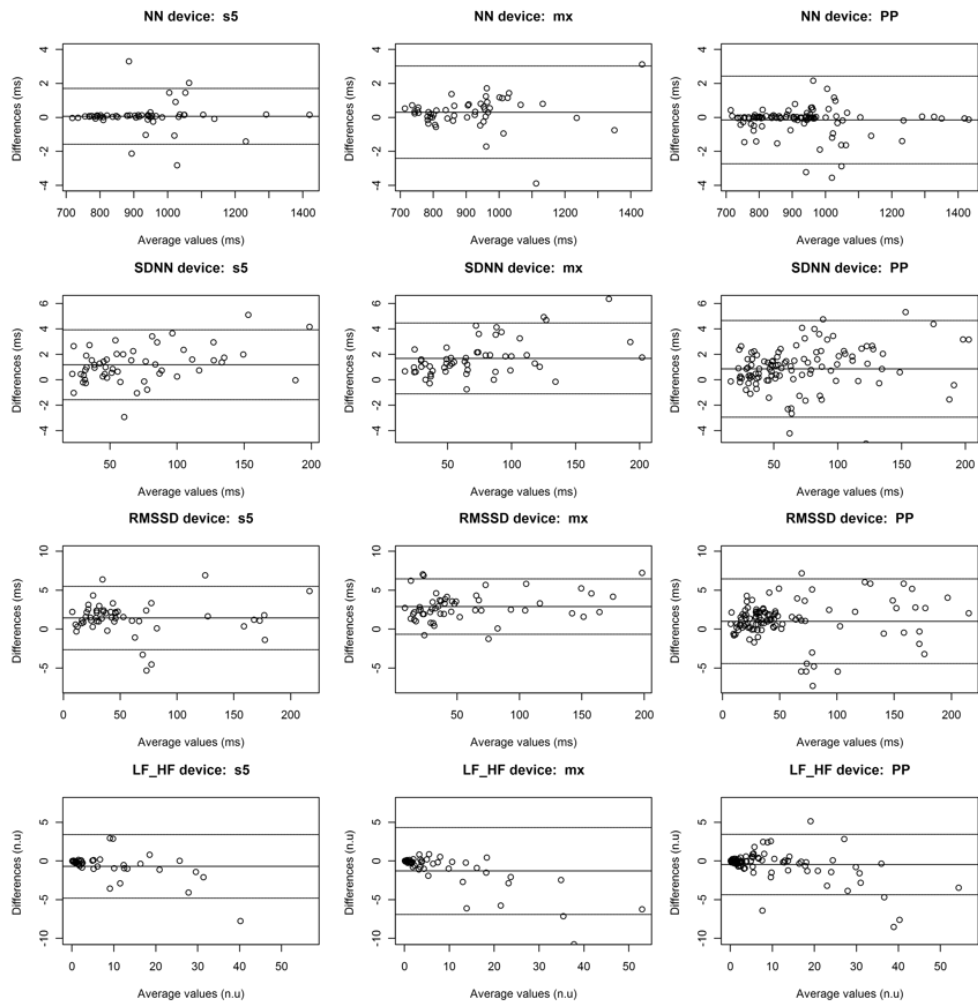


Figura 4.14: Sesgo y límites de acuerdo del Bland-Altman de los índices VFC. La primera columna muestra los resultados para SPP_{S5-RR} , la segunda para SPP_{MX-RR} y la tercera para $PP-RR$

en el PPG (RR-PP) es debido a la variabilidad del PAT. Por lo tanto, el error en el smartphone (RR-SPP) no sólo está asociado con el propio dispositivo sino también está relacionado con el PAT y como la medida ha sido realizada.

La SDE para diferentes modelos de smartphone y ritmos respiratorios ha sido evaluada para estudiar su influencia en las medidas de SPPG y PPG. Se ha encontrado que las diferencias en el SDE pueden ser significativas entre algunos ritmos respiratorios pero son pequeñas y van desde 0.5 ms a 1 ms. La SDE de ambos modelos de smartphone es similar.

Los índices de VFC de las series SPP, PP y RR han sido comparados. El análisis Bland-Altman realizado muestra no sólo la exactitud sino también la precisión. Los LA de los índices VFC cuando se comparan SPP-RR son un poco menos exactos pero más precisos que los PP-RR, esto puede ser debido a que el PP fue medido en el dedo medio mientras que el SPP fue medido en el dedo índice. El ratio BA muestra buen acuerdo para el NN,

SDNN y RMSSD pero acuerdo insuficiente para el LF/HF en las tres comparaciones: SPP-RR, PP-RR y el SPP-PP. El LF/HF del PP y SPP es menor en comparación con el LF/HF obtenido del RR. Esta disminución puede estar asociada a los componentes de la arritmia sinusal respiratoria. Estas componentes (actividad de alta frecuencia) pueden ser más pronunciadas en la señal del PPG y SPPG y por eso el LF/HF del PP y SPP puede decrecer [187].

El BA muestra que para el SPP-RR, el sesgo y la precisión del NN, SDNN, RMSSD y LF/HF son bastante similares para el S5 y el MX.

Los LA y el ratio BA de los índices VFC para el SPP-RR son mejores que los reportados previamente en un estudio similar [84]. Una posible explicación para estas diferencias puede ser que la frecuencia de muestreo de la cámara del smartphone del mencionado estudio no era estable y oscilaba entre los 20 y 30 Hz mientras que en nuestro estudio estaba fijada a 30 Hz.

4.4.9 Conclusiones

El análisis de las medidas realizadas demuestran que el smartphone es un instrumento de medida válido para la medida de algunos parámetros de la VFC. Se han evaluado las diferencias entre el PPG basado en la cámara del smartphone, ECG y PPG de referencia. El error entre las series SPP y RR se ha minimizado con el punto fiducial en el máximo de la primera derivada (FP2). El error debido a que la frecuencia de muestreo de la cámara del smartphone es 30 Hz ha sido compensado mediante la interpolación empleada. El SDE obtenido entre el SPP y RR está alrededor de 5.4 ms, es similar al obtenido entre el PP y RR. Además, se han obtenido SDE similares para ambos modelos de smartphone y los tres patrones respiratorios. Por otro lado, se han obtenido los LA entre los métodos: smartphone, ECG y PPG. Se han obtenido LA similares para el SPP-RR y PP-RR para los índices de VFC. Se ha encontrado buen acuerdo del SPP-RR para NN, SDNN y RMSSD pero insuficiente acuerdo del LF/HF. Finalmente, la diferencias del SPP-RR entre los modelos de smartphone de los LA son pequeñas. Por lo tanto, el smartphone puede ser utilizado para medir NN, SDNN y RMSSD de forma precisa y las diferencias entre modelos de smartphone son pequeñas.

Capítulo 5

Análisis conjunto de la actividad física y ritmo cardíaco mediante smartphones

5.1 Introducción

Durante los últimos años se ha registrado un envejecimiento progresivo de la población en Europa. Además, se estima que este envejecimiento va en aumento, en concreto se calcula que el porcentaje de gente mayor de 60 años pasará de un 14 % en 2010 a un 22 % en 2050 [1]. De ahí que, a distintos niveles, organismos oficiales estén promocionando la adherencia a hábitos de vida saludables en este colectivo, entre ellos la práctica de actividad física. Según varios estudios, la práctica de ejercicio físico está asociada a distintos beneficios para la salud: menor incidencia de enfermedades cardiovasculares, mayor vitalidad y mejor estado general de salud, menor riesgo de caída y menor incidencia de enfermedades mentales.

En nuestro caso, el objetivo es integrar en una única aplicación la recogida y el procesamiento de la actividad física y la realización de la medida del pulso cardíaco utilizando únicamente los sensores del smartphone. A la hora de diseñar la interfaz de usuario se ha tenido en cuenta que esta aplicación iba a ser utilizada por gente mayor.

En la literatura podemos encontrar algunos trabajos acerca de la medida simultánea de movimiento corporal y ritmo cardíaco [44, 188, 189]. Estos estudios están enfocados a la medida simultánea con dos aparatos para cuantificar la cantidad de actividad a partir del uso de acelerómetros y ritmo cardíaco. En nuestro caso se trata de integrar en una única aplicación la medición continuada de la actividad física y las medidas puntuales de ritmo cardíaco. Todo ello con la finalidad de que los datos que se recogen sirvan para

poder evaluar cómo afecta la cantidad de actividad física realizada al ritmo cardíaco y a los índices de VFC derivados.

Existen evidencias que indican que la interacción social está vinculada a la salud y al bienestar en las diferentes etapas de la vida. En concreto, la presencia de vínculos sociales está asociada con menores niveles de estrés y una menor mortalidad [190, 191]. Además, el hecho de estar apoyado por un red social está asociado con un aumento de la satisfacción de la vida, autoestima y felicidad [192]. La interacción social puede servir de apoyo cuando éste es necesario, por ejemplo en circunstancias de estrés pueden reducir el malestar. Asimismo, la interacción social facilita el acompañamiento, lo que permite participar en actividades placenteras de forma compartida, de esta manera se incrementa el sentimiento positivo [193].

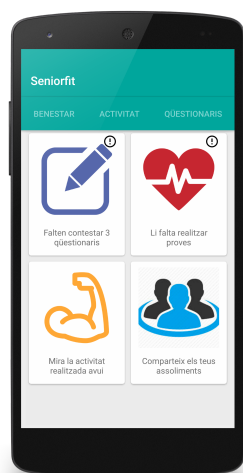
La aparición de las redes sociales virtuales ha permitido conectar a millones de personas de distintos puntos del planeta a través de Internet. Aunque existen algunos estudios acerca de las consecuencias para la salud del uso de estas redes sociales, no existe un consenso en cuanto a los perjuicios o beneficios que tienen para la salud [194]. En este sistema se ha diseñado una red social para realzar la interacción de la comunidad de usuarios y que la aplicación sirva como complemento a los vínculos sociales ya existentes.

5.2 Integración de ambos sistemas en una única aplicación

Se ha desarrollado una única aplicación para evaluar estos dos factores: actividad física y VFC. Esta aplicación se ha desarrollado para ser utilizada en un proyecto de mayor envergadura, en el que se pretende evaluar la adherencia a hábitos de vida saludables mediante el análisis de variables psicológicas y la interacción social de forma conjunta con la actividad física y el VFC. Por ello en esta aplicación también se trata el estado psicológico y se ha diseñado una red social para evaluar la interacción social. Como se puede observar en la figura 5.1(a), la aplicación consta de cuatro partes, cada una enfocada a cada uno de los aspectos de la vida saludable mencionados. Esta división clarifica la navegación durante el uso de la aplicación. En las siguientes subsecciones se explican en detalle cada una de las funcionalidades.

5.2.1 Cuestionarios psicológicos

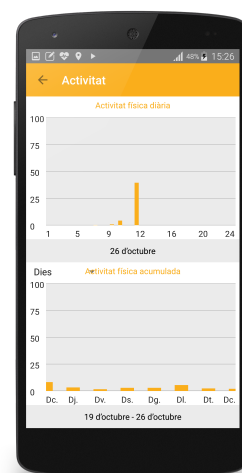
En la aplicación se han implementado una serie de cuestionarios que el propio usuario ha de responder para evaluar distintos aspectos psicológicos. Se han implementado tres tipos de cuestionarios que se deben responder con una periodicidad determinada. Por un lado, para la evaluación del estado de ánimo la aplicación implementa una versión reducida a 6 ítems del cuestionario del Perfil de Estados de Ánimo (PEA) [195], que es una adaptación en castellano del POMS *Profile of Mood States* que tiene 65 ítems



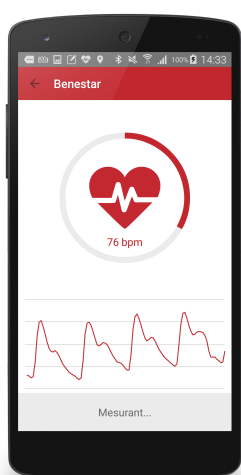
(a) Panel de navegación principal



(b) Cuestionario PEA



(c) Datos de actividad física



(d) Medida del ritmo cardíaco

Figura 5.1: Diseño final de las diferentes partes de la aplicación

[196]. Este cuestionario se debe responder diariamente y utilizar esta versión reducida permite reducir considerablemente el tiempo de respuesta por parte del usuario. Este cuestionario proviene del ámbito de la psicología clínica y es ampliamente utilizado para medir los estados emocionales que presentan las personas en diferentes situaciones. Este cuestionario se muestra en la figura 5.1(b).

Además, en la aplicación se ha diseñado un cuestionario sobre la percepción subjetiva de la salud basado en el SF-12 que consta de 12 preguntas [197] y el cuestionario SCOFF sobre hábitos alimentarios que se deben responder mensualmente [198]. Se ha programado el lanzamiento de notificaciones por parte de la aplicación para recordar al usuario

que se deben rellenar los cuestionarios.

5.2.2 Ritmo cardíaco

En la parte dedicada al ritmo cardíaco, la aplicación permite realizar la medida del pulso cardíaco mediante fotopleletismografía utilizando la cámara del smartphone y consultar un histórico de las medidas realizadas. Estas medidas se han realizado con la ayuda de una carcasa que facilita la correcta ubicación del dedo sobre la cámara. Además, antes de empezar a medir el ritmo cardíaco se le pide al usuario que indique su estado de ánimo en un rango de 6 elementos de serio a alegre. Durante la medida, el usuario puede ver la señal de pulso y el ritmo cardíaco tal y como muestra la figura 5.1(d). Al finalizar la medida, a partir de los datos registrados, se calcula, se almacena y se muestra al usuario, el ritmo medio y un índice de bienestar asociado a esta medida que se explicará más adelante. La aplicación permite visualizar estos datos promediados por días, semanas y meses y ver la evolución a lo largo del tiempo. Además, en este caso mediante notificaciones la aplicación también recuerda al usuario que debe realizar dos medidas de pulso cardíaco diariamente en distintas franjas horarias.

5.2.3 Actividad física

La aplicación recoge datos de acelerometría, GPS y presión atmosférica adquiridos por los sensores internos del móvil de forma continua. Los datos de acelerometría son procesados por la aplicación a medida que se van recogiendo mediante un algoritmo para calcular un índice de actividad en función del movimiento del smartphone. Este algoritmo ha sido diseñado para que el smartphone pueda ser llevado tanto en el bolsillo como en una bolsa. La figura 5.1(c) muestra el aspecto de la visualización de los datos de actividad, la aplicación muestra la actividad realizada por horas del día actual y un histórico de la actividad acumulada por días, semanas o meses.

5.2.4 Interacción social

Se ha diseñado un red social en la aplicación para favorecer el contacto entre todos los usuarios y que puedan ayudarse y animarse entre ellos. Por un lado, la red social permite colgar anuncios que pueden ver todos los usuarios. En la red social también se van publicando los logros de los usuarios en función de la evolución de sus indicadores de actividad física y bienestar. Por otro lado, la red social ofrece un chat para que los usuarios puedan conversar. Además se pueden añadir comentarios y “me gusta” a los chats y anuncios [199, 200].

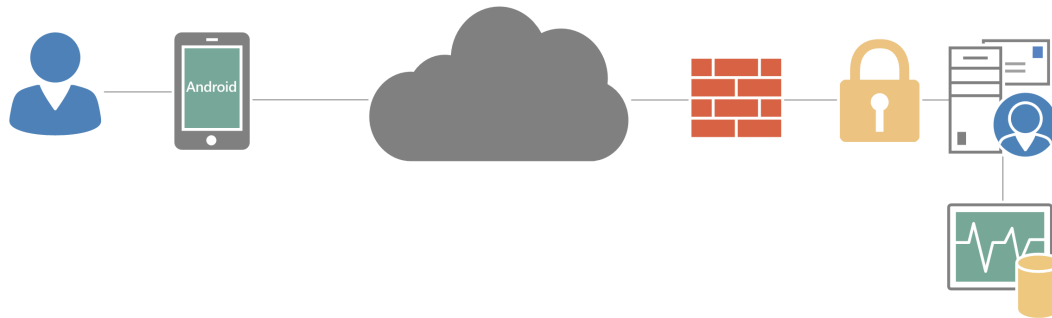


Figura 5.2: Diagrama de la arquitectura del sistema

5.3 Implementación de la aplicación

5.3.1 Arquitectura del sistema

La figura 5.2 muestra un diagrama de la arquitectura del sistema. La arquitectura implementada está basada en el modelo cliente-servidor. En nuestro caso el cliente es la aplicación móvil. Esta arquitectura nos permite enviar los datos adquiridos por la aplicación móvil a un servidor para almacenar los datos de múltiples usuarios de forma estructurada y poder ser analizados posteriormente. La aplicación se ha desarrollado para el sistema operativo de Android en JAVA (v. 1.7) utilizando el Android SDK (v.23). Esta aplicación es compatible con las versiones 5.x y 6.x de Android y para el desarrollo de la parte de conexión con el servidor hemos usado la librería *Retrofit 2* y para las gráficas la librería *GraphView*. La transmisión de los datos se realiza de forma segura realizando la comunicación por SSL certificado por terceros. Todos los datos que se envían están identificados mediante un número identificador de cada usuario para anonimizarlos.

Por otro lado, el servidor se encarga de recoger y almacenar de forma estructurada toda la información que va recogiendo la aplicación móvil. De esta forma los datos pueden ser analizados y visualizados por un tercero para monitorizar y evaluar los resultados tan pronto como se van obteniendo. Se ha montando un servidor Apache Tomcat, se ha desarrollado un servicio web tipo REST en Java-EE utilizando la librería Spark. La base de datos se ha diseñado en MySQL.

Además la red social está alojada en el mismo servidor. Para implementarla hemos utilizado el motor de redes sociales “Elgg” (v. 2.0.3). La aplicación móvil envía los datos utilizando la red wifi para evitar el consumo de datos móviles.

5.3.2 Índices utilizados

A partir de los datos de acelerometría del smartphone se estima el índice de actividad física cada minuto. Cada eje del acelerómetro (x,y,z) es interpolado linealmente a 20

Hz para obtener las tres señales muestreadas a frecuencia constante, se ha escogido esta frecuencia porque es suficiente para analizar los movimientos de la persona a nivel de actividad física. A continuación se calcula la media cuadrática de los tres ejes y la señal resultante se filtra paso alto a 0.76 Hz con un filtro Butterworth de orden 2 para eliminar la componente gravitatoria. Finalmente, el índice de actividad es el cociente entre el número de muestras que superan el umbral de 0.88 m/s^2 y 1200 (número de muestras en 1 minuto). Tanto la recogida de datos como el procesado se realizan en un hilo secundario de la aplicación de forma transparente al usuario.

El pulso cardíaco se obtiene en tiempo real y de forma no invasiva mediante fotople-tismografía, utilizando la cámara trasera y el flash LED del smartphone. La aplicación procesa en tiempo real las imágenes de la cámara sobre la cual el sujeto coloca la yema del dedo índice. A partir del promediado del canal verde de cada fotograma se obtiene la señal de la onda de pulso (PPG) a una frecuencia de muestreo de 30 Hz. Entonces, se extrae la serie temporal RR del intervalo de tiempo entre latidos a partir del tiempo transcurrido entre dos puntos fiduciales consecutivos del PPG. La serie PP obtenida se utiliza para calcular el ritmo cardíaco medio y el índice de bienestar asociado al ritmo cardíaco. El índice de bienestar propuesto se define como:

$$100 * \sqrt[3]{\min(1, \frac{SDSD}{PPSD}) * \min(1, \frac{avPP}{1000}) * \min(1, 20 * \frac{PPSD}{avPP})} \quad (5.1)$$

donde $SDSD$ es la desviación estándar de la diferencia entre intervalos PP consecutivos en milisegundos, $PPSD$ es la desviación estándar de todos los intervalos PP en milisegundos y $avPP$ es la duración media de todos los intervalos PP en milisegundos. Previamente, se detectan y se corrigen los artefactos de la serie PP utilizando el test de Grubbs [201].

Teniendo en cuenta que las medidas se realizan de forma libre y sin supervisión, se ha introducido un paso previo de comprobación de la correcta colocación del dedo sobre la cámara. Esta comprobación está basada en el análisis de los valores RGB y de los pulsos detectados para comprobar que el dedo está correctamente puesto sobre la cámara. Cada medida tiene una duración de 90 s y la aplicación también ofrece la visualización de la onda de pulso durante la medida.

5.4 Medidas realizadas

Las medidas se han realizado con un grupo de once personas mayores, la media de edad fue de 67.2 años y en un rango de edades de 63 a 76 años. La distribución por sexos fue ocho mujeres y tres hombres. Cinco personas tenían un nivel de estudios primario, tres tenían un nivel asimilable a la actual educación secundaria obligatoria, dos habían terminado el bachillerato y uno poseía estudios superiores. Nueve voluntarios

presentaban problemas visuales y dos de ellos problemas auditivos, seis tenían problemas para dormir y dos de ellos declararon problemas de adormecimiento durante el día. Uno expresó sufrir dolores de cabeza, uno fatiga y uno temblores. Ninguno manifestó epilepsia o infecciones importantes en el pasado. Tres voluntarios declararon sufrir alguna vez mareos o vértigos (un voluntario incluso perdió recientemente la consciencia). Por último, tres voluntarios toman medicación para la hipertensión, uno es diabético y sólo tres de entre los once no toman medicamentos rutinariamente.

Antes de comenzar la tanda de medidas, los voluntarios contestaron a un test cognitivo del que se desprendió que ninguno de ellos presenta problemas en este sentido. Además, en una entrevista previa se observó que en la rutina diaria, los voluntarios caminan entre 1 km y 5 km (media: 2.59 km, sd: 1.46 km), cinco de ellos realizan alguna otra actividad física semanal y sólo tres voluntarios consideran que realizan suficiente actividad física.

Antes de empezar las medidas, se les impartió una breve sesión explicativa acerca del funcionamiento de la aplicación. Las medidas consistieron en llevar encima el smartphone durante el día para que fuera recogiendo datos sobre actividad física. Además, consistieron en realizar diariamente dos medidas de ritmo cardíaco, una a primera hora de la mañana y otra a última hora de la tarde. La duración de cada medida fue de 90 segundos. Todas las pruebas se hicieron con el mismo modelo de smartphone, un Samsung Galaxy S5 que les proporcionamos a los participantes durante las pruebas. Las pruebas se realizaron en dos turnos debido a que disponíamos de seis smartphones únicamente. El primer turno duró 3 semanas y el segundo 6 semanas. Una vez finalizadas las pruebas, se procedió a la recogida de los smartphones y los voluntarios respondieron a un cuestionario de satisfacción sobre la aplicación.

Los participantes dieron su consentimiento informado para la realización de las pruebas de forma escrita. Los datos de uno de los sujetos fueron descartados porque no contestó los cuestionarios PEA.

5.5 Análisis de los datos

El PEA evalúa el estado de ánimo global a partir de diferentes factores que representan estados afectivos o emocionales: tensión, depresión, hostilidad, vigor y fatiga que constituyen medidas independientes del estado de ánimo:

- Tensión (T): las puntuaciones altas indican niveles altos de tensión y de ansiedad
- Depresión (D): las puntuaciones altas indican estados de ánimo, depresivos y de infravaloración
- Hostilidad (H): las puntuaciones altas indican antipatía hacia los otros, incluso cólera

- Vigor (V): las puntuaciones altas indican un estado de ánimo de actividad y de predisposición para la acción.
- Fatiga (F): las puntuaciones altas indican cansancio, fatiga y flojedad.
- Confusión (C): las puntuaciones altas indican desconcierto y confusión.

Para los factores T, D, H, F y C las puntuaciones altas indican una mayor alteración del estado de ánimo, ya que las puntuaciones elevadas en sus ítems se relacionan con aspectos negativos. Por el contrario, el factor V es el único que se interpreta como un estado de ánimo positivo cuanto mayor sea su puntuación. Los sujetos valoraron cada factor con un valor entre cero y diez. Hemos estimado una puntuación final de cada cuestionario mediante la media de todos los ítems, excepto el V que previamente fue transformado restando a diez el cuantificador V.

Por un lado, se ha calculado para cada sujeto el ajuste lineal de la media actividad, del índice de bienestar y del PEA con el tiempo. Obteniendo el p-valor y la pendiente de la recta para cada ajuste. De esta manera se pretende obtener un indicador de la tendencia a lo largo del tiempo del índice analizado y si esa tendencia es significativa.

Por otro lado, hemos realizado el análisis de regresión para estudiar la relación entre las variables de actividad física, PEA y ritmo cardíaco. Se ha investigado si existe una asociación entre dos variables analizando los p-valores en el análisis de la regresión lineal. El p-valor para cada relación comprueba la hipótesis nula de que el coeficiente es igual a cero (no hay relación). Un p-valor bajo (<0.05) indica que se puede rechazar la hipótesis nula, indicando que es probable que exista relación entre ambas variables.

5.6 Resultados

En la tabla 5.1 se muestra la media y la desviación estándar de la actividad diaria. La tabla muestra la pendiente de la regresión entre la actividad acumulada y el tiempo para cada uno de los sujetos. Se han calculado los p-valores usando la t-test de Student para estimar la significancia de la regresión.

Observando que los p-valores de la regresión son significativos y que la pendiente de la curva es negativa, podemos apreciar que la actividad física decae a lo largo del tiempo en todos los usuarios, aunque en diferentes proporciones.

En cuanto a la regresiones entre distintas variables, se ha encontrado que la desviación estándar de la actividad desciende con la pendiente de la actividad significativamente ($p < 0.05$), esto nos muestra que los más inconstantes son los que más reducen la actividad con el tiempo.

Tabla 5.1: Resultados del índice de actividad

Sujeto	Promedio	Desviación estándar	p-valor (lineal)	pendiente (lineal)
1	0.144	0.1202	0	-1.69E-07
2	0.2465	0.1124	0	-1.10E-07
3	0.2577	0.0863	0	-1.13E-07
4	0.2499	0.1728	0	-1.58E-07
5	0.4447	0.0967	0	-7.76E-08
6	0.3153	0.1018	0	-6.27E-08
7	0.3274	0.0799	0	-2.57E-08
8	0.3448	0.0518	0	-1.37E-08
9	0.2971	0.0955	0	-3.15E-08
10	0.1224	0.0737	0	-1.10E-08

Tabla 5.2: Resultados del índice de bienestar

Sujeto	Valor medio	Desviación estándar	p-valor (lineal)	pendiente (lineal)
1	54,7471	10,007	0,1521	-1,42E-05
2	64,4852	13,5003	0,4951	-2,00E-05
3	52,3361	12,4762	0,2485	1,81E-05
4	78,1344	10,7141	0,6479	-4,10E-06
5	51,7649	11,4658	0,0242	-8,74E-06
6	64,0871	10,7196	0,1471	1,45E-04
7	59,4156	13,0021	0,008	-2,58E-05
9	53,3095	15,0583	0,1016	-3,97E-05
10	64,4361	14,771	0,0311	-4,33E-05

Respecto a las medidas de ritmo cardíaco, la tabla 5.2 muestra los valores medios y desviación estándar del indicador de bienestar para cada sujeto. También muestra la pendiente y el p-valor de la regresión del índice de bienestar con el tiempo.

Se ha descartado el sujeto 8 porque no había realizado las medidas de ritmo cardíaco. Se ha encontrado que en los sujetos 5,7 y 10 el indicador de bienestar tiene tendencia a descender a lo largo del tiempo.

No se han encontrado relaciones entre los indicadores de actividad y el ritmo cardíaco. No obstante, haría falta medir un número mayor de sujetos para poder afirmar la existencia o no de este tipo de relaciones.

En la tabla 5.3 podemos encontrar los resultados para el valor medio y desviación estándar del PEA, p-valores y pendiente para el ajuste lineal de los valores del PEA. Para los sujetos 1,4,7,9 y 10 se encuentra una tendencia lineal significativa. En todos los sujetos en los que se encontró un cambio significativo en el estado de ánimo, este cambio ha sido una mejora del estado de ánimo, pendiente negativa.

Además, se han obtenido los p-valores y la pendiente de recta de regresión de las regresiones de los índices de actividad y PEA. En cuanto a los resultados significativos

Tabla 5.3: Resultados del Perfil de Estado de Ánimo

Sujeto	Valor medio	Desviación estándar	p-valor (lineal)	Pendiente (lineal)
1	3.6136	0.7566	9.14E-05	-1.00E-06
2	3.5556	0.832	0.2516	-2.94E-07
3	4.3258	0.7246	0.1224	4.38E-07
4	1.1818	0.388	3.59E-04	-4.80E-07
5	2.8784	0.6043	0.5889	-5.94E-08
6	4.3636	0.4977	0.7384	6.73E-08
7	1.8241	0.527	0.0412	-1.98E-07
8	1.2222	0.5008	0.8372	-1.96E-08
9	2.5303	0.3513	8.36E-04	-4.12E-07
10	2.6019	0.5244	0.007	-2.55E-07

Tabla 5.4: Regresión del índice de actividad y el PEA

Sujeto	p-valor (lineal) Act y PEA	pendiente(lineal) Act y PEA
1	3,28E-04	4,0201
2	0,6056	0,9263
3	0,1706	-2,435
4	0,2186	0,5677
5	0,2605	1,0962
6	0,1618	-1,64
7	0,9389	-0,0933
8	0,9477	-0,1315
9	0,8335	-0,1679
10	0,6291	-0,5957

tenemos que la asociación del valor medio de actividad y p-valor del PEA es significativa ($p < 0.05$) y con pendiente positiva. Esto nos indica que aquellos que son más activos son menos propensos a cambios en el estado de ánimo.

Los resultados del SF-12 y SCOFF no se han analizado porque debido a que se responden mensualmente y las pruebas duraron entre 3 y 6 semanas los datos disponibles son muy pocos para poder analizarlos.

La tabla 5.4 muestra los p-valores de las regresiones del índice de actividad y el PEA. No se ha encontrado una relación directa entre los cambios en el índice de actividad y PEA. Sin embargo, existe asociación positiva entre la pendiente del índice de actividad y los p-valores de la regresión del índice de actividad y PEA, esto nos indica que aquellos con mayores reducciones de la actividad física, muestran cambios más significativos en el estado de ánimo asociado a la actividad física. Además, se ha encontrado una asociación negativa significativa entre la pendiente del PEA y la pendiente de la regresión de la actividad física y el PEA. Esto nos indica que las mejoras en el PEA están asociadas a una mayor influencia de la actividad en este cambio.

Tabla 5.5: Resultados del cuestionario de satisfacción

Preguntas	Mediana	Máximo	Mínimo
La medida de ritmo cardíaco es difícil de realizar	1	1	1
He comprendido el resultado de la medida de ritmo cardíaco	5	5	1
La visualización de los resultados de ritmo cardíaco es correcta	5	5	2
La visualización de los resultados de actividad física es correcta	5	5	1
Los resultados de actividad física cotejan con el ejercicio que he realizado	5	5	3
El uso de la aplicación ha influido en mi rutina diaria	1	5	1
Los cuestionarios son difíciles de responder	1	5	1
Creo que la aplicación es útil	5	5	3

Finalmente, respecto al cuestionario de satisfacción, los usuarios valoraron muy bien la aplicación. La tabla 5.5 muestra los resultados a las afirmaciones propuestas donde 1 debe entenderse como total desacuerdo mientras que 5 significa total acuerdo. De estos resultados se desprende que la dificultad de la medida y de los cuestionarios es baja, la comprensión de la medida del ritmo cardíaco es alta y la visualización de los datos de ritmo cardíaco y de actividad es correcta.

Aunque se ha diseñado una red social y se ha implementado un sistema de logros para ser publicados en la red social diseñada, finalmente, la red social no ha sido utilizada por las personas que participaron en los ensayos porque la red social requiere del uso de Internet y los smartphones no disponían de tarjeta SIM para conectarse a Internet. El buen funcionamiento de la arquitectura cliente-servidor para el volcado de datos sí ha sido probado, puesto que al finalizar las pruebas todos los smartphone se conectaron a una red wifi y los datos se volcaron al servidor de forma correcta.

Finalmente, se analizó el rendimiento de la aplicación en cuanto al consumo de batería, la aplicación consumía un 17% de batería al día, teniendo en cuenta que el Samsung S5 tiene unos 2.800 mAh, el consumo es de un total de unos 475 mAh. A nivel de memoria RAM utilizó un máximo de 20 MB y el espacio de almacenamiento ocupado durante 3 semanas fue de 527 MB. Por tanto podemos decir que esta aplicación puede ser utilizada con el funcionamiento normal del smartphone.

5.7 Conclusiones

El análisis conjunto de la actividad física y el ritmo cardíaco se ha llevado a cabo utilizando únicamente los sensores internos del smartphone. Se ha diseñado y desarrollado una aplicación para evaluar la adherencia a un estilo de vida saludable mediante la

cuantificación de la actividad física a partir de la señal del acelerómetro y mediante la evaluación indicadores de bienestar a partir del ritmo cardíaco utilizando la cámara del smartphone. Además, la aplicación permite la evaluación subjetiva del estado de ánimo mediante cuestionarios.

Un grupo de 11 personas mayores han utilizado la aplicación diseñada para realizar medidas de actividad física, ritmo cardíaco y del estado de ánimo de forma conjunta durante un tiempo de entre 3 y 6 semanas. Los datos de las medidas se han recogido de forma adecuada.

A partir de los resultados recogidos se han analizado los indicadores de actividad, ritmo cardíaco y estado de ánimo, se ha encontrado que el nivel de actividad física desciende a lo largo del tiempo y se ha encontrado alguna asociación entre la constancia en la práctica de la actividad física y los cambios en el estado de ánimo. Sin embargo, estas asociaciones se han de tomar con precaución debido al reducido número de sujetos que han participado en este estudio.

Además se ha evaluado mediante un cuestionario el nivel de usabilidad y la experiencia del usuario. Se ha encontrado que el grado de satisfacción con la aplicación es alto y los usuarios han indicado que ha resultado sencilla de utilizar.

Finalmente, el sistema desarrollado supone un punto de partida para evaluar la adherencia a un estilo de vida activo de forma unificada con una única aplicación. Es necesario ampliar el estudio a más personas para realizar un análisis más profundo, comparando los índices de ritmo cardíaco, actividad física y los resultados de los cuestionarios psicológicos.

Capítulo 6

Detección temprana de somnolencia en conductores

6.1 Motivación

El ritmo circadiano es el reloj biológico que regula las funciones vitales durante el día y la noche. La alteración del ritmo circadiano o la desincronización de éste con el día y la noche es una de las principales causas de somnolencia diurna. Además, llevar un estilo de vida poco saludable: vida sedentaria, estrés, etc., puede provocar alteraciones en los ritmos circadianos que desencadenen en alteraciones del sueño. Los trabajadores a turnos son uno de los colectivos que más desórdenes del sueño sufren debido al desajuste de los ritmos circadianos [202]. Entre estos colectivos están los enfermeros [203], operarios de cadenas de montaje, etc. La somnolencia es particularmente pronunciada durante el turno de noche y puede llevar a incidentes como el quedarse dormido durante el trabajo [204].

Un estilo de vida poco saludable puede estar asociado con alteraciones en el sueño lo que puede provocar somnolencia diurna [39]. La obesidad es uno de los mayores factores de riesgo para la incidencia y la cronicidad de la excesiva somnolencia diurna (ESD), mientras que la pérdida de peso está también asociada con su remisión [205]. La obesidad y el ESD están asociados con la incidencia de depresión para ambos sexos y el sobrepeso también está asociado con la depresión en mujeres [206].

Otro factor de riesgo de somnolencia al volante es la apnea obstructiva del sueño (OSA). Las personas que padecen este trastorno del sueño tienen entre el doble y el triple de riesgo de quedarse dormidas mientras conducen [207]. La OSA puede ser tratada mediante una máquina de presión positiva en las vías aéreas. Además, en personas obesas la severidad de la OSA puede ser reducida mediante la pérdida de peso [208]. También la práctica de ejercicio físico tiene beneficios para las personas que padecen

OSA, porque se reduce la somnolencia diurna y se aumentan la eficiencia del sueño y el consumo de oxígeno máximo [209].

Esta tesis se enfoca en los conductores de coche, camión o autocar. Estos colectivos tienen una jornada laboral que se caracteriza por ser extensa y en ocasiones el trabajo es monótono y una pequeña distracción puede poner en grave peligro su vida y la de otras personas. Por eso se ha validado un sistema de detección temprana de somnolencia a partir del análisis de la variabilidad del período respiratorio. Se pretende que este sistema pueda ser de utilidad tanto para colectivos de trabajadores a turnos como para personas que fruto de una vida poco saludable tengan tendencia a la somnolencia diurna.

Aunque este capítulo está enfocado a la población de conductores profesionales, las conclusiones del mismo se pueden hacer extensivas a la población general. En este capítulo se propone un algoritmo para la detección de los eventos de lucha por no dormirse a partir del análisis de la señal respiratoria. También se describe la validación del algoritmo a partir de la realización de pruebas en un entorno controlado.

6.2 Métodos

6.2.1 Participantes

20 voluntarios adultos (10 mujeres y 10 hombres) de edades comprendidas entre los 20 y 60 años participaron en el estudio. 6 personas fueron clasificadas según el índice de masa corporal (IMC) como delgadas ($<18.5 \text{ kg/m}^2$), 7 de peso normal ($18.5 \text{ a } 25 \text{ kg/m}^2$) y 7 con sobrepeso ($25 \text{ a } 30 \text{ kg/m}^2$). Todos los participantes estaban sanos y no presentaban en su historial ningún trastorno de sueño. Además, todos estaban en posesión del carnet de conducir y no bebieron ninguna bebida alcohólica ni energéticas en las 6 horas previas a la prueba. Este estudio fue llevado a cabo de acuerdo con los principios de la Declaración de Helsinki [177] y todos los participantes recibieron información detallada del estudio y dieron su consentimiento informado. Los participantes realizaron dos pruebas, cada una en días distintos. Del total de las pruebas, cuatro no se grabaron adecuadamente así que un total de 36 pruebas fueron consideradas para este estudio.

6.2.2 Sistema de medida

Para evaluar la exactitud del algoritmo de detección de somnolencia propuesto, es necesario disponer de un conjunto de datos en los que los voluntarios del estudio deben conducir mientras están luchando por no dormirse. Debido a que esta situación puede ser peligrosa para los voluntarios, los test fueron llevados a cabo en un simulador de conducción bajo condiciones controladas. Además las condiciones del test han sido diseñadas para observar este comportamiento de lucha por no dormirse mientras se conduce.



(a) Simulador de conducción: pantalla frontal y chasis del coche (b) El sujeto llevaba tres bandas RIP alrededor del abdomen, diafragma y pecho

Figura 6.1: Setup de medida para detectar somnolencia del conductor en las instalaciones de IBV

Por un lado, el protocolo del ensayo está enfocado en detectar somnolencia al volante en condiciones difíciles para mantenerse alerta: horas nocturnas, carreteras aburridas y poco transitadas. El sistema de simulación consiste en una pantalla frontal con un chasis de un coche para ofrecer una experiencia inmersiva en las instalaciones del Instituto de Biomecánica de Valencia (IBV), este entorno se muestra en la figura 6.1(a). Además, una cámara de vídeo enfocada a la cara del sujeto grabó el ensayo. Estas grabaciones fueron usadas por los observadores externos para generar la señal de control de somnolencia.

Por otro lado, en la respiración están involucrados varios músculos que actúan en la inspiración y la espiración. La respiración es llevada a cabo primeramente por el diafragma, un músculo que separa la cavidad torácica de la cavidad abdominal. En concreto, la contracción y relajación del diafragma produce cambios de volumen en las cavidades torácica y abdominal. Por eso, la señal respiratoria puede ser obtenida mediante el registro de los desplazamientos del diafragma, abdomen y la cavidad torácica. La pletismografía inductiva respiratoria (RIP) es un método no invasivo ampliamente utilizado para medir la respiración [210]. Además, estudios previos han demostrado que RIP puede ser también utilizado para cuantificar la variabilidad de la respiración ciclo a ciclo [211]. Como podemos ver en la figura 6.1(b), en nuestro sistema de medida, el sujeto lleva puesto tres bandas RIP ubicadas en el tórax, diafragma y abdomen para garantizar la mejor calidad de señal sin tener en cuenta las características antropométricas. Las señales respiratorias han sido adquiridas por tres canales sincronizados del monitor Bitmed EximPro (Bitmed, Sibel S.A., Barcelona, Spain) a una frecuencia de muestreo de 40 Hz, la cual satisface el criterio de Nyquist para señales respiratorias.

6.2.3 Protocolo de ensayo

Cada participante realizó dos ensayos, cada uno en días diferentes. En uno de ellos, el sujeto tuvo privación de sueño durante la última noche así que se pidió a los sujetos que no durmieran en las 24 horas previas al ensayo. El otro ensayo se realizó bajo una situación normal de sueño, en la que el sujeto debía haber dormido al menos 6 horas la noche anterior al ensayo. Entonces, la diferencia entre ambos ensayos fue la condición inicial de sueño del sujeto. Además, se pidió a los sujetos que permanecieran sentados en el asiento del simulador de conducción, llevando puesto el cinturón de seguridad y manteniendo ambas manos sobre el volante y los pies en los pedales. Durante los primeros 5 minutos del ensayo, se pidió a los sujetos que permanecieran quietos y en silencio. Después de esto, se pusieron a conducir durante una hora y media. Al final, se les pidió que permanecieran sentados en simulador con los ojos cerrados durante 5 minutos.

6.2.4 Observadores externos

Se pidió a observadores externos previamente entrenados que clasificaran el estado de alerta del sujeto como “somnoliento” o “no somnoliento” minuto a minuto a partir de las grabaciones faciales de vídeo para generar la señal de control. Para esta clasificación, se ha considerado que el estado somnoliento está caracterizado por: pérdida de expresión facial, bostezos y parpadeo lento. De lo contrario, el estado de alerta del sujeto se clasifica como “no somnoliento”. Además, se les pidió que clasificaran minuto a minuto la calidad de la señal de respiración como “buena” o “mala” a partir de la inspección visual de la señal respiratoria.

A partir de las clasificaciones minuto a minuto del estado de alerta del sujeto, es esperable que una correcta detección de somnolencia esté ubicada en el mismo minuto que un evento de somnolencia en la señal de control. Sin embargo, en experimentos previos se ha encontrado que las clasificaciones de lucha por no dormirse pueden tener algunas limitaciones [212]. La lucha por no dormirse es un fenómeno que incluye varios eventos como suspiros y parpadeos, que pueden ser claramente identificados a través de grabaciones faciales de vídeo por un observador externo. Sin embargo, es complicado para el observador externo clasificar cuando aparecen los primeros signos de somnolencia, tales como cortos periodos de largos parpadeos, algunos bostezos y cambios en la posición del cuerpo. Aunque, los estados de somnolencia pueden ser identificados claramente, el punto de inicio a veces no parece claro. Por ello, se ha definido una región llamada tiempo de avance (AT) para reducir la incertidumbre en las clasificaciones de somnolencia. Los estados clasificados dentro de AT minutos antes del primer minuto clasificado como somnoliento, permanecen fuera de los resultados de rendimiento del algoritmo de clasificación de somnolencia para reducir la incertidumbre asociada a la clasificaciones de los observadores.

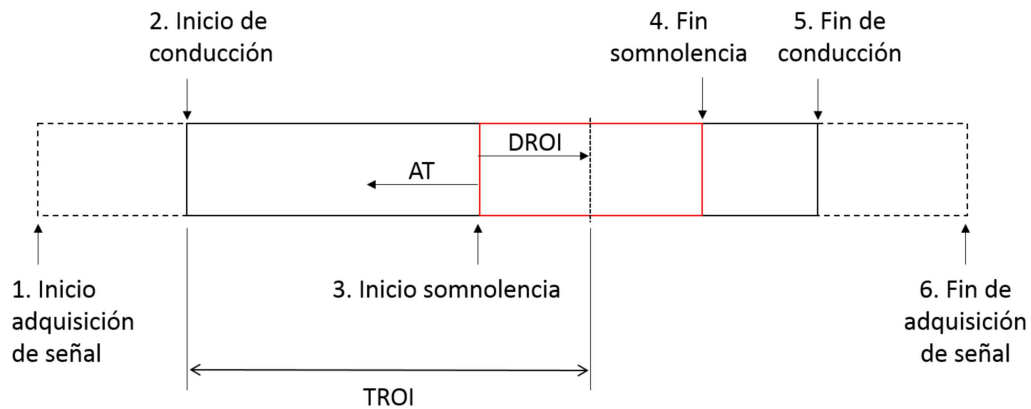


Figura 6.2: Cronograma de los eventos registrados sobre la medida y la región de tiempo de interés (TROI) definida

6.2.5 Región de tiempo de interés (TROI)

Para analizar el rendimiento del algoritmo propuesto algunos eventos deben ser registrados durante el ensayo. Estos eventos se definen como: evento 1 es el inicio de adquisición de la señal, el evento 2 ocurre cuando el sujeto empieza a conducir, el evento 3 ocurre cuando el sujeto empieza a luchar por no dormirse, el evento 4 tiene lugar cuando el sujeto se despierta, el evento 5 ocurre cuando el sujeto deja de conducir y el evento 6 es el final de la adquisición de señal. Además, AT es el espacio de tiempo anterior a la ocurrencia del evento 3. La región de interés de somnolencia (DROI) es el intervalo de tiempo dentro de cada ensayo del sujeto que ha sido identificado para detectar somnolencia. La DROI abarca desde el primer minuto en que los observadores clasifican al sujeto como somnoliento hasta 5 minutos más tarde.

La figura 6.2 muestra el cronograma de eventos en el ensayo, debemos resaltar que estos eventos no tienen por que ocurrir en el mismo momento para todos los ensayos. Así que los eventos han sido marcados para cada sujeto por los observadores externos a partir de la grabaciones en vídeo. Aunque, el tiempo de los eventos 2 y 5 están fijados por el protocolo, estos tiempos pueden variar de uno a dos minutos entre sujetos, pero los eventos 3 y 4 dependen del estado de los sujetos en cada test.

Se ha definido un punto de inicio y final como el que limita el espacio de tiempo de señal respiratoria que será procesada por el algoritmo diseñado. Se ha establecido el evento 2 como el punto de inicio porque el algoritmo está diseñado para ser usado en condiciones de conducción. El punto final ha sido establecido como el final de la DROI. Se ha considerado que el estado de alerta del sujeto ya no es apto para conducir después de que hayan pasado cinco minutos desde que hayan aparecido los primeros síntomas de somnolencia. Entonces, la DROI ha sido definida para analizar la señal hasta un tiempo asumible después de que el sujeto empieza a dormirse, así la DROI ha sido establecida

Tabla 6.1: Cantidad de datos disponibles para ser analizados según los criterios de inclusión

Criterio de inclusión	Número de tests seleccionados ^a	Datos desde 2 hasta 5 ^b			Datos dentro de TROI ^b		
		DE min	SO min	SO %	DE min	SO min	SO %
N/A	36	2246	1035	31.6	2015	103	4.7
Alguna alerta de somnolencia	21	872	1035	54.3	424	103	19.5
Alerta de somnolencia después de los primeros 10 minutos	15	768	592	43.5	404	73	15.3

^a Número de tests según los diferentes criterios de presencia de somnolencia.

^b Cantidad de datos de señal respiratoria para cada estado de alerta para las diferentes regiones de tiempo de test. DE significa despierto y SO significa somnoliento.

a 5 minutos para todos los sujetos. El intervalo de tiempo total seleccionado para cada test se ha llamado región de tiempo de interés (TROI).

El análisis de la señal respiratoria dentro de la TROI proporciona resultados más fiables que el análisis durante todo el test porque el principal objetivo del algoritmo diseñado es la detección del primer tramo de tiempo donde aparecen signos de somnolencia. Se han obtenido las diferentes TROI para cada sujeto a partir de las clasificaciones de los observadores externos. Hemos incluido en el análisis las medidas en las que hay presencia de somnolencia y la somnolencia aparece más tarde de los 10 primeros minutos desde que se empieza a conducir. Se han escogido 10 minutos porque la inicialización del algoritmo dura 5 minutos y la duración de AT, que ha sido establecida al mismo valor que la DROI, debe ser añadida para obtener resultados más fiables. La tabla 6.1 muestra un resumen de la base de datos: el número de participantes, la cantidad de minutos clasificados como somnoliento y despierto para las diferentes condiciones y TROI. Finalmente, cabe destacar que el porcentaje de tiempo en el que los sujetos estaban somnolientos muestra que se obtienen resultados más realistas a partir del análisis durante la TROI que desde el evento 2 al 5. Por lo tanto, el algoritmo propuesto ha sido optimizado y evaluado a partir 15 diferentes tests y 477 minutos de señal respiratoria.

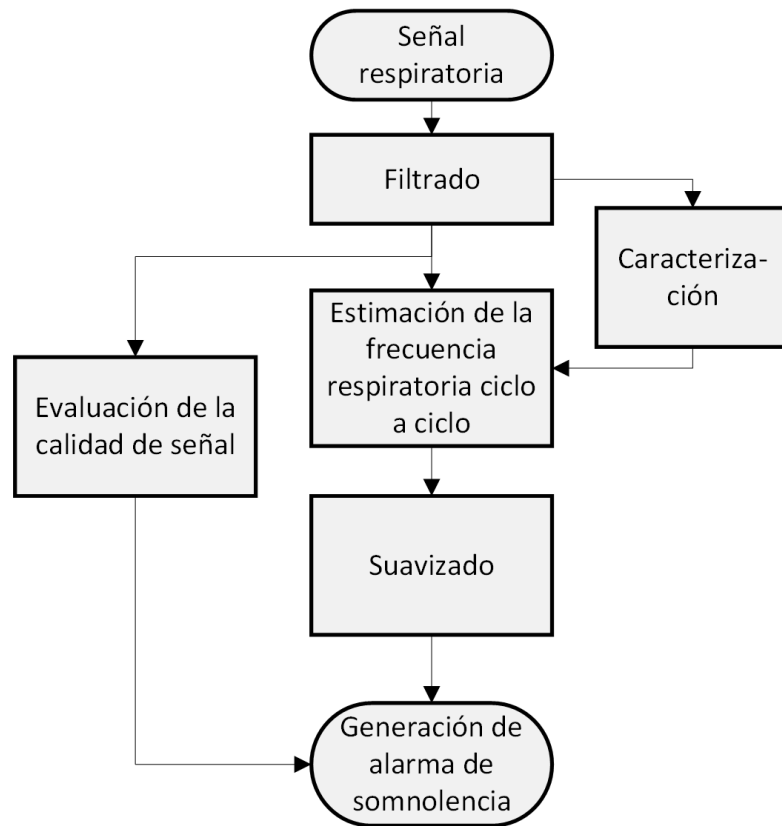


Figura 6.3: Bloques principales del algoritmo propuesto de detección de somnolencia a partir del esfuerzo torácico (TEDD)

6.3 Algoritmo propuesto

En esta sección describiremos el algoritmo para detección de somnolencia en tiempo real a partir del índice de detección de somnolencia derivado del esfuerzo torácico (TEDD). Cabe destacar que el algoritmo TEDD propuesto ha sido diseñado para analizar no solo señales de esfuerzo torácico, sino también señales de esfuerzo diafragmático y abdominal. Principalmente, el algoritmo está basado en el análisis de la variabilidad del ritmo respiratorio (RRV) a lo largo del tiempo, así como la presencia de artefactos en la señal respiratoria. La figura 6.3 muestra los principales bloques del algoritmo propuesto. Los diferentes pasos del algoritmo de detección de somnolencia y de la evaluación de la calidad de la señal respiratoria se explican más adelante.

El algoritmo propuesto para detectar somnolencia mientras se conduce está basado en la detección de la lucha por no dormirse a partir del análisis de la RRV. Debido a que la señal respiratoria se adquiere con una banda inductiva alrededor del cuerpo del conductor mientras conduce, pueden aparecer artefactos en la señal respiratoria debido a movimientos del cuerpo, a la tos y al habla. La amplitud de la señal respiratoria puede variar entre sujetos y a lo largo del tiempo, por eso el algoritmo diseñado está basado

en variaciones en frecuencia y no en amplitud. Por ello, la señal se filtra y se comprime para mitigar estos efectos.

Durante los episodios de lucha por no dormirse, la RRV aumenta pero no de forma igual para todo el mundo. Entonces, el algoritmo busca para cada sujeto un nivel inicial de RRV, el cual se toma como referencia. El algoritmo busca un intervalo de señal respiratoria que se caracteriza por: estabilidad en frecuencia y amplitud, media del ritmo respiratorio dentro del rango de respiración normal y ausencia de artefactos causados por ronquidos y movimientos. El nivel de RRV del intervalo de señal seleccionado como referencia, será usado por los siguientes pasos del algoritmo para evaluar el incremento de RRV a lo largo del tiempo. A este proceso se le llama *caracterización*. La señal de referencia se utiliza para obtener un umbral con el que estimar el ritmo respiratorio.

Una vez obtenida la señal de referencia, el ritmo respiratorio es estimado ciclo a ciclo a partir del tiempo transcurrido entre dos cruces por el umbral de la señal respiratoria. El ritmo respiratorio ha sido promediado con una ventana deslizante porque puede estar mal estimado debido a varios tipos de artefactos no relacionados con la dinámica de la respiración. Este proceso se llama *estimación de la frecuencia respiratoria ciclo a ciclo*.

Se ha observado que la RRV aumenta mientras el conductor está luchando por no dormirse y que el estado de la alerta del conductor decrece lentamente. De ahí que el RRV haya sido suavizado y el algoritmo busca una secuencia de picos en la serie de RRV durante un periodo de tiempo para identificar los episodios de lucha. El método cuasi-pico descrito más abajo ha sido utilizado para este fin, este proceso ha sido llamado *suavizado y búsqueda del máximo*.

Uno de los mayores inconvenientes de la detección de somnolencia a partir del análisis de la RRV es que una elevada RRV está causada no sólo por la lucha por no dormirse sino también por movimientos del conductor, hablar o toser. Por lo tanto, el algoritmo evalúa la calidad de la señal respiratoria a partir de la presencia o ausencia de estos eventos y el índice de calidad ha sido combinado con la detección del cuasipico para evitar que los eventos que no están relacionados con episodios de somnolencia puedan confundir las RRV relacionadas con somnolencia.

Como ha sido mencionado anteriormente, el estado de alerta decrece lentamente mientras se conduce y el objetivo de este algoritmo es alertar cuando el sujeto no esté apto para conducir, por eso el algoritmo genera un indicador binario minuto a minuto. Se ha encontrado un umbral optimizado para convertir el nivel promediado en un minuto de la RRV en un indicador binario de detección de somnolencia. La exactitud del algoritmo ha sido evaluada a partir de la comparación del indicador binario con las clasificaciones de los observadores externos. Los diferentes bloques mencionados del algoritmo TEDD propuesto se describen con más profundidad a continuación y están ilustrados en la figura 6.4.

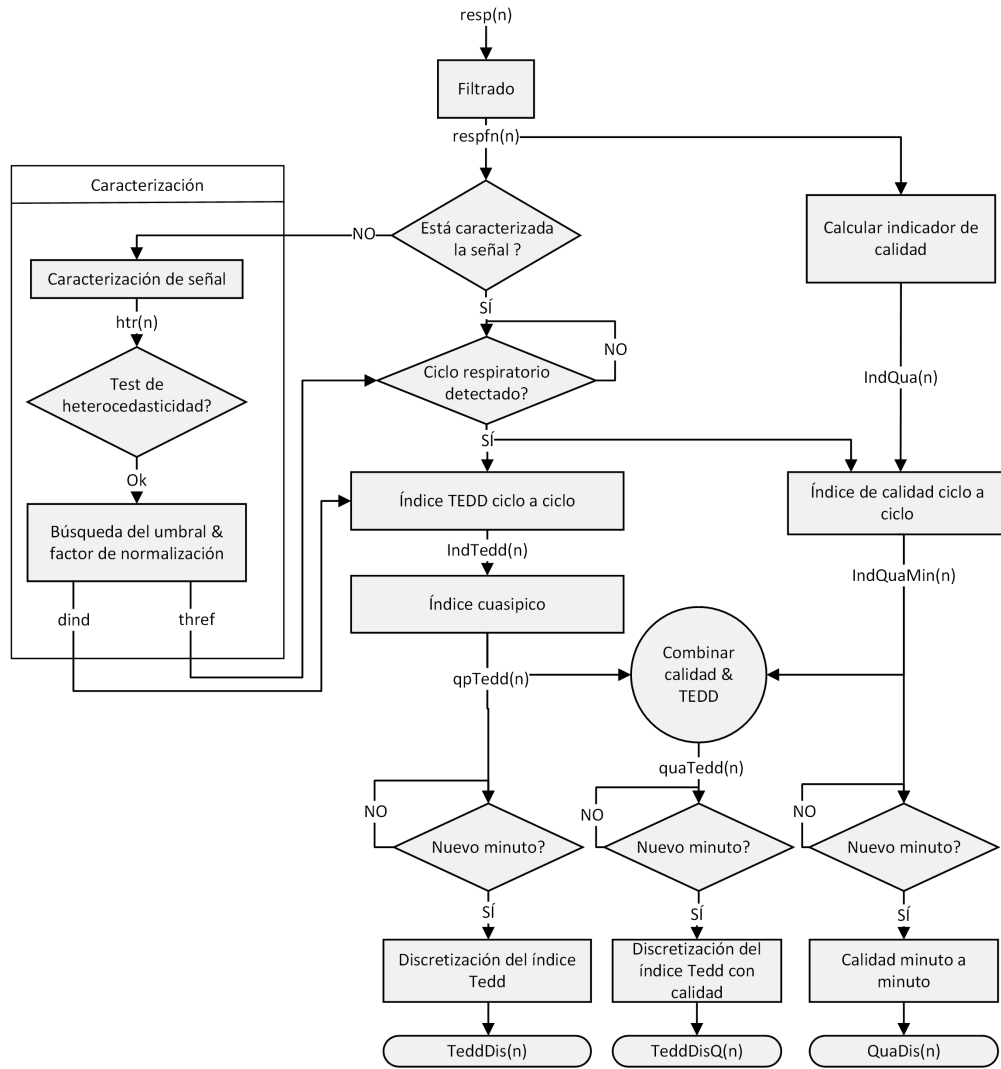
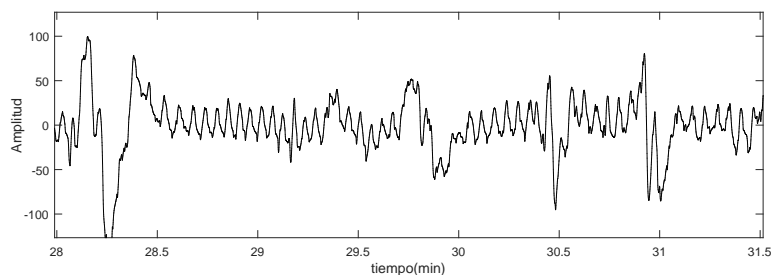


Figura 6.4: Diagrama de flujo detallado del algoritmo TEDD

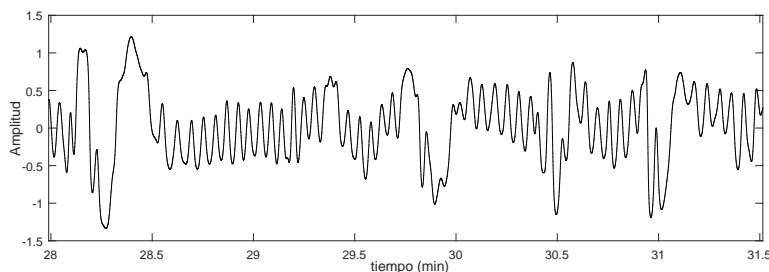
6.3.1 Filtrado de la señal respiratoria

La frecuencia respiratoria normal varía según la edad y el sexo. El ritmo respiratorio normal tiene un rango de 14 a 20 respiraciones por minuto en adultos [213]. El bloque de filtrado de la señal respiratoria consiste en un filtrado paso bajo a una frecuencia de corte de 0.5 Hz para reducir el ruido, seguido por un filtro paso alto a frecuencia de corte de 0.05 Hz para eliminar la continua. Ambos filtros han sido diseñados como un filtro Butterworth de cuarto orden. Finalmente, hemos aplicado una función de realzado no lineal para reducir artefactos basado en la arcotangente que se define como:

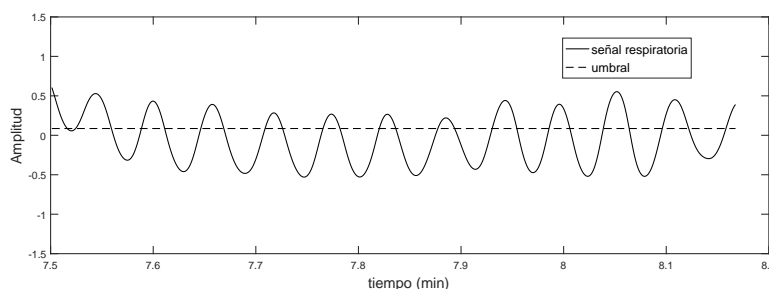
$$respfn(n) = \arctan \frac{respf(n)}{\sqrt{\frac{\sum_{k=1}^{300 \cdot f_s} (respf(k) - respfm)^2 \cdot \sqrt{2}}{300 \cdot f_s}}} \quad (6.1)$$



(a) Señal respiratoria en crudo



(b) Señal respiratoria filtrada



(c) Señal respiratoria de referencia y umbral obtenido

Figura 6.5: Ejemplo de una señal respiratoria

donde f_s es la frecuencia de muestreo, $resp_f$ es señal filtrada y $resp_fm$ es el promedio de $resp_f$ dentro de las $300 * f_s$ primeras muestras. La transformación no lineal normaliza la señal mediante la desviación estándar y calcula el arcotangente del cociente. En concreto, reduce la amplitud de los artefactos en la señal respiratoria. La figura 6.5, ilustra un ejemplo de señal respiratoria en crudo y la señal después de haber sido filtrada.

6.3.2 Caracterización de la variabilidad del ritmo respiratorio del sujeto

Este bloque se encarga de encontrar el patrón de referencia del estado inicial del sujeto. Esto nos permite tratar con las diferencias entre sujetos y con diferentes estados de alerta iniciales de los sujetos. Los parámetros constantes del algoritmo se han obtenido en pruebas previas [214]. El algoritmo busca una región estable como patrón de referencia durante un máximo de 5 minutos en la señal desde el inicio del registro. La estimación de

la duración de los ciclos respiratorios está basada en el tiempo transcurrido entre cruces por el umbral de la señal respiratoria, así que el patrón de referencia también se usará para establecer este umbral. La RRV de referencia también se obtiene a partir del patrón de referencia. La señal respiratoria se analiza en intervalos de 40 segundos de señal con una ventana deslizante y una estrategia de búsqueda del patrón basada en el análisis de la estacionaridad de la señal a partir del test de heteroscedasticidad. El tramo de señal más estacionario será seleccionado como patrón de referencia. Por un lado, el indicador de estacionaridad $mHtr$ ha sido estimado como:

$$\begin{aligned}
 htr(n) &= \frac{\sum_{j=1}^n (x(j) - \bar{x})^2}{\sum_{k=1}^N (x(k) - \bar{x})^2} \\
 mHtr &= \max_{k=1..N} htr(k)
 \end{aligned} \tag{6.2}$$

donde x es $respfn$, N es el número de muestras ($N = 40 * f_s$) y \bar{x} es la media de $respfn$ desde 1 hasta N . Por otro lado, se han utilizado dos índices más para encontrar el patrón de referencia: el ritmo respiratorio medio llamado mRR en Hz y la desviación estándar del ritmo respiratorio en segundos llamado $stdRR$. En este bloque, el ritmo respiratorio se calcula a partir de los cruces por cero de $respfn$. La señal seleccionada como referencia será el primer intervalo de 40 s de señal que cumpla las siguientes condiciones: $mHtr$ debe ser menor de 0.03, los valores de mRR deben estar dentro del rango de 0.04 Hz a 0.5 Hz y el $stdRR$ debe ser menor de 0.7 s. Si el índice $mHtr$ de todos los intervalos de tiempo es mayor que 0.03, se selecciona el intervalo de señal respiratoria con el menor $mHtr$ que cumple las condiciones mRR y $stdRR$ mencionadas. En experimentos previos, se ha comprobado que estas tres condiciones son adecuadas para obtener un patrón de referencia estable y cuasiperiódico.

Se ha utilizado el filtro de media móvil (MA), el cual se define como:

$$MA_N(x(k)) = \frac{1}{N} * \sum_{j=k-N}^k x(j) \tag{6.3}$$

donde x es el dato de entrada y N es la longitud del filtro MA.

Una vez que la señal de referencia es obtenida, se utiliza para obtener el umbral para la detección de ciclo respiratorio llamado $thref$. Se obtiene a partir del valor del percentil 60 de la señal respiratoria de referencia. Un ejemplo de señal respiratoria que se toma como referencia y el $thref$ obtenido se muestra en la figura 6.5. Además, a partir de este tramo de señal se calcula un índice de referencia de RRV mediante la siguiente ecuación:

$$dind = \frac{1}{M} \sum_{k=5}^M |MA_4(T(k)) - MA_4(T(k-1))| \tag{6.4}$$

donde $dind$ es el índice de referencia RRV, M es el número de ciclos de respiración detectados dentro del patrón de referencia y $T(k)$ es la duración de la respiración del k -ésimo ciclo de respiración detectado. k empieza en 5 para eliminar el transitorio inicial del MA de longitud 4. Además, $dind$ ha sido acotado inferiormente a un nivel basal ($dind=7$) para normalizar los patrones respiratorios estables y periódicos. El algoritmo guarda el $dind$ obtenido para normalizar el índice RRV que se obtiene a lo largo del registro.

6.3.3 Evaluación de la variabilidad del ritmo respiratorio ciclo a ciclo e indicador de somnolencia

La finalidad del algoritmo es obtener un indicador que responda a los cambios en el ritmo respiratorio asociado a la fatiga. Esta respuesta debe ser acumulativa, indicando un incremento o decremento del nivel de fatiga. El indicador propuesto, que está basado en la RRV, se describe como:

$$\begin{aligned} T_{ma}(n) &= MA_{WLBC}(T(n)) \\ IndTedd(n) &= \frac{MA_{WLD}(|T_{ma}(n) - T_{ma}(n-1)|)}{dind} \end{aligned} \quad (6.5)$$

donde $IndTedd$ es el indicador, $WLBC=4$ es el tamaño de la ventana del MA de los ciclos de respiración y WLD es el tamaño de la ventana del MA de la diferencia entre ciclos de respiración sucesivos. El valor del parámetro de WLD será encontrado mediante el método de optimización. $T(n)$ se obtiene a partir del intervalo de tiempo entre cruces positivos consecutivos por el $thref$ y el RRV se ha calculado a partir de la primera derivada del ritmo respiratorio estimado. Entonces, MA_{WLD} suaviza los cambios de RRV para obtener la tendencia del RRV y ha sido normalizado por el estado inicial de referencia.

Después de esto, se aplica un detector de cuasipico (QPD) al $IndTedd$ para estimar los cambios de RRV de forma acumulativa. QPD es un método de detección de picos empleado para medir emisiones conducidas y radiadas de equipos electrónicos en los test de compatibilidad electromagnética. El QPD tiene se caracteriza por un ataque rápido (tiempo de carga) y una respuesta de lento decaimiento (tiempo de descarga) [215], el cual es resultado de un filtro paso bajo de componentes R-C con diferente constante de tiempo para la subida y la bajada. El $qpTedd$ se obtiene como sigue:

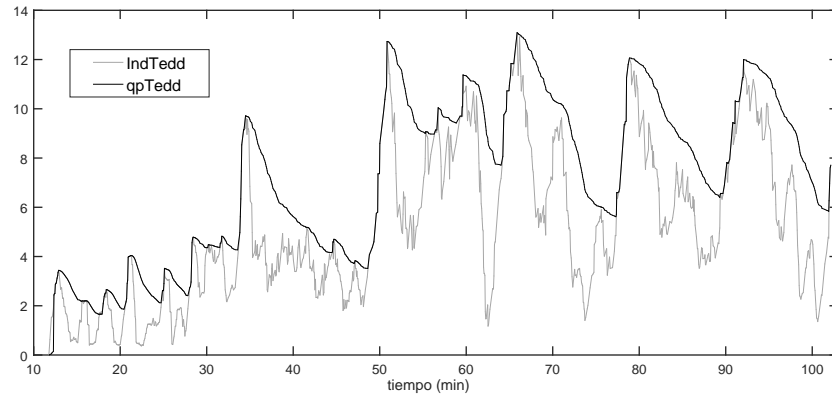


Figura 6.6: Comparación de los índices $IndTedd$ y $qpTedd$ que se han obtenido para un sujeto

$$qp(x_n) = \begin{cases} \alpha * x_n + (1 - \alpha) * qp_{n-1} & \text{para } x_n \geq qp_{n-1} \\ \beta * x_n + (1 - \beta) * qp_{n-1} & \text{sino} \end{cases} \quad (6.6)$$

$$qpTedd_n = qp(IndTedd_n)$$

donde el QPD diseñado tiene una tasa de aumento más rápida que la de decaimiento, $\alpha=1$ y $\beta=0.02$ respectivamente. Entonces, $qpTedd$ proporciona una indicación acumulativa de los cambios de RRV a lo largo del tiempo mediante la obtención de una mayor ponderación a los picos de $IndTedd$ más cercanos. La figura 6.6 ilustra el resultado de aplicar el método basado en QPD descrito sobre el $IndTedd$.

Finalmente, el $qpTedd$ fue discretizado mediante el promediado sobre intervalos de un minuto sin solapamiento. El índice discretizado se ha llamado $TeddDis$, el cual no solo tiene en cuenta la amplitud de los picos sino también la duración de los cambios de RRV. Se debe remarcar que los eventos de lucha por no dormirse son clasificados por varios observadores externos a partir de los vídeos de la cara del conductor y estos vídeos son generalmente clasificados minuto a minuto [216]. Por ello, el $TeddDis$ es un indicador de los cambios RRV para un intervalo de tiempo específico cuya resolución es la adecuada para ser comparado con las clasificaciones de los observadores externos. Por lo tanto, $TeddDis$ ha sido utilizado para clasificar el estado de alerta del sujeto basado en un umbral de decisión ($ThTedd$), el cual ha sido optimizado mediante la maximización del rendimiento de clasificación definido en la sección 6.4.1.

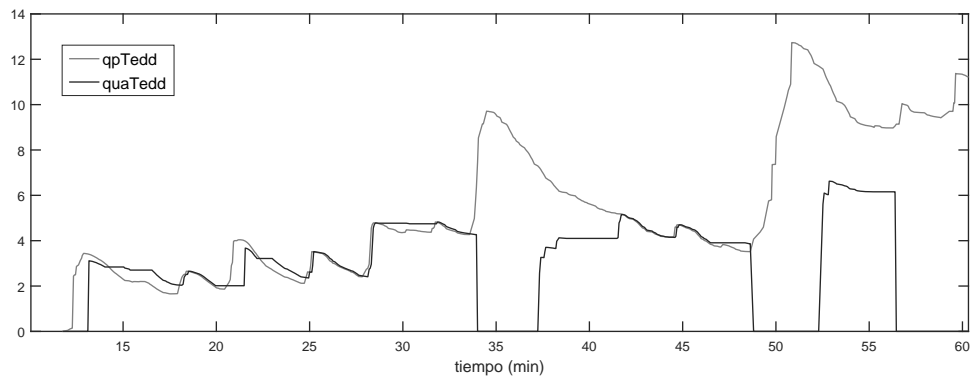
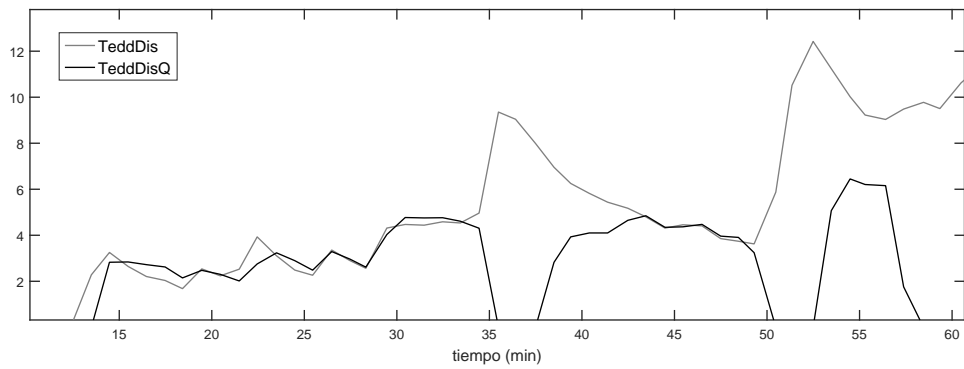
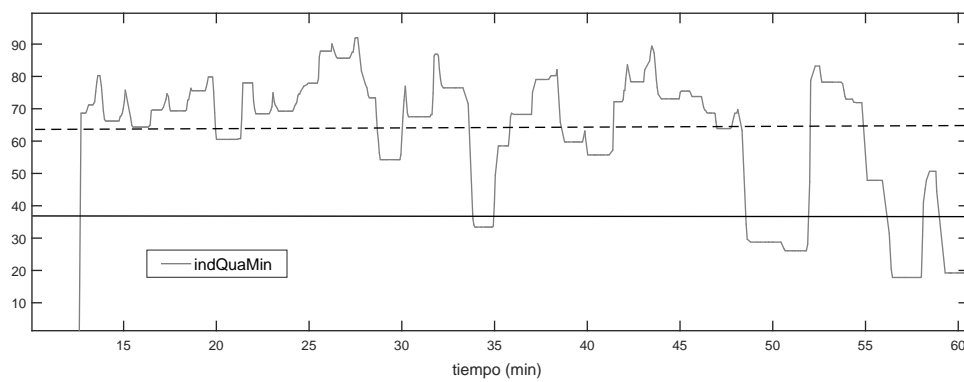
(a) Comparación del cuasipico del TEDD con y sin calidad: *quaTedd* y *qpTedd* respectivamente(b) Comparación del TEDD discretizado (*TeddDis*) y TEDD con calidad (*TeddDisQ*)(c) *IndQuaMin* a lo largo del tiempo, la línea discontinua es *QuaTh* y la línea sólida es *QuaTh/2*

Figura 6.7: Un ejemplo de los resultados del algoritmo TEDD combinado con la evaluación de la calidad para un sujeto

6.3.4 Algoritmo de evaluación de calidad de señal y combinación con el algoritmo de detección de somnolencia

El indicador propuesto para la evaluación de la calidad de la señal respiratoria está basado en el análisis de los cambios en la forma de onda respiratoria. El indicador de calidad se define como sigue:

$$qua(n) = \frac{MA_{WLR}(|x(n)|)}{\max_{k=n-WLR\dots n} x(k) - \min_{k=n-WLR\dots n} x(k)} \quad (6.7)$$

donde x es *respf_n* definido en la ecuación (6.1), k es la k -ésima muestra respiratoria y WLR es el tamaño de la ventana de MA en muestras. El mejor WLR ha sido encontrado mediante la optimización desde 20 a 60 segundos. Un índice normalizado *quan* ha sido definido para analizar los cambios de calidad sobre el tiempo como:

$$quan(n) = \left| \frac{qua(n)}{quaRef} - 1 \right| \quad (6.8)$$

donde *quaRef* es la media aritmética de *qua* sobre los primeros cinco minutos. Además, el *quan* ha sido filtrado para suavizar cambios repentinos y ha sido escalado. En concreto, puede ser expresado como:

$$IndQua(n) = 100 * \left(1 - \frac{MA_{WLQ}(quan(n))}{0,6} \right) \quad (6.9)$$

donde WLQ es el tamaño de la ventana del MA aplicado a *quan*. El mejor WLQ también ha sido buscado desde 20 a 60 segundos. Finalmente, la evaluación de la calidad de la señal respiratoria ha sido combinada con la evaluación de los cambios de RRV. Por ello, se han combinado en el bloque del detector de cuasipico para acumular los cambios en RRV de forma adecuada evitando estimaciones erróneas debida a baja calidad de la señal respiratoria. *quaTedd_n* denota la combinación de los índices TEDD y de calidad, que ha sido definida como:

$$IndQuaMin_n = \min_{k=n_{NBC}\dots n} IndQua(k)$$

$$quaTedd_n = \begin{cases} y_n & \text{si } x_n \geq QuaTh \\ y_{n-1} & \text{si } \frac{QuaTh}{2} \leq x_n < QuaTh \\ 0 & \text{sino} \end{cases} \quad (6.10)$$

donde x_n es *IndQuaMin_n*, y_n es *qpTedd_n* y n_{NBC} es NBC ciclos de respiración anteriores a n . *IndQua* ha sido minimizado tomando las muestras previas para tener un

Tabla 6.2: Métricas de rendimiento de clasificación utilizadas para el ajuste de parámetros y la evaluación del modelo

Métrica de rendimiento	Descripción
Verdadero positivo (VP)	# estados de somnolencia correctamente reconocidos
Verdadero negativo (VN)	# estados de no somnolencia correctamente reconocidos
Falso positivo (FP)	# estados de no somnolencia incorrectamente reconocidos
Falso negativo (FN)	# estados de somnolencia incorrectamente reconocidos
Sensibilidad (SEN)	$\frac{VP}{VP+FN}$
Especificidad (ESP)	$\frac{VN}{VN+FP}$
F1-score (F1)	$\frac{2*VP}{2*VP+FN+FP}$
G-media (G)	$\sqrt{SEN * ESP}$
Exactitud (E)	$\frac{VP+VN}{VP+VN+FP+FN}$
Kappa de Cohen (K)	$pa = \frac{(VN+FP)*(VN+FP)}{VP+VN+FN+FP}$
	$pb = \frac{(VP+FN)*(VP+FP)}{VP+VN+FN+FP}$
	$pe = \frac{pa+pb}{VP+VN+FN+FP}$
	$K = \frac{E-pe}{1-pe}$

indicador conservador porque patrones de señal respiratoria mala pueden aparecer en ciclos de respiratorios consecutivos. Como en la discretización de *qpTedd*, *quaTedd* ha sido discretizado minuto a minuto mediante el promediado sobre intervalos de un minuto no solapados. El índice discretizado ha sido llamado *TeddDisQ*. La figura 6.7 muestra el efecto de la combinación de la evaluación de la calidad de la señal con el *qpTedd* a lo largo del tiempo para un test de un sujeto.

6.4 Evaluación del rendimiento del algoritmo

6.4.1 Métricas de rendimiento

Las clasificaciones generadas por los observadores externos fueron usadas como referencia para evaluar la correctitud del algoritmo de detección de somnolencia propuesto. En concreto, se ha evaluado mediante el cálculo del número de estados de somnolencia correctamente reconocidos (verdaderos positivos), el número de estados de no somnolencia correctamente reconocidos (verdaderos negativos) y los estados que o fueron incorrectamente clasificados como somnolientos (falsos positivos) o bien no fueron reconocidos como somnolencia (falso negativo). Estas variables pueden ser representadas como una

matriz que se conoce como matriz de confusión [217]. Las métricas que han sido utilizadas para evaluar el rendimiento de clasificación del algoritmo TEDD se muestran en la tabla 6.2. La sensibilidad (SEN) y especificidad (ESP) son métricas ampliamente utilizadas conjuntamente para evaluar el rendimiento de algoritmos de clasificación binarios. Por un lado la SEN mide la proporción de positivos, en nuestro caso detección de somnolencia que son correctamente detectados por el algoritmo y la ESP mide la proporción de negativos (detección de no somnolencia) que son adecuadamente identificados. Por otro lado, se han utilizado también otras métricas tales como: F1-score (F1), G-media (G) y el coeficiente kappa de Cohen (K). G es insensible a los cambios en la distribución de las clases, por el contrario F1 se adecua bien para tareas con muchas instancias negativas. El coeficiente K mide el acuerdo entre clasificadores y ha sido utilizada la siguiente escala de valoración del grado de acuerdo: ≤ 0 =pobre, .01-.20=ligero, .21-.40=discreto, .41-.60=moderado, .61-.80=sustancial, y .81-1=casi perfecto [218]. Además, K ha sido empleado para medir el rendimiento del clasificador, porque ha sido medido el acuerdo entre los estados de alerta del sujeto predecidos y observados, corrigiendo el acuerdo que ocurre por suerte [219].

6.4.2 Optimización y validación del algoritmo

Los datos han sido divididos en conjunto de entrenamiento y conjunto de test para obtener una estimación no sesgada de la exactitud del algoritmo optimizado. Además, los datos de entrenamiento deben ser a su vez divididos en conjunto de entrenamiento y conjunto de validación para obtener una estimación no sesgada de la exactitud durante el proceso de optimización (p.e seleccionar el mejor tamaño de WLD).

Una de las principales limitaciones de realizar una única partición es que puede que no tengamos suficientes datos para hacer unos conjuntos de entrenamiento, validación y test suficientemente grandes. Por ello, se ha llevado a cabo un método de validación cruzada para usar los datos disponibles para validación y testeo de forma eficiente. El método de validación cruzada que consiste en dejar un sujeto fuera (LOSOCV) fue utilizado para elegir los mejores parámetros del modelo y para evaluar el rendimiento del clasificador de somnolencia seleccionado. En la técnica LOSOCV, los datos de todos los sujetos menos uno, fueron usados para entrenar el clasificador y los datos del sujeto restante fueron utilizados para test. Este proceso se repitió para todos los sujetos hasta que cada sujeto ha formado parte una vez del conjunto de test [156]. Por ello, este método de validación cruzada es adecuado para nuestro escenario porque no se desperdician los datos de ningún sujeto.

Los resultados de rendimiento pueden ser demasiado optimistas porque, los resultados de la validación cruzada estaban sujetos a la maximización. Por lo tanto, se ha llevado a cabo la validación cruzada en bucle anidado para combinar la selección del modelo y la evaluación del rendimiento. El método consiste en dos bucles de validación cruzada

- 1: Dividir el conjunto de datos en N partes
- 2: **for** $i = 1$ hasta N **do**
- 3: Usamos $N-1$ como conjunto de validación
- 4: Usamos el restante como conjunto de test
- 5: Dividimos el conjunto de validación en $N-1$
- 6: **for** $j = 1$ hasta $N - 1$ **do**
- 7: Usamos $N-2$ para entrenar el clasificador
- 8: Usamos el restante para evaluar el clasificador
- 9: **end for**
- 10: Seleccionamos el clasificador con mejor AUC
- 11: Construimos el clasificador seleccionado con el conjunto de validación
- 12: Almacenamos las predicciones en el conjunto de test
- 13: **end for**
- 14: Estimamos la media de las métricas de rendimiento en el conjunto de test

Figura 6.8: Algoritmo de LOSOCV anidada para selección y evaluación del modelo

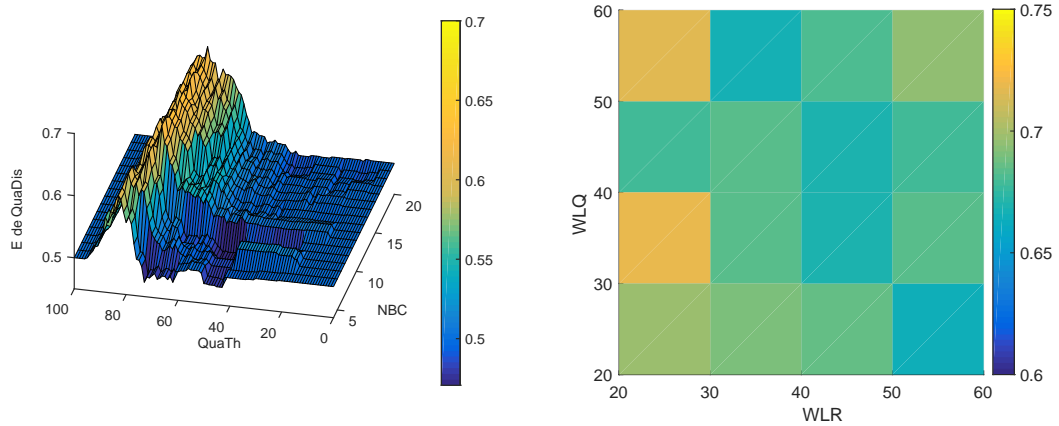
anidados: uno interno para seleccionar los parámetros del algoritmo con el mejor rendimiento en validación cruzada y otro externo para evaluar el modelo ajustado en el conjunto de test restante. Entonces, se obtienen los resultados del algoritmo de clasificación de somnolencia con los parámetros seleccionados para la estimar las métricas de rendimiento. El pseudocódigo del procedimiento de LOSOCV ha sido descrito en la figura 6.8.

6.5 Resultados y discusión

6.5.1 Algoritmo de evaluación de calidad de señal: optimización y evaluación del rendimiento

El algoritmo lleva a cabo una clasificación binaria de la calidad de la señal respiratoria como buena o mala, que está basada en el parámetro *QuaDis* que ha sido definido en la sección 6.3.4. Los resultados de clasificación fueron utilizados para optimizar los parámetros del algoritmo. Estos resultados han sido obtenidos mediante la comparación de las clasificaciones del algoritmo de evaluación de la calidad de la señal con la calidad de señal clasificadas por los observadores externos.

Los parámetros optimizados son: *WLR*, *WLQ*, *NBC* y *QuaTh*. Los valores de los rangos de los parámetros *WLR* y *WLQ*, se han buscado entre 20 y 60 segundos para abarcar varios períodos de respiración sin introducir mucho retraso en el algoritmo. Por ello, hemos evaluado los resultados de clasificación con diferentes valores de estos parámetros para obtener los mejores resultados. El algoritmo ha sido optimizado siguiendo los pasos de abajo a arriba. Primero, el *QuaTh* que maximiza la E ha sido encontrado para cada uno de los restantes parámetros. Un ejemplo se muestra en la figura 6.9(a). Esta figura



(a) Ejemplo de E del algoritmo de clasificación de ca- (b) Resultados de clasificación de F1 para diferentes
 lidad de la señal para diferentes umbrales de decisión valores de WLR y WLQ . Los parámetros NBC y $QuaTh$
 ($QuaTh$) y NBC de $IndQuaMin$. En este ejemplo, los han sido optimizados para cada WLR y WLQ
 demás parámetros que son optimizados han sido fijados
 ($WLR=20$ y $WLQ=20$)

Figura 6.9: Resultados de optimización del algoritmo de calidad de señal

Tabla 6.3: Métricas de rendimiento del algoritmo de clasificación de calidad de señal

	Métricas de rendimiento				Parámetros optimizados			
	E	ESP	SEN	F1	WLR	WLQ	NBC	QuaTh
MA	0.63	0.36	0.81	0.70	20	50	11	75
MI	0.63	0.35	0.83	0.72				

muestra la E de clasificación para distintos valores de NBC y $QuaTh$ cuando WLR y WLQ están fijados. Segundo, el $QuaTh$ obtenido ha sido usado para encontrar el mejor NBC que maximiza el F1 para cada combinación de WLQ y WLR . En caso de que F1 similares se hayan obtenido para diferentes NBC , se ha escogido el NBC más bajo para reducir la complejidad del algoritmo. Tercero, se ha encontrado el WLR y WLQ que proporciona los mejores resultados de F1. La figura 6.9(b) muestra los valores de F1 para distintos valores de estos parámetros con el NBC previamente optimizado.

Se debe tener en cuenta que F1 es más sensible a clasificaciones de buena calidad y G proporciona resultados balanceados para las clasificaciones de buena y mala calidad. Por ello, F1 ofrece un buen compromiso en elevados ratios de sensibilidad y exactitud de las predicciones, por eso los parámetros han sido optimizados mediante el criterio de maximización de F1. La tabla 6.3 resume los resultados obtenidos del algoritmo de clasificación de calidad de señal. La tabla muestra los resultados de forma micro-promediada (MI) y macro-promediada (MA), porque mientras que las métricas MA se utilizan cuando se entrenan clasificadores independientes para cada clase, las métricas

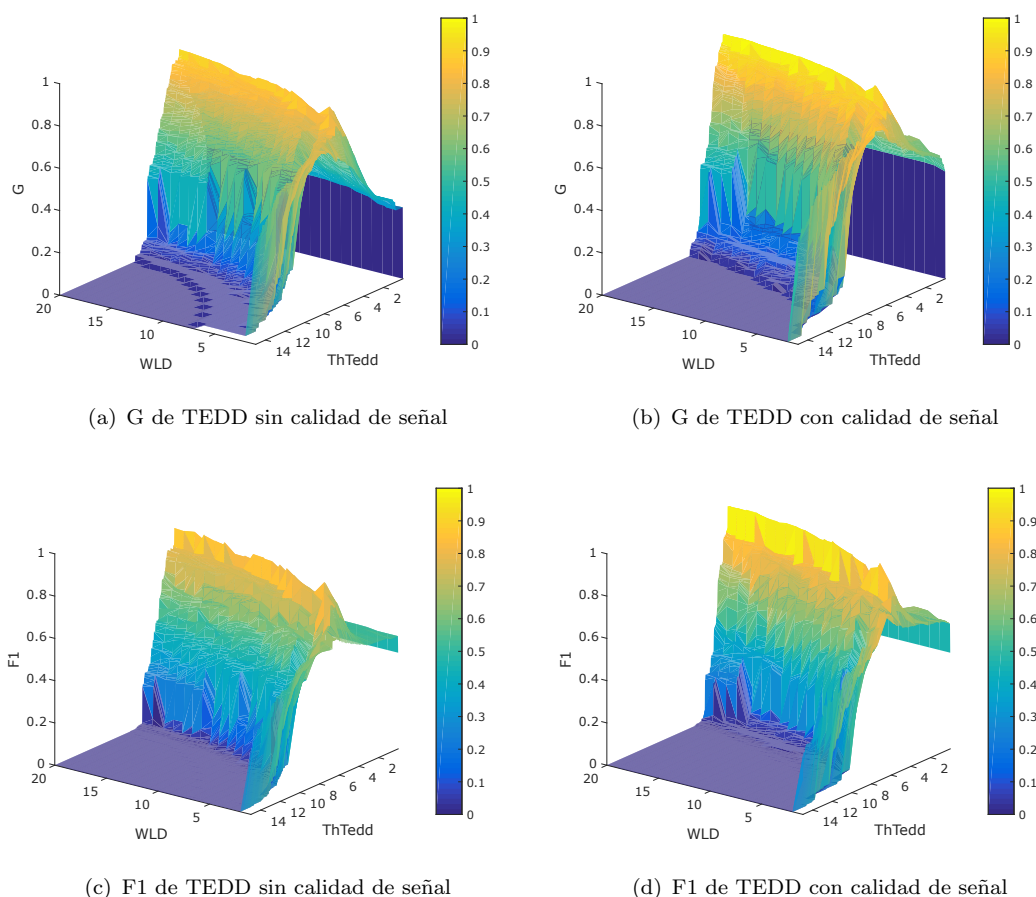


Figura 6.10: Resultados de clasificación de somnolencia del TEDD variando los valores $ThTedd$ y WLD para encontrar los mejores parámetros

MI se utilizan cuando todas las clases se entrenan conjuntamente. Entre MI y MA, MI ha sido usado para ajustar los parámetros porque el mismo clasificador ha sido entrenado para todos los sujetos y los resultados MA han sido utilizados para evaluar el rendimiento global sin sesgo de la clase más representada. La tabla 6.3 muestra que los resultados obtenidos para MI y MA son bastante similares. Además, el algoritmo de calidad optimizado será aplicado al algoritmo TEDD para obtener resultados de clasificación de somnolencia más fiables.

6.5.2 Algoritmo TEDD: resultados del algoritmo de clasificación

6.5.2.1 Ajuste de parámetros del algoritmo TEDD

El algoritmo propuesto ha sido optimizado mediante el ajuste de los parámetros WLD y $ThTedd$, que se han definido en la sección 6.3.3, para obtener el algoritmo con los mejores resultados de clasificación de somnolencia. El método LOSOCV ha sido utilizado para seleccionar los mejores WLD y $ThTedd$ y para evaluar el rendimiento del clasificador de

Tabla 6.4: Resultados de detección de somnolencia para diferente tiempo de avance (AT). Algoritmo TEDD (TeddDis) y TEDD combinado con evaluación de calidad de señal (TeddDisQ)

Método TEDD & max. criterio	AT=5					AT=4				
	ESP	SEN	F1	G	K	ESP	SEN	F1	G	K
TeddDis F1	80.2 ± 10.1	89.9 ± 14.9	77.2 ± 15.0	84.9 ± 8.8	0.59 ± 0.22	75.7 ± 10.5	89.9 ± 14.9	73.7 ± 17.6	82.5 ± 8.9	0.55 ± 0.23
TeddDis G	80.2 ± 10.1	89.9 ± 14.9	77.2 ± 15.0	84.9 ± 8.8	0.59 ± 0.22	75.7 ± 10.5	89.9 ± 14.9	73.7 ± 17.6	82.5 ± 8.9	0.55 ± 0.23
TeddDisQ F1	96.6 ± 3.6	90.3 ± 14.3	83.7 ± 12.7	93.4 ± 7.6	0.75 ± 0.19	93.1 ± 5.1	90.3 ± 14.3	80.4 ± 14.0	91.7 ± 7.7	0.69 ± 0.20
TeddDisQ G	96.6 ± 3.6	90.3 ± 14.3	83.7 ± 12.7	93.4 ± 7.6	0.75 ± 0.19	93.1 ± 5.1	90.3 ± 14.3	80.4 ± 14.0	91.7 ± 7.7	0.69 ± 0.20

somnolencia seleccionado. Por lo tanto, cada *WLD* diferente ha sido considerado como un modelo diferente. Entonces el *WLD* con los mejores resultados ha sido seleccionado en el bucle interior mientras que los resultados de rendimiento generalizados han sido obtenidos en el bucle exterior. Los valores de *WLD* van desde 1 hasta 20, porque los valores mayores de 20 pueden causar un retraso en algoritmo no deseable. La figura 6.10 muestra el G y el F1 del algoritmo TEDD con y sin el algoritmo de evaluación de calidad de la señal variando los valores *WLD* y *ThTedd* durante la fase de ajuste en el conjunto de validación. Esta figura también revela que diferentes *ThTedd* óptimos deben ser encontrados para cada *WLD*. Mientras que el mejor *ThTedd* ha sido buscado maximizando G o F1, la optimización del *WLD* maximiza el área bajo la curva ROC (AUC). AUC es un indicador del rendimiento del clasificador independiente del umbral de clasificación escogido, en nuestro caso *ThTedd*. El AUC ha sido calculado usando el estadístico Wilcoxon-Mann-Whitney, el cual es más eficiente que integrar la curva ROC [220].

6.5.2.2 Evaluación de los resultados del algoritmo TEDD

Los resultados del algoritmo de detección de somnolencia combinados con el algoritmo de evaluación de la calidad de la señal respiratoria y sin esta combinación han sido obtenidos. Estos resultados han sido obtenidos para evaluar la mejora con la contribución del algoritmo de clasificación de calidad de la señal. Además, se ha evaluado el efecto de elegir el criterio de maximización de G o F1 para optimizar *ThTedd*. Los intervalos de confianza (IC) de cada métrica de rendimiento de clasificación han sido obtenidos para comparar y evaluar distintos algoritmos. La incertidumbre de la exactitud ha sido evaluada teniendo en cuenta los efectos de variabilidad entre sujetos. El método de análisis de efectos aleatorios ha sido utilizado para estimar la media e intervalo de confianza de las métricas teniendo en cuenta esta variabilidad. Por lo tanto, la media y

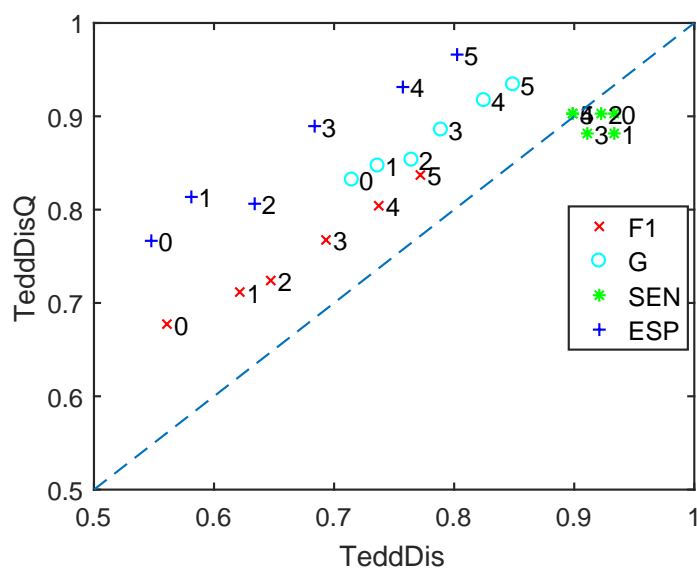


Figura 6.11: Métricas de rendimiento de clasificación de somnolencia del algoritmo TEDD para *TeddDis* y *TeddDisQ*. Se ha seguido el criterio de maximización de F1 para encontrar el mejor *ThTedd*. El número sobre el símbolo indica el valor de AT

el 95% IC de los resultados del algoritmo en validación cruzada han sido obtenidos sobre las particiones a partir del *pooling* a través de todos los sujetos usando el análisis de efectos aleatorios [221]. Los valores de estas métricas para diferentes métodos del TEDD (*TeddDis* y *TeddDisQ*), diferentes criterios de maximización (F1 y G) y para AT=5 y AT=4 están en la tabla 6.4. Por un lado, esta tabla muestra que se ha conseguido el mismo resultado de clasificación para ambos criterios de maximización.

Por otro lado, el test de Welch se ha utilizado para testear la significancia de las diferencias entre los métodos *TeddDis* y *TeddDisQ*. Este test es comúnmente utilizado para comparar las medias de dos muestras independientes con varianzas diferentes. Es una adaptación de t-test de Student y es más fiable cuando las varianzas de las muestras no son iguales [222]. En concreto, se ha encontrado que la ESP, F1, G, *K* obtenidas para *TeddDisQ* para ambos criterios de maximización son significativamente mejores que los de *TeddDis* para ambos AT ($p < 0.05$), mientras que se mantiene la misma SEN. Además, esta tabla muestra que la ESP de *TeddDisQ* ha aumentado de 80.2% a 96.6% mientras las SEN se mantiene alrededor del 90%. Finalmente, la *K* obtenida muestra acuerdo moderado para *TeddDis* y acuerdo substancial para *TeddDisQ*.

La figura 6.11 muestra los resultados de varias métricas de rendimiento de clasificación de *TeddDis* y *TeddDisQ* variando AT. Ambos han sido optimizados con el mismo criterio (maximización del F1) para comparar sólo el efecto del cambio de AT. Los valores de AT van desde 5 minutos a 0 minutos y el número sobre el marcador indica los minutos de AT. Esta figura muestra que ESP, F1, G decrecen cuando AT decrece, porque los FP aumentan cerca del evento de somnolencia mientras se mantiene el mismo SEN para

TeddDis y *TeddDisQ*. Además, los resultados del *TeddDisQ* son mejores que *TeddDis* para todos los diferentes valores de AT. Además, se debe destacar que AT solo afecta a las clasificaciones etiquetadas como negativas (VN y FP) por eso el incremento del número de FP cercanos al evento de somnolencia podría empobrecer el resultado de ESP del algoritmo. En concreto, para AT=5 el FPR (FPR=1-ESP) obtenido es del 3 % mientras que el FPR aumenta hasta el 23 % para AT=0.

En resumen, los resultados obtenidos son similares para los diferentes criterios de entrenamiento pero se ha encontrado una mejora significativa con la combinación de la calidad de la señal respiratoria. Por lo tanto el algoritmo puede ser maximizado con diferentes criterios siempre que el criterio sea insensible a la distribución de las clases. Finalmente, los valores óptimos para la detección de somnolencia del algoritmo TEDD han sido: $WLD=17$ y $ThTedd=3.025$.

6.6 Conclusiones

Se ha propuesto un nuevo algoritmo (TEDD) para la monitorización del estado de alerta del conductor a través de la identificación de la lucha por no dormirse basado en los cambios de la RRV. La RRV se estima a partir del esfuerzo torácico registrado con una banda inductiva. Además, también se ha propuesto un método para proporcionar un nivel de confianza de la calidad de la señal respiratoria. El nivel de calidad de señal obtenido ha sido combinado con el algoritmo de detección de somnolencia para mejorar los resultados de clasificación, esto se ha conseguido a base de reducir el número de falso positivos inducidos por cambios en el RRV medido y que no están asociados a somnolencia sino a movimientos corporales o a hablar.

El algoritmo diseñado ha sido validado en un simulador de conducción y los episodios de somnolencia han sido identificados para cada una de las 15 sesiones. El gold estándar de los eventos de somnolencia ha sido generado por los observadores externos a partir de las grabaciones del sujeto mientras conduce. Los resultados han sido obtenidos siguiendo el método LOSOCV para conseguir una estimación no sesgada del rendimiento generalizado del algoritmo. La mejora de los resultados de clasificación de somnolencia debido al algoritmo de calidad de la señal ha sido evaluada mejorando la especificidad de un 80.2 % a 96.6 % y las diferencias entre los criterios de optimización y AT han sido discutidas.

Finalmente, el rendimiento generalizado de detección de somnolencia ha sido evaluado, consiguiendo una sensibilidad de $96.6 \% \pm 3.6 \%$, una especificidad de $90.3 \% \pm 14.3 \%$ y acuerdo Kappa de Cohen de $0.75 \pm 0.19 \%$ en media a través de todos los sujetos. Por lo tanto, el algoritmo propuesto puede ser un valioso sistema de seguridad en vehículos para alertar de la somnolencia mientras se conduce.

Capítulo 7

Conclusiones

En esta tesis se han tratado los aspectos de la actividad física y la VFC relacionados con la adherencia a un estilo de vida saludable. También se ha realizado un análisis combinado de ambas variables junto con variables psicológicas para evaluar estos aspectos relacionados con la vida saludable de forma conjunta. Además, se ha utilizado un sensor externo de respiración para detectar la somnolencia al volante. Estos aspectos se han tratado en distintos capítulos de la tesis que tienen sus propias conclusiones y aquí se presenta un resumen de los mismos.

Se ha desarrollado y evaluado un algoritmo para la evaluación del gasto calórico asociado a la actividad realizada a partir de los datos de acelerometría. Estos datos se han obtenido utilizando únicamente el sensor del smartphone y registrando las actividades realizadas con una aplicación que implementa los cuestionarios del RSAF. Para ello se han utilizado algoritmos de aprendizaje automático y técnicas de minería de datos. Se han comparado distintos algoritmos de aprendizaje automático. Se han realizado pruebas con 26 sujetos distintos para validar los algoritmos y se ha conseguido una F1-score de 75.6 % en la clasificación de actividades. Además, el gasto calórico se cuantifica con un error del $3.18\% \pm 3.18\%$. A nivel de clasificación de actividades según el gasto calórico se ha alcanzado una F1-score del 85.9 %. Como trabajo futuro, para mejorar los resultados se podrían usar los datos de otros sensores que tiene incorporado el propio smartphone como el barómetro o el vector de rotación. Además, a nivel de algoritmos se podrían explorar soluciones basadas en algoritmos de *deep learning*. También sería interesante utilizar los datos procedentes de los sensores del *smartwatch* para realizar un análisis combinándolos con los del smartphone.

Se ha diseñado una aplicación móvil para la evaluación de la VFC basado en la técnica de fotoplethismografía utilizando cámara del propio smartphone como sensor. Se han realizado medidas de comparación del sistema diseñado con el ECG y un dispositivo de fotoplethismografía de referencia con un grupo de 23 personas. Los resultados obtenidos muestran en mediana una desviación estándar del error de 5.4 ms entre las series latido

a latido, entre el smartphone y el ECG mientras que el error entre el ECG y el sensor de fotopletimografía de referencia es de 4.9 ms. Los resultados han mostrado un grado de acuerdo alto entre el smartphone y los otros métodos para distintas respiraciones y distintos móviles en posición sentada. Se ha obtenido un grado de acuerdo alto para el NN, SDNN y RMSSD pero los valores obtenidos para el LF/HF no son buenos. Validar este método con un mayor número de modelos de smartphone distintos sería una línea de trabajo futuro.

Acerca de las relaciones entre actividad física, VFC y variables psicológicas, se ha diseñado un sistema que permite recoger datos para estudiarlas. Se ha realizado medidas con 11 personas en las que se ha validado el correcto funcionamiento del sistema y se ha obtenido un alto grado de usabilidad. Se ha encontrado que el nivel de la actividad física desciende a lo largo del tiempo y también se ha encontrado alguna asociación entre la constancia en la práctica de la actividad física y los cambios en el estado de ánimo. Sin embargo, estas asociaciones se han de tomar con precaución debido al reducido número de sujetos que han participado en este estudio. Para evaluar las relaciones es necesario un mayor número de personas para extraer conclusiones al respecto. Por tanto, el sistema desarrollado es un punto de partida para evaluar la adherencia a un estilo de vida saludable de forma unificada utilizando una única aplicación. Como trabajo futuro quedaría realizar una evaluación con número mayor de personas y durante un espacio de tiempo prolongado.

La somnolencia en conducción puede estar favorecida por un estilo de vida poco saludable. En esta tesis se ha diseñado y validado un algoritmo para la detección de somnolencia al volante. Este algoritmo está basado en el análisis de la variabilidad del ritmo respiratorio y en el análisis de la calidad de la señal respiratoria para generar una alarma al conductor cuando aparecen los primeros signos de lucha por no dormirse durante la conducción. Se han realizado pruebas con 15 sujetos y se han comparado con clasificaciones de los observadores externos. Se ha obtenido una sensibilidad de 90.3%, una especificidad de 96.6% y el grado de acuerdo Kappa de Cohen es 0.75. Una línea de trabajo futuro sería evaluar el rendimiento del algoritmo en condiciones de conducción real y evaluar este rendimiento con un número mayor de personas. También se podría realizar esta evaluación con pruebas que tengan una mayor duración.

Lista de publicaciones

Revista

- Miguel A. García-González, Mireya Fernández-Chimeno, Federico Guede-Fernández, Víctor Ferrer-Mileo, Ariadna Argelagós-Palau, Laura Álvarez-Gómez, Eva Parrado, Jordi Moreno, Lluís Capdevila, Juan Ramos-Castro. **A methodology to quantify the differences between alternative methods of heart rate variability measurement.** *Physiological Measurement* 37 (1) Enero 2015. DOI: 10.1088/0967-3334/37/1/128, Publicado.

Congreso

- Federico Guede-Fernández, Marc Pous-Solà, Miguel A. García-González, Lluís Capdevila-Ortís, Juan Ramos-Castro, Mireya Fernández-Chimeno. **Using smartphone bases biodevices for analyzing physiological, psychological and behavioral user's habits.** *6th International Conference on Biomedical Electronics and Devices (BIODEVICES)*. Barcelona, España. Febrero 2013. Presentado (póster).
- Federico Guede-Fernández, Bernat Carbonés, Lluís Capdevila, Miguel A. García-González, Juan J. Ramos-Castro, Mireya Fernández-Chimeno. **Assessment of Energy Expended in Physical Activity by a Smartphone-Based System.** *6th European Conference of the IFMBE (MBEC)*. Dubrovnik, Croacia. Septiembre 2014. Presentado (presentación oral).
- Federico Guede-Fernández, Víctor Ferrer-Mileo, Juan Ramos-Castro, Mireya Fernández-Chimeno, Miguel A. García-González. **Real time heart rate variability assessment from Android smartphone camera photoplethysmography: Postural and device influences** *37th Annual International Conference of the IEEE EMBS*. Milán, Italia. Agosto 2015. Presentado (presentación oral).
- Víctor Ferrer-Mileo, Federico Guede-Fernández, Mireya Fernández-Chimeno, Juan Ramos-Castro, Miguel A. García-González. **Accuracy of heart rate variability estimation by photoplethysmography using an smartphone: Processing**

optimization and fiducial point selection *37th Annual International Conference of the IEEE EMBS*. Milán, Italia. Agosto 2015. Presentado (presentación oral).

- Federico Guede-Fernández, Víctor Ferrer-Mileo, Mireya Fernández-Chimeno, Juan Ramos-Castro, Miguel A. García-González. **SeniorFit : Una aplicación móvil para el seguimiento de la adherencia a estilos de vida saludable para gente mayor** *XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB)*. Valencia, España. Noviembre 2016. Presentado (Presentación oral).

Bibliografía

- [1] Organización Mundial de la Salud (OMS), “OMS — Datos interesantes acerca del envejecimiento,” 2012.
- [2] World Health Organization (WHO), “Adherence to long term therapies: evidence for action,” tech. rep., 2003.
- [3] A. M. Martinelli, “An explanatory model of variables influencing health promotion behaviors in smoking and nonsmoking college students,” *Public Health Nursing*, vol. 16, no. 4, pp. 263–269, 1999.
- [4] H. J. Spear and P. A. Kulbok, “Adolescent health behaviors and related factors: A review,” *Public Health Nursing*, vol. 18, no. 2, pp. 82–93, 2001.
- [5] S. H. Kelder, C. L. Perry, K. I. Klepp, and L. L. Lytle, “Longitudinal tracking of adolescent smoking, physical activity, and food choice behaviors,” *American Journal of Public Health*, vol. 84, no. 7, pp. 1121–1126, 1994.
- [6] B. K. Lenz, “Tobacco, depression, and lifestyle choices in the pivotal early college years,” *Journal of American College Health*, vol. 52, no. 5, pp. 213–219, 2004.
- [7] J. E. Gangwisch, S. B. Heymsfield, B. Boden-Albala, R. M. Buijs, F. Kreier, T. G. Pickering, A. G. Rundle, G. K. Zammit, and D. Malaspina, “Short Sleep Duration as a Risk Factor for Hypertension,” *Hypertension*, vol. 47, no. 5, pp. 833–839, 2006.
- [8] J. E. Gangwisch, D. Malaspina, K. Posner, L. A. Babiss, S. B. Heymsfield, J. B. Turner, G. K. Zammit, and T. G. Pickering, “Insomnia and Sleep Duration as Mediators of the Relationship Between Depression and Hypertension Incidence,” *American Journal of Hypertension*, vol. 23, pp. 62–69, jan 2010.
- [9] D. A. Beihl, A. D. Liese, and S. M. Haffner, “Sleep duration as a risk factor for incident type 2 diabetes in a multiethnic cohort,” *Annals of epidemiology*, vol. 19, no. 5, pp. 351–357, 2009.
- [10] Diabetes Prevention Program Research Group, “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” *New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.

- [11] Diabetes Prevention Program Research Group, “10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study,” *The Lancet*, vol. 374, no. 9702, pp. 1677–1686, 2009.
- [12] G. Li, P. Zhang, J. Wang, E. W. Gregg, W. Yang, Q. Gong, H. Li, H. Li, Y. Jiang, and Y. An, “The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-up study,” *The Lancet*, vol. 371, no. 9626, pp. 1783–1789, 2008.
- [13] J. Pinilla, P. Barber, and B. Gonzalez Lopez-Valcarcel, “El coste de la enfermedad potencialmente prevenible en España,” tech. rep., Fundación MAPFRE, Madrid, 2017.
- [14] J. M. Jakicic, “The Effect of Physical Activity on Body Weight,” *Obesity*, vol. 17, pp. S34–S38, dec 2009.
- [15] D. Aune, T. Norat, M. Leitzmann, S. Tonstad, and L. J. Vatten, “Physical activity and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis,” *European journal of epidemiology*, vol. 30, pp. 529–542, jul 2015.
- [16] Y.-S. Lee and S.-B. Cho, *Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer*, pp. 460–467. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
- [17] L. Liu, Y. Shi, T. Li, Q. Qin, J. Yin, S. Pang, S. Nie, and S. Wei, “Leisure time physical activity and cancer risk: evaluation of the WHO’s recommendation based on 126 high-quality epidemiological studies,” *British journal of sports medicine*, vol. 50, pp. 372–378, mar 2016.
- [18] F. Sofi, A. Capalbo, F. Cesari, R. Abbate, and G. F. Gensini, “Physical activity during leisure time and primary prevention of coronary heart disease: an updated meta-analysis of cohort studies,” *European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology*, vol. 15, pp. 247–257, jun 2008.
- [19] A. L. Rebar, R. Stanton, D. Geard, C. Short, M. J. Duncan, and C. Vandelanotte, “A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations,” *Health psychology review*, vol. 9, no. 3, pp. 366–378, 2015.
- [20] G. Mammen and G. Faulkner, “Physical activity and the prevention of depression: a systematic review of prospective studies,” *American journal of preventive medicine*, vol. 45, pp. 649–657, nov 2013.
- [21] S. Shiffman, A. A. Stone, and M. R. Hufford, “Ecological momentary assessment,” *Annual review of clinical psychology*, vol. 4, pp. 1–32, 2008.

- [22] O. Marquet, C. Alberico, and A. J. Hipp, "Pokémon GO and physical activity among college students. A study using Ecological Momentary Assessment," *Computers in Human Behavior*, vol. 81, pp. 215–222, 2018.
- [23] J. A. Emerson, S. Dunsiger, and D. M. Williams, "Reciprocal within-day associations between incidental affect and exercise: An EMA study.," *Psychology & health*, vol. 33, pp. 130–143, jan 2018.
- [24] G. F. Dunton, "Ecological Momentary Assessment in Physical Activity Research.," *Exercise and sport sciences reviews*, vol. 45, pp. 48–54, jan 2017.
- [25] B. H. Dobkin and A. Dorsch, "The promise of mHealth: Daily activity monitoring and outcome assessments by wearable sensors," *Neurorehabilitation and neural repair*, vol. 25, no. 9, pp. 788–798, 2011.
- [26] J. P. Higgins, "Smartphone Applications for Patients'Health and Fitness.," *The American journal of medicine*, vol. 129, pp. 11–19, jan 2016.
- [27] A. M. Muller, S. Alley, S. Schoeppe, and C. Vandelanotte, "The effectiveness of e-& mHealth interventions to promote physical activity and healthy diets in developing countries: A systematic review.," *The international journal of behavioral nutrition and physical activity*, vol. 13, p. 109, oct 2016.
- [28] G. A. O'Reilly and D. Spruijt-Metz, "Current mHealth technologies for physical activity assessment and promotion.," *American journal of preventive medicine*, vol. 45, pp. 501–507, oct 2013.
- [29] S. S. Coughlin, M. Whitehead, J. Q. Sheats, J. Mastromonico, and S. Smith, "A Review of Smartphone Applications for Promoting Physical Activity.," *Jacobs journal of community medicine*, vol. 2, no. 1, 2016.
- [30] World Health Organization (WHO), "Mental health: strengthening our response," 2016.
- [31] N. Montano, A. Porta, C. Cogliati, G. Costantino, E. Tobaldini, K. R. Casali, and F. Iellamo, "Heart rate variability explored in the frequency domain: a tool to investigate the link between heart and behavior.," *Neuroscience and biobehavioral reviews*, vol. 33, pp. 71–80, feb 2009.
- [32] T. Chandola, A. Heraclides, and M. Kumari, "Psychophysiological biomarkers of workplace stressors," *Neuroscience & Biobehavioral Reviews*, vol. 35, no. 1, pp. 51–57, 2010.
- [33] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," *Biomedical Signal Processing and Control*, vol. 18, pp. 370–377, 2015.

- [34] A. H. Kemp and D. S. Quintana, "The relationship between mental and physical health: insights from the study of heart rate variability," *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, vol. 89, pp. 288–296, sep 2013.
- [35] J. A. Chalmers, D. S. Quintana, M. J.-A. Abbott, and A. H. Kemp, "Anxiety Disorders are Associated with Reduced Heart Rate Variability: A Meta-Analysis," *Frontiers in Psychiatry*, vol. 5, p. 80, 2014.
- [36] A. K. Schuster, J. E. Fischer, J. F. Thayer, D. Mauss, and M. N. Jarczok, "Decreased heart rate variability correlates to increased cardiovascular risk," *International Journal of Cardiology*, vol. 203, pp. 728–730, 2016.
- [37] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart Rate Variability : Standards of Measurement, Physiological Interpretation, and Clinical Use," 1996.
- [38] D. Moss and F. Shaffer, "The Application of Heart Rate Variability Biofeedback to Medical and Mental Health Disorders," *Biofeedback*, vol. 45, pp. 2–8, mar 2017.
- [39] D. Foley, S. Ancoli-Israel, P. Britz, and J. Walsh, "Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey," *Journal of psychosomatic research*, vol. 56, no. 5, pp. 497–502, 2004.
- [40] D. Ndahimana and E.-K. Kim, "Measurement Methods for Physical Activity and Energy Expenditure: a Review.," *Clinical nutrition research*, vol. 6, pp. 68–80, apr 2017.
- [41] A. C. Pinheiro Volp, F. C. Esteves de Oliveira, R. Duarte Moreira Alves, E. A. Esteves, and J. Bressan, "Energy expenditure: components and evaluation methods.," *Nutricion hospitalaria*, vol. 26, no. 3, pp. 430–440, 2011.
- [42] P. Trumbo, S. Schlicker, A. A. Yates, and M. Poos, "Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids," *Journal of the American Dietetic Association*, vol. 102, no. 11, pp. 1621–1630, 2002.
- [43] K. R. Westerterp, "Diet induced thermogenesis," *Nutrition & Metabolism*, vol. 1, no. 1, p. 5, 2004.
- [44] P. S. Freedson and K. Miller, "Objective monitoring of physical activity using motion sensors and heart rate.," *Research quarterly for exercise and sport*, vol. 71, pp. S21–9, jun 2000.
- [45] A. Luke, K. C. Maki, N. Barkey, R. Cooper, and D. McGee, "Simultaneous monitoring of heart rate and motion to assess energy expenditure.," *Medicine and science in sports and exercise*, vol. 29, pp. 144–148, jan 1997.

- [46] P. Ainslie, T. Reilly, and K. Westerterp, “Estimating human energy expenditure: a review of techniques with particular reference to doubly labelled water,” *Sports medicine (Auckland, N.Z.)*, vol. 33, no. 9, pp. 683–698, 2003.
- [47] S. M. Ceesay, A. M. Prentice, K. C. Day, P. R. Murgatroyd, G. R. Goldberg, W. Scott, and G. B. Spurr, “The use of heart rate monitoring in the estimation of energy expenditure: a validation study using indirect whole-body calorimetry,” *The British journal of nutrition*, vol. 61, pp. 175–186, mar 1989.
- [48] Y. Schutz, R. L. Weinsier, and G. R. Hunter, “Assessment of free-living physical activity in humans: an overview of currently available and proposed new measu-res.,” *Obesity research*, vol. 9, pp. 368–379, jun 2001.
- [49] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett Jr, C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, and A. S. Leon, “2011 Compendium of Physical Activities: a second update of codes and MET values,” *Medicine & science in sports & exercise*, vol. 43, no. 8, pp. 1575–1581, 2011.
- [50] B. E. Ainsworth, W. L. Haskell, M. C. Whitt, M. L. Irwin, A. M. Swartz, S. J. Strath, W. L. O’Brien, D. R. Bassett, K. H. Schmitz, P. O. Emplaincourt, D. R. Jacobs, and A. S. Leon, “Compendium of physical activities: an update of activity codes and MET intensities,” *Medicine and science in sports and exercise*, vol. 32, no. 9 Suppl, pp. S498–504, 2000.
- [51] M. D. Mifflin, S. T. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. O. Koh, “A new predictive equation for resting energy expenditure in healthy individuals,” *The American Journal of Clinical Nutrition*, vol. 51, pp. 241–247, feb 1990.
- [52] A. Ruiz-Zafra, E. Orantes-González, M. Noguera, K. Benghazi, and J. Heredia-Jimenez, “A Comparative Study on the Suitability of Smartphones and IMU for Mobile, Unsupervised Energy Expenditure Calculi,” *Sensors (Basel, Switzerland)*, vol. 15, no. 8, pp. 18270–18286, 2015.
- [53] J. J. Guiry, P. van de Ven, J. Nelson, L. Warmerdam, and H. Riper, “Activity recognition with smartphone support,” *Medical Engineering & Physics*, vol. 36, no. 6, pp. 670–675, 2014.
- [54] S. Rousset, R. Guidoux, L. Paris, N. Farigon, M. Miolanne, C. Lahaye, M. Duclos, Y. Boirie, and D. Saboul, “A Novel Smartphone Accelerometer Application for Low-Intensity Activity and Energy Expenditure Estimations in Overweight and Obese Adults,” *Journal of Medical Systems*, vol. 41, p. 117, jul 2017.
- [55] E. Mitchell, D. Monaghan, and N. O’Connor, “Classification of Sporting Activities Using Smartphone Accelerometers,” *Sensors*, vol. 13, pp. 5317–5337, apr 2013.

- [56] A. Pande, J. Zhu, A. K. Das, Y. Zeng, P. Mohapatra, and J. J. Han, "Using Smartphone Sensors for Improving Energy Expenditure Estimation," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 3, pp. 1–12, 2015.
- [57] M. B. D. Rosario, K. Wang, J. Wang, Y. Liu, M. Brodie, K. Delbaere, N. H. Lovell, S. R. Lord, and S. J. Redmond, "A comparison of activity classification in younger and older cohorts using a smartphone," *Physiological Measurement*, vol. 35, p. 2269, nov 2014.
- [58] R. Guidoux, M. Duclos, G. Fleury, P. Lacomme, N. Lamaudière, P.-H. Manenq, L. Paris, L. Ren, and S. Rousset, "A smartphone-driven methodology for estimating physical activities and energy expenditure in free living conditions," *Journal of Biomedical Informatics*, vol. 52, no. Supplement C, pp. 271–278, 2014.
- [59] A. Moncada-Torres, K. Leuenberger, R. Gonzenbach, A. Luft, and R. Gassert, "Activity classification based on inertial and barometric pressure sensors at different anatomical locations," *Physiological Measurement*, vol. 35, no. 7, p. 1245, 2014.
- [60] L. Atallah, B. Lo, R. King, and G. Z. Yang, "Sensor Positioning for Activity Recognition Using Wearable Accelerometers," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 4, pp. 320–329, 2011.
- [61] S. Pirttikangas, K. Fujinami, and T. Nakajima, "Feature Selection and Activity Recognition from Wearable Sensors," in *Ubiquitous Computing Systems* (H. Y. Youn, M. Kim, and H. Morikawa, eds.), (Berlin, Heidelberg), pp. 516–527, Springer Berlin Heidelberg, 2006.
- [62] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions," in *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, pp. 4 pp.–116, 2006.
- [63] G. Plasqui, "Smart approaches for assessing free-living energy expenditure following identification of types of physical activity," *Obesity Reviews*, vol. 18, pp. 50–55, 2017.
- [64] E. B. Hekler, M. P. Buman, L. Grieco, M. Rosenberger, S. J. Winter, W. Haskell, and A. C. King, "Validation of Physical Activity Tracking via Android Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living Validation Studies," 2015.
- [65] A. Wang, G. Chen, J. Yang, S. Zhao, and C.-Y. Chang, "A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone," *IEEE Sensors Journal*, vol. 16, pp. 4566–4578, jun 2016.

- [66] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, “Complex Human Activity Recognition Using Smartphone and Wrist-Worn Motion Sensors,” *Sensors*, vol. 16, no. 4, 2016.
- [67] C. Catal, S. Tufekci, E. Pirit, and G. Kocabag, “On the use of ensemble of classifiers for accelerometer-based activity recognition,” *Applied Soft Computing*, vol. 37, no. Supplement C, pp. 1018–1022, 2015.
- [68] R. K. Dishman, Y. Nakamura, M. E. Garcia, R. W. Thompson, A. L. Dunn, and S. N. Blair, “Heart rate variability, trait anxiety, and perceived stress among physically fit men and women,” *International Journal of Psychophysiology*, vol. 37, no. 2, pp. 121–133, 2000.
- [69] G. Rodas, C. Pedret, J. Ramos-Castro, and L. Capdevila, “Variabilidad de la frecuencia cardíaca: concepto, medidas y relación con aspectos clínicos (I),” *Archivos de medicina del deporte: revista de la Federación Española de Medicina del Deporte y de la Confederación Iberoamericana de Medicina del Deporte*, ISSN 0212-8799, N^o. 123, 2008, pags. 41-48, vol. 123, 2008.
- [70] F. Shaffer and J. P. Ginsberg, “An Overview of Heart Rate Variability Metrics and Norms,” *Frontiers in Public Health*, vol. 5, p. 258, sep 2017.
- [71] E. Parrado, M. A. Garcia, J. Ramos, J. C. Cervantes, G. Rodas, and L. Capdevila, “Comparison of Omega Wave System and Polar S810i to detect R-R intervals at rest,” *International journal of sports medicine*, vol. 31, pp. 336–341, may 2010.
- [72] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, pp. R1–R39, mar 2007.
- [73] M. Elgendi, “On the Analysis of Fingertip Photoplethysmogram Signals,” *Current Cardiology Reviews*, vol. 8, pp. 14–25, jun 2012.
- [74] X. Li, J. Chen, G. Zhao, and M. Pietikäinen, “Remote Heart Rate Measurement from Face Videos under Realistic Situations,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4264–4271, 2014.
- [75] S. Kwon, H. Kim, and K. S. Park, “Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2174–2177, 2012.
- [76] S. Kwon, J. Kim, D. Lee, and K. Park, “ROI analysis for remote photoplethysmography on facial video,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4938–4941, 2015.
- [77] W. J. Jiang, S. C. Gao, P. Wittek, and L. Zhao, “Real-time quantifying heart beat rate from facial video recording on a smart phone using Kalman filters,” in

- 2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 393–396, 2014.
- [78] C. Huang, X. Yang, and K. T. T. Cheng, “Accurate and efficient pulse measurement from facial videos on smartphones,” in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–8, 2016.
- [79] S. W. Cheatham, M. J. Kolber, and M. P. Ernst, “Concurrent validity of resting pulse-rate measurements: a comparison of 2 smartphone applications, the polar h7 belt monitor, and a pulse oximeter with bluetooth,” *Journal of sport rehabilitation*, vol. 24, no. 2, pp. 171–178, 2015.
- [80] T. Coppetti, A. Brauchlin, S. Muggler, A. Attinger-Toller, C. Templin, F. Schonrath, J. Hellermann, T. F. Luscher, P. Biaggi, and C. A. Wyss, “Accuracy of smartphone apps for heart rate measurement.,” *European journal of preventive cardiology*, vol. 24, pp. 1287–1293, aug 2017.
- [81] B. P. Yan, C. K. Chan, C. K. Li, O. T. To, W. H. Lai, G. Tse, Y. C. Poh, and M.-Z. Poh, “Resting and Postexercise Heart Rate Detection From Fingertip and Facial Photoplethysmography Using a Smartphone Camera: A Validation Study.,” *JMIR mHealth and uHealth*, vol. 5, p. e33, mar 2017.
- [82] C. G. Scully, Jinseok Lee, J. Meyer, A. M. Gorbach, D. Granquist-Fraser, Y. Mendelson, and K. H. Chon, “Physiological Parameter Monitoring from Optical Recordings With a Mobile Phone,” *IEEE Transactions on Biomedical Engineering*, vol. 59, pp. 303–306, feb 2012.
- [83] J. A. J. Heathers, “Smartphone-enabled pulse rate variability: An alternative methodology for the collection of heart rate variability in psychophysiological research,” *International Journal of Psychophysiology*, vol. 89, pp. 297–304, sep 2013.
- [84] R.-C. Peng, X.-L. Zhou, W.-H. Lin, and Y.-T. Zhang, “Extraction of Heart Rate Variability from Smartphone Photoplethysmograms,” *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 1–11, jan 2015.
- [85] J. Pumprla, K. Howorka, D. Groves, M. Chester, and J. Nolan, “Functional assessment of heart rate variability: physiological basis and practical applications,” *International Journal of Cardiology*, vol. 84, no. 1, pp. 1–14, 2002.
- [86] D. E. McMillan, “Interpreting heart rate variability sleep/wake patterns in cardiac patients.,” *The Journal of cardiovascular nursing*, vol. 17, pp. 69–81, oct 2002.
- [87] L. C. M. Vanderlei, C. M. Pastre, R. A. Hoshi, T. D. de Carvalho, and M. F. de Godoy, “Basic notions of heart rate variability and its clinical applicability.,” *Revista brasileira de cirurgia cardiovascular : orgao oficial da Sociedade Brasileira de Cirurgia Cardiovascular*, vol. 24, no. 2, pp. 205–217, 2009.

- [88] R. Mccraty and F. Shaffer, “Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk,” *Global Advances in Health and Medicine*, vol. 4, pp. 46–61, jan 2015.
- [89] B. Martínez-Pérez, I. de la Torre-Díez, M. López-Coronado, and J. Herreros-González, “Mobile Apps in Cardiology: Review,” *JMIR mhealth and uhealth*, vol. 1, p. e15, jul 2013.
- [90] M. J. Gregoski, M. Mueller, A. Vertegel, A. Shaporev, B. B. Jackson, R. M. Frenzel, S. M. Sprehn, and F. A. Treiber, “Development and Validation of a Smartphone Heart Rate Acquisition Application for Health Promotion and Wellness Telehealth Applications,” *International Journal of Telemedicine and Applications*, vol. 2012, pp. 1–7, 2012.
- [91] M.-Z. Poh and Y. C. Poh, “Validation of a Standalone Smartphone Application for Measuring Heart Rate Using Imaging Photoplethysmography.,” *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, vol. 23, pp. 678–683, aug 2017.
- [92] T. Vandenberk, J. Stans, G. Van Schelvergem, C. Pelckmans, J. P. C. Smeets, D. Lanssens, H. De Canniere, V. Storms, M. I. Thijs, and M. P. Vandervoort, “Clinical Validation of Heart Rate Apps: Mixed-Methods Evaluation Study,” *JMIR Mhealth Uhealth*, vol. 5, no. 8, p. e129, 2017.
- [93] K. Mitchell, M. Graff, C. Hedt, and J. Simmons, “Reliability and validity of a smartphone pulse rate application for the assessment of resting and elevated pulse rate.,” *Physiotherapy theory and practice*, vol. 32, pp. 494–499, aug 2016.
- [94] M. E. Losa-Iglesias, R. Becerro-de Bengoa-Vallejo, and K. R. Becerro-de Bengoa-Losa, “Reliability and concurrent validity of a peripheral pulse oximeter and health-app system for the quantification of heart rate in healthy adults,” *Health informatics journal*, vol. 22, no. 2, pp. 151–159, 2014.
- [95] B. De Ridder, B. Van Rompaey, K. J. Kampen, S. Haine, and T. Dilles, “Smartphone Apps Using Photoplethysmography for Heart Rate Monitoring: Meta-Analysis,” *JMIR Cardio*, vol. 2, p. e4, feb 2018.
- [96] J. B. Bolkhovsky, C. G. Scully, and K. H. Chon, “Statistical analysis of heart rate and heart rate variability monitoring through the use of smart phone cameras.,” *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2012, pp. 1610–1613, 2012.
- [97] L. Drijckoningen, F. Lenaerts, J. Van der Auwera, K. Leysen, D. Nuyens, P. Vandervoort, and L. Grieten, “Validation of a smartphone based photoplethysmographic

- beat detection algorithm for normal and ectopic complexes,” in *Computing in Cardiology Conference (CinC), 2014*, pp. 845–848, IEEE, 2014.
- [98] N. Koenig, A. Seeck, J. Eckstein, A. Mainka, T. Huebner, A. Voss, and S. Weber, “Validation of a new heart rate measurement algorithm for fingertip recording of video signals with smartphones,” *Telemedicine and e-Health*, vol. 22, no. 8, pp. 631–636, 2016.
- [99] K. Matsumura and T. Yamakoshi, “iPhysioMeter: a new approach for measuring heart rate and normalized pulse volume using only a smartphone,” *Behavior research methods*, vol. 45, no. 4, pp. 1272–1278, 2013.
- [100] Y. Nam, Y. Kong, B. Reyes, N. Reljin, and K. H. Chon, “Monitoring of heart and breathing rates using dual cameras on a smartphone,” *PloS one*, vol. 11, no. 3, p. e0151013, 2016.
- [101] L.-M. Po, X. Xu, L. Feng, Y. Li, K.-W. Cheung, and C.-H. Cheung, “Frame adaptive ROI for photoplethysmography signal extraction from fingertip video captured by smartphone,” in *Circuits and Systems (ISCAS), 2015 IEEE International Symposium on*, pp. 1634–1637, IEEE, 2015.
- [102] D. J. Plews, B. Scott, M. Altini, M. Wood, A. E. Kilding, and P. B. Laursen, “Comparison of heart-rate-variability recording with smartphone photoplethysmography, polar H7 chest strap, and electrocardiography,” *International Journal of Sports Physiology and Performance*, vol. 12, pp. 1324–1328, mar 2017.
- [103] A. Bánhalmi, J. Borbás, M. Fidrich, V. Bilicki, Z. Gingl, and L. Rudas, “Analysis of a Pulse Rate Variability Measurement Using a Smartphone Camera,” *Journal of Healthcare Engineering*, vol. 2018, 2018.
- [104] N. Bruining, E. Caiani, C. Chronaki, P. Guzik, and E. van der Velde, “Acquisition and analysis of cardiovascular signals on smartphones: potential, pitfalls and perspectives,” *European Journal of Preventive Cardiology*, vol. 21, pp. 4–13, nov 2014.
- [105] H. H. Nguyen and J. N. A. Silva, “Use of smartphone technology in cardiology,” *Trends in cardiovascular medicine*, vol. 26, no. 4, pp. 376–386, 2016.
- [106] J. D. Slater, “A definition of drowsiness: One purpose for sleep?,” *Medical Hypotheses*, vol. 71, pp. 641–644, nov 2008.
- [107] N. H. T. S. Administration and Others, “Drowsy driving and automobile crashes,” *NCSDR/NHTSA Expert Panel on Driver Fatigue and Sleepiness. DOT Report HS*, vol. 808, p. 707, 1998.
- [108] AAA Foundation for Traffic Safety, “2016 Traffic Safety Culture Index,” tech. rep., AAA Foundation for Traffic Safety, feb 2017.

- [109] World Health Organization. Violence and Injury Prevention and World Health Organization, *Global status report on road safety 2015*. World Health Organization, 2015.
- [110] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat, and L. Smith, “Fatigue and road safety: a critical analysis of recent evidence,” *London: Department for Transport*, no. 21, 2011.
- [111] B. C. Teff, “Acute Sleep Deprivation and Risk of Motor Vehicle Crash Involvement,” Tech. Rep. December, AAA Foundation for Traffic Safety, 2016.
- [112] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, “Steering in a random forest: ensemble learning for detecting drowsiness-related lane departures,” *Human factors*, vol. 56, pp. 986–998, aug 2014.
- [113] W. W. Wierwille, S. S. Wreggit, C. L. Kirn, L. A. Ellsworth, and R. J. Fairbanks, “Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness. final report,” tech. rep., U.S. Department of Transportation, National Highway Traffic Safety Administration, 1994.
- [114] T. D’Orazio, M. Leo, C. Guaragnella, and A. Distante, “A visual approach for driver inattention detection,” *Pattern Recognition*, vol. 40, pp. 2341–2355, aug 2007.
- [115] J. Jo, S. J. Lee, K. R. Park, I.-J. Kim, and J. Kim, “Detecting driver drowsiness using feature-level fusion and user-specific classification,” *Expert Systems with Applications*, vol. 41, no. 4, pp. 1139–1152, 2014.
- [116] S. J. Lee, J. Jo, H. G. Jung, K. R. Park, and J. Kim, “Real-Time Gaze Estimator Based on Driver’s Head Orientation for Forward Collision Warning System,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, pp. 254–267, mar 2011.
- [117] A. Garces Correa, L. Orosco, and E. Laciari, “Automatic detection of drowsiness in EEG records based on multimodal analysis,” *Medical engineering & physics*, vol. 36, pp. 244–249, feb 2014.
- [118] J. W. Fu, M. Li, and B. L. Lu, *Detecting Drowsiness in Driving Simulation Based on EEG*, pp. 21–28. Dordrecht: Springer Netherlands, 2008.
- [119] M. Hajinorozi, Z. Mao, T. P. Jung, C. T. Lin, and Y. Huang, “EEG-based prediction of driver’s cognitive performance by deep convolutional neural network,” *Signal Processing: Image Communication*, vol. 47, pp. 549–555, sep 2016.
- [120] L. Schmitt, J. Regnard, M. Desmarests, F. Mauny, L. Mourot, J.-P. Fouillot, N. Coulmy, and G. Millet, “Fatigue shifts and scatters heart rate variability in elite endurance athletes,” *PloS one*, vol. 8, no. 8, p. e71588, 2013.

- [121] M. Ako, T. Kawara, S. Uchida, S. Miyazaki, K. Nishihara, J. Mukai, K. Hirao, J. Ako, and Y. Okubo, “Correlation between electroencephalography and heart rate variability during sleep.,” *Psychiatry and clinical neurosciences*, vol. 57, pp. 59–65, feb 2003.
- [122] J. van den Berg, G. Neely, U. Wiklund, and U. Landström, “Heart rate variability during sedentary work and sleep in normal and sleep-deprived states.,” *Clinical physiology and functional imaging*, vol. 25, pp. 51–7, jan 2005.
- [123] Z. Mao, X.-p. Yan, and C.-z. Wu, “Driving Fatigue Identification Method Based on Physiological Signals,” in *Plan, Build, and Manage Transportation Infrastructure in China*, pp. 341–352, ASCE, 2008.
- [124] H. De Rosario, J. S. Solaz, N. Rodriguez, and L. M. Bergasa, “Controlled inducement and measurement of drowsiness in a driving simulator,” *IET intelligent transport systems*, vol. 4, no. 4, pp. 280–288, 2010.
- [125] P. Daponte, L. De Vito, F. Picariello, and M. Riccio, “State of the art and future developments of measurement applications on smartphones,” *Measurement*, vol. 46, no. 9, pp. 3291–3307, 2013.
- [126] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, “A survey of mobile phone sensing,” *Communications Magazine, IEEE*, vol. 48, no. 9, pp. 140–150, 2010.
- [127] F. Lamonaca, G. Polimeni, K. Barbé, and D. Grimaldi, “Health parameters monitoring by smartphone for quality of life improvement,” *Measurement*, vol. 73, pp. 82–94, 2015.
- [128] V. P. Cornet and R. J. Holden, “Systematic review of smartphone-based passive sensing for health and wellbeing.,” *Journal of biomedical informatics*, vol. 77, pp. 120–132, jan 2018.
- [129] N. Noury, K. A. Quach, M. Berenguer, M. J. Bouzi, and H. Teyssier, “A feasibility study of using a smartphone to monitor mobility in elderly,” in *2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 423–426, 2012.
- [130] K. Muralidharan, A. J. Khan, A. Misra, R. K. Balan, and S. Agarwal, “Barometric Phone Sensors: More Hype Than Hope!,” in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications, HotMobile ’14*, (New York, NY, USA), pp. 12:1—12:6, ACM, 2014.
- [131] S. Vanini, F. Faraci, A. Ferrari, and S. Giordano, “Using barometric pressure data to recognize vertical displacement activities on smartphones,” *Computer Communications*, vol. 87, pp. 37–48, 2016.

- [132] M. A. Habib, M. S. Mohktar, S. B. Kamaruzzaman, K. S. Lim, T. M. Pin, and F. Ibrahim, “Smartphone-based solutions for fall detection and prevention: challenges and open issues.,” *Sensors (Basel, Switzerland)*, vol. 14, pp. 7181–7208, apr 2014.
- [133] R. Y. W. Lee and A. J. Carlisle, “Detection of falls using accelerometers and mobile phone technology.,” *Age and ageing*, vol. 40, pp. 690–696, nov 2011.
- [134] S. R. Lord, C. Sherrington, H. B. Menz, and J. C. T. Close, *Falls in older people: risk factors and strategies for prevention*. Cambridge University Press, 2007.
- [135] S. Brownsell and M. S. Hawley, “Automatic fall detectors and the fear of falling.,” *Journal of telemedicine and telecare*, vol. 10, no. 5, pp. 262–266, 2004.
- [136] S. Saeb, E. G. Lattie, S. M. Schueller, K. P. Kording, and D. C. Mohr, “The relationship between mobile phone location sensor data and depressive symptom severity.,” *PeerJ*, vol. 4, p. e2537, 2016.
- [137] F. Wahle, T. Kowatsch, E. Fleisch, M. Rufer, and S. Weidt, “Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild.,” *JMIR mHealth and uHealth*, vol. 4, p. e111, sep 2016.
- [138] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, and D. C. Mohr, “Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study.,” *Journal of medical Internet research*, vol. 17, p. e175, jul 2015.
- [139] F. Miskelly, “Electronic tracking of patients with dementia and wandering using mobile phone technology.,” *Age and ageing*, vol. 34, pp. 497–499, sep 2005.
- [140] D. D. McManus, J. W. Chong, A. Soni, J. S. Saczynski, N. Esa, C. Napolitano, C. E. Darling, E. Boyer, R. K. Rosen, K. C. Floyd, and K. H. Chon, “PULSE-SMART: Pulse-Based Arrhythmia Discrimination Using a Novel Smartphone Application.,” *Journal of cardiovascular electrophysiology*, vol. 27, pp. 51–57, jan 2016.
- [141] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, “StressSense,” in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, vol. Pittsburgh, (New York, New York, USA), p. 351, ACM Press, 2012.
- [142] M. C. Townsend, “Spirometry in the occupational health setting—2011 update.,” *Journal of occupational and environmental medicine*, vol. 53, pp. 569–584, may 2011.
- [143] E. C. Larson, M. Goel, G. Boriello, S. Heltshe, M. Rosenfeld, and S. N. Patel, “SpiroSmart: Using a Microphone to Measure Lung Function on a Mobile Phone.,”

- in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12*, (New York, NY, USA), pp. 280–289, ACM, 2012.
- [144] F. W. Booth, C. K. Roberts, and M. J. Laye, “Lack of exercise is a major cause of chronic diseases,” *Comprehensive Physiology*, vol. 2, pp. 1143–1211, apr 2012.
- [145] Organización Mundial de la Salud (OMS), “Actividad física,” 2017.
- [146] A. Middelweerd, J. S. Mollee, C. N. van der Wal, J. Brug, and S. J. Te Velde, “Apps to promote physical activity among adults: a review and content analysis,” *International journal of behavioral nutrition and physical activity*, vol. 11, no. 1, p. 97, 2014.
- [147] E. Parrado, J. C. Cervantes, M. Ocaña, M. Pintanel, M. Valero, and L. Capdevila, “Evaluación de la conducta activa: el registro semanal de actividad física (RSAF),” *Revista de psicología del deporte*, vol. 18, no. 2, 2009.
- [148] C. V. C. Bouten, K. T. M. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, “A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity,” *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 3, pp. 136–147, 1997.
- [149] K. Ellis, J. Kerr, S. Godbole, G. Lanckriet, D. Wing, and S. Marshall, “A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers,” *Physiological measurement*, vol. 35, no. 11, p. 2191, 2014.
- [150] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC press, 1984.
- [151] L. Breiman, “Bagging predictors,” *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [152] L. Breiman, “Arcing classifier (with discussion and a rejoinder by the author),” *The annals of statistics*, vol. 26, no. 3, pp. 801–849, 1998.
- [153] M. Skurichina and R. P. W. Duin, “Bagging, boosting and the random subspace method for linear classifiers,” *Pattern Analysis & Applications*, vol. 5, no. 2, pp. 121–135, 2002.
- [154] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, “RUSBoost: Improving classification performance when training data is skewed,” in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pp. 1–4, IEEE, 2008.
- [155] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

- [156] P. Refaeilzadeh, L. Tang, and H. Liu, *Cross-Validation*, pp. 532–538. Boston, MA: Springer US, 2009.
- [157] F. H. Joanneum, “Cross-validation explained,” *Graz, Austria: Institute for Genomics and Bioinformatics, Graz University of Technology*, 2005.
- [158] S. Varma and R. Simon, “Bias in error estimation when using cross-validation for model selection,” *BMC bioinformatics*, vol. 7, no. 1, p. 91, 2006.
- [159] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001.
- [160] A. L. Wheat and K. T. Larkin, “Biofeedback of heart rate variability and related physiology: a critical review.,” *Applied psychophysiology and biofeedback*, vol. 35, pp. 229–242, sep 2010.
- [161] E. Gil, M. Orini, R. Bailon, J. M. Vergara, L. Mainardi, and P. Laguna, “Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions.,” *Physiological measurement*, vol. 31, pp. 1271–1290, sep 2010.
- [162] C.-C. Chuang, J.-J. Ye, W.-C. Lin, K.-T. Lee, and Y.-T. Tai, “Photoplethysmography variability as an alternative approach to obtain heart rate variability information in chronic pain patient.,” *Journal of clinical monitoring and computing*, vol. 29, pp. 801–806, dec 2015.
- [163] A. V. J. Challoner, “Photoelectric plethysmography for estimating cutaneous blood flow,” *Non-invasive physiological measurements*, vol. 1, pp. 125–151, 1979.
- [164] J. Weinman, *Photoplethysmography Manual of Psycho-Physiological Methods*. 1967.
- [165] R. R. Anderson and J. A. Parrish, “The optics of human skin,” *Journal of investigative dermatology*, vol. 77, no. 1, pp. 13–19, 1981.
- [166] L. S. D’Ágrosa and A. B. Hertzman, “Opacity pulse of individual minute arteries.,” *Journal of applied physiology*, vol. 23, no. 5, pp. 613–620, 1967.
- [167] R. W. Treskes, E. T. van der Velde, R. Barendse, and N. Bruining, “Mobile health in cardiology: a review of currently available medical apps and equipment for remote monitoring.,” *Expert review of medical devices*, vol. 13, pp. 823–830, sep 2016.
- [168] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, “Remote plethysmographic imaging using ambient light,” dec 2008.
- [169] H. Liu, Y. Wang, and L. Wang, “The Effect of Light Conditions on Photoplethysmographic Image Acquisition Using a Commercial Camera,” 2014.

- [170] K. Matsumura, P. Rolfe, J. Lee, and T. Yamakoshi, “iPhone 4s photoplethysmography: which light color yields the most accurate heart rate and normalized pulse volume using the iPhysioMeter Application in the presence of motion artifact?,” *PloS one*, vol. 9, no. 3, p. e91205, 2014.
- [171] J. H. Tulen, F. Boomsma, and A. J. Man in ’t Veld, “Cardiovascular control and plasma catecholamines during rest and mental stress: effects of posture,” *Clinical Science*, vol. 96, pp. 567 LP – 576, jun 1999.
- [172] J. E. Sanderson, L. Y. Yeung, D. T. Yeung, R. L. Kay, B. Tomlinson, J. A. Critchley, K. S. Woo, and L. Bernardi, “Impact of changes in respiratory frequency and posture on power spectral analysis of heart rate and systolic blood pressure variability in normal subjects and patients with heart failure.,” *Clinical science (London, England : 1979)*, vol. 91, pp. 35–43, jul 1996.
- [173] F. L. Young and A. S. Leicht, “Short-term stability of resting heart rate variability: influence of position and gender,” *Applied Physiology, Nutrition, and Metabolism*, vol. 36, pp. 210–218, apr 2011.
- [174] F. Mahananto, T. Igasaki, and N. Murayama, “Potential force dynamics of heart rate variability reflect cardiac autonomic modulation with respect to posture, age, and breathing pattern,” *Computers in Biology and Medicine*, vol. 64, pp. 197–207, sep 2015.
- [175] D. Nunan, G. Donovan, D. G. Jakovljevic, L. D. Hodges, G. R. H. Sandercock, and D. A. Brodie, “Validity and reliability of short-term heart-rate variability from the Polar S810,” *Medicine and science in sports and exercise*, vol. 41, pp. 243–250, jan 2009.
- [176] Y. Maeda, M. Sekine, and T. Tamura, “The Advantages of Wearable Green Reflected Photoplethysmography,” *Journal of Medical Systems*, vol. 35, pp. 829–834, oct 2011.
- [177] World Medical Association, “World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects.,” *Bulletin of the World Health Organization*, vol. 79, pp. 373–374, jul 2001.
- [178] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm.,” *IEEE transactions on bio-medical engineering*, vol. 32, pp. 230–236, mar 1985.
- [179] V. Ferrer-Mileo, F. Guede-Fernandez, M. Fernandez-Chimeno, J. Ramos-Castro, and M. Garcia-Gonzalez, “Accuracy of heart rate variability estimation by photoplethysmography using an smartphone: Processing optimization and fiducial point selection,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2015-Novem, 2015.

- [180] Y. C. Chiu, P. W. Arand, S. G. Shroff, T. Feldman, and J. D. Carroll, "Determination of pulse wave velocities with computerized algorithms," *American heart journal*, vol. 121, no. 5, pp. 1460–1470, 1991.
- [181] E. Kazanavicius, R. Gircys, A. Vrubliauskas, and S. Lugin, "Mathematical methods for determining the foot point of the arterial pulse wave and evaluation of proposed methods," *Information Technology and control*, vol. 34, no. 1, 2005.
- [182] M. García-González, M. Fernández-Chimeno, F. Guede-Fernández, V. Ferrer-Mileo, A. Argelagós-Palau, L. Álvarez-Gómez, E. Parrado, J. Moreno, L. Capdevila, and J. Ramos-Castro, "A methodology to quantify the differences between alternative methods of heart rate variability measurement," *Physiological Measurement*, vol. 37, no. 1, 2015.
- [183] G. A. Mack and J. H. Skillings, "A Friedman-Type Rank Test for Main Effects in a Two-Factor ANOVA," *Journal of the American Statistical Association*, vol. 75, pp. 947–951, dec 1980.
- [184] S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979.
- [185] P. Nemenyi, "Distribution-free multiple comparisons," in *Biometrics*, vol. 18, p. 263, INTERNATIONAL BIOMETRIC SOC 1441 I ST, NW, SUITE 700, WASHINGTON, DC 20005-2210, 1962.
- [186] S. W. Weinschenk, R. D. Beise, and J. Lorenz, "Heart rate variability (HRV) in deep breathing tests and 5-min short-term recordings: agreement of ear photoplethysmography with ECG measurements, in 343 subjects," *European Journal of Applied Physiology*, vol. 116, pp. 1527–1535, aug 2016.
- [187] W. Karlen, C. J. Brouse, E. Cooke, J. M. Ansermino, and G. A. Dumont, "Respiratory rate estimation using respiratory sinus arrhythmia from photoplethysmography," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1201–1204, IEEE, aug 2011.
- [188] W. L. Haskell, M. C. Yee, A. Evans, and P. J. Irby, "Simultaneous measurement of heart rate and body motion to quantitate physical activity.," *Medicine and science in sports and exercise*, vol. 25, pp. 109–115, jan 1993.
- [189] H. M. Romero-Ugalde, M. Garnotel, M. Doron, P. Jallon, G. Charpentier, S. Franc, E. Huneker, C. Simon, and S. Bonnet, "An original piecewise model for computing energy expenditure from accelerometer and heart rate signals.," *Physiological measurement*, vol. 38, pp. 1599–1615, jul 2017.
- [190] J. S. House, K. R. Landis, and D. Umberson, "Social relationships and health.," *Science (New York, N.Y.)*, vol. 241, pp. 540–545, jul 1988.

- [191] T. E. Seeman, "Social ties and health: the benefits of social integration.," *Annals of epidemiology*, vol. 6, pp. 442–451, sep 1996.
- [192] M. Pinguart and S. Sorensen, "Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis.," *Psychology and aging*, vol. 15, pp. 187–224, jun 2000.
- [193] K. S. Rook, "Social Networks in Later Life: Weighing Positive and Negative Effects on Health and Well-Being.," *Current directions in psychological science*, vol. 24, pp. 45–51, feb 2015.
- [194] D. Felmlee and R. Faris, "Interaction in Social Networks," in *Handbook of Social Psychology* (J. DeLamater and A. Ward, eds.), pp. 439–464, Dordrecht: Springer Netherlands, 2013.
- [195] S. Shacham, "A Shortened Version of the Profile of Mood States," *Journal of Personality Assessment*, vol. 47, pp. 305–306, jun 1983.
- [196] D. M. McNair, *Manual profile of mood states*. Educational & Industrial testing service, 1971.
- [197] J. Alonso, "Cols. Versión española de SF-12v2 TM," *Health Survey*, vol. 2002, 1992.
- [198] J. Garcia-Campayo, C. Sanz-Carrillo, J. A. Ibanez, S. Lou, V. Solano, and M. Alda, "Validation of the Spanish version of the SCOFF questionnaire for the screening of eating disorders in primary care.," *Journal of psychosomatic research*, vol. 59, pp. 51–55, aug 2005.
- [199] C. Costello, *Elgg 1.8 social networking*. Packt Publishing Ltd, 2012.
- [200] M.-Y. Hsieh, Y.-T. Tsai, C.-H. Hsu, C.-H. Hung, and K.-C. Li, "Design and Implementation of Multimedia Social Services on Elgg," in *Intelligent Technologies and Engineering Systems* (J. Juang and Y.-C. Huang, eds.), (New York, NY), pp. 11–18, Springer New York, 2013.
- [201] F. E. Grubbs, "Procedures for Detecting Outlying Observations in Samples," *Technometrics*, vol. 11, pp. 1–21, feb 1969.
- [202] T. Åkerstedt, "Shift work and disturbed sleep/wakefulness," *Sleep medicine reviews*, vol. 2, no. 2, pp. 117–128, 1998.
- [203] D. R. Gold, S. Rogacz, N. Bock, T. D. Tosteson, T. M. Baum, F. E. Speizer, and C. A. Czeisler, "Rotating shift work, sleep, and accidents related to sleepiness in hospital nurses.," *American journal of public health*, vol. 82, no. 7, pp. 1011–1014, 1992.

- [204] T. Åkerstedt, "Sleepiness as a consequence of shift work," *Sleep*, vol. 11, no. 1, pp. 17–34, 1988.
- [205] J. Fernandez-Mendoza, A. N. Vgontzas, I. Kritikou, S. L. Calhoun, D. Liao, and E. O. Bixler, "Natural history of excessive daytime sleepiness: role of obesity, weight loss, depression, and sleep propensity," *Sleep*, vol. 38, no. 3, pp. 351–360, 2015.
- [206] C. LaGrotte, J. Fernandez-Mendoza, S. L. Calhoun, D. Liao, E. O. Bixler, and A. N. Vgontzas, "The relative association of obstructive sleep apnea, obesity and excessive daytime sleepiness with incident depression: a longitudinal, population-based study," *International Journal of Obesity*, vol. 40, no. 9, p. 1397, 2016.
- [207] K. P. Strohl, D. B. Brown, N. Collop, C. George, R. Grunstein, F. Han, L. Kline, A. Malhotra, A. Pack, and B. Phillips, "An official American Thoracic Society Clinical Practice Guideline: sleep apnea, sleepiness, and driving risk in noncommercial drivers. An update of a 1994 Statement," *American journal of respiratory and critical care medicine*, vol. 187, no. 11, pp. 1259–1266, 2013.
- [208] A. Romero-Corral, S. M. Caples, F. Lopez-Jimenez, and V. K. Somers, "Interactions between obesity and obstructive sleep apnea: implications for treatment," *Chest*, vol. 137, no. 3, pp. 711–719, 2010.
- [209] F. M. D. de Andrade and R. P. Pedrosa, "The role of physical exercise in obstructive sleep apnea," *Jornal Brasileiro de Pneumologia*, vol. 42, no. 6, pp. 457–464, 2016.
- [210] P.-Y. Carry, P. Baconnier, A. Eberhard, P. Cotte, and G. Benchetrit, "Evaluation of Respiratory Inductive Plethysmography: Accuracy for Analysis of Respiratory Waveforms," *Chest*, vol. 111, no. 4, pp. 910–915, 1997.
- [211] M. N. Fiamma, Z. Samara, P. Baconnier, T. Similowski, and C. Straus, "Respiratory inductive plethysmography to assess respiratory variability and complexity in humans," *Respiratory Physiology and Neurobiology*, vol. 156, no. 2, pp. 234–239, 2007.
- [212] A. Anund, C. Fors, D. Hallvig, T. Åkerstedt, and G. Kecklund, "Observer Rated Sleepiness and Real Road Driving: An Explorative Study," *PLoS ONE*, vol. 8, p. e64782, may 2013.
- [213] P. Hilton, *Fundamental nursing skills*. John Wiley & Sons, 2006.
- [214] N. Rodriguez-Ibanez, M. A. Garcia-Gonzalez, M. Fernandez-Chimeno, and J. Ramos-Castro, "Drowsiness detection by thoracic effort signal analysis in real driving environments," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 6055–6058, IEEE, aug 2011.

- [215] P. Russer, “EMC measurements in the time-domain,” in *General Assembly and Scientific Symposium, 2011 XXXth URSI*, pp. 1–35, IEEE, 2011.
- [216] C. Ahlstrom, C. Fors, A. Anund, and D. Hallvig, “Video-based observer rated sleepiness versus self-reported subjective sleepiness in real road driving,” *European Transport Research Review*, vol. 7, no. 4, pp. 1–9, 2015.
- [217] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
- [218] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, pp. 159–174, mar 1977.
- [219] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000.
- [220] D. J. Hand, *Construction and assessment of classification rules*. Wiley, 1997.
- [221] M. Borenstein, L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein, “Fixed-Effect Versus Random-Effects Models,” in *Introduction to Meta-Analysis*, pp. 77–86, John Wiley & Sons, Ltd, 2009.
- [222] G. D. Ruxton, “The unequal variance t-test is an underused alternative to Student’s t-test and the Mann–Whitney U test,” *Behavioral Ecology*, vol. 17, pp. 688–690, jul 2006.