

EFFECTO DEL TAMAÑO DE VENTANA EN EL PROCESO DE
RECONOCIMIENTO DE LA ACTIVIDAD DIARIA: ESTUDIO Y MÉTODOS



Juan Manuel Gálvez Gómez
Departamento de Arquitectura y Tecnología de Computadores
Universidad de Granada

Proyecto Fin de Carrera de Ingeniería Electrónica

Septiembre 2013



UNIVERSIDAD DE GRANADA

INGENIERIA ELECTRÓNICA

EFFECTO DEL TAMAÑO DE VENTANA EN EL PROCESO DE
RECONOCIMIENTO DE LA ACTIVIDAD DIARIA: ESTUDIO Y MÉTODOS

Autor: JUAN MANUEL GÁLVEZ GÓMEZ

Directores: ORESTI BAÑOS LEGRÁN
HÉCTOR POMARES CINTAS

Departamento: ARQUITECTURA Y TECNOLOGÍA DE COMPUTADORES

Palabras clave: Reconocimiento de la actividad, Acelerómetro, Ventanado, Métrica F, Fusión

Resumen: El reconocimiento de actividades se está posicionando como una buena solución para monitorizar la ejecución de actividades a través de acelerómetros. Sin embargo, no hay un estudio previo concluyente que indique un tamaño de ventana óptimo para la monitorización de cada actividad. En este trabajo se propone el estudio del efecto del tamaño de ventana sobre el resto del proceso de inferencia y detección de actividades. Asimismo, y a partir de los resultados obtenidos, se plantea la definición e implementación de un sistema que haga uso de métodos de optimización del reconocimiento en base a dicho tamaño de segmentación de los datos.

KEYWORDS: Activity Recognition, Accelerometer, Windowing, F-measure, Fusion

Abstract: Activity recognition is positioning as a good solution to monitor activities execution through accelerometers. However, there is not any conclusive previous study indicating the optimum window size for monitoring the activity. In this work it is proposed to study the effect of window size on the rest of the process of inference and detection activities. Likewise, from the obtained results, it is suggested the definition and implementation of a system that makes use of optimization recognizing methods based on this size of data segmentation.



UNIVERSIDAD DE GRANADA

INGENIERIA ELECTRÓNICA

PROYECTO FIN DE CARRERA

AUTOR: JUAN MANUEL GÁLVEZ GÓMEZ

TÍTULO: EFECTO DEL TAMAÑO DE VENTANA EN EL PROCESO DE RECONOCIMIENTO DE LA ACTIVIDAD DIARIA: ESTUDIO Y MÉTODOS

TRIBUNAL :

D/Dña. MIGUEL DAMAS HERMOSO (ATC)

D/Dña. FRANCISCO GÓMEZ MULA (ATC)

D/Dña. JAVIER RAMÍREZ PÉREZ DE HINESTROSA (TSTC)

CALIFICACIÓN:

Presentado en Granada a 20 de SEPTIEMBRE de 2013

Evaluado en Granada a 24 de SEPTIEMBRE de 2013

El Presidente

El Vocal

El Secretario



UNIVERSIDAD DE GRANADA

INGENIERIA ELECTRÓNICA

AUTORIZACIÓN DE LECTURA DE
TRABAJO FIN DE CARRERA

D. ORESTI BAÑOS LEGRÁN y D. HÉCTOR POMARES CINTAS, profesores del Departamento de ARQUITECTURA Y TECNOLOGÍA DE COMPUTADORES de la Universidad de Granada, como director/es del Trabajo Fin de Carrera titulado “EFECTO DEL TAMAÑO DE VENTANA EN EL PROCESO DE RECONOCIMIENTO DE LA ACTIVIDAD DIARIA: ESTUDIO Y MÉTODOS” y realizado por el alumno D. JUAN MANUEL GÁLVEZ GÓMEZ

CERTIFICA/N: que el citado Trabajo Fin de Carrera, ha sido realizado y redactado por dicho alumno y autorizan su presentación.

Granada,

Fdo.

Fdo.

ORESTI BAÑOS LEGRÁN

HÉCTOR POMARES CINTAS



UNIVERSIDAD DE GRANADA

INGENIERIA ELECTRÓNICA

AUTORIZACIÓN DE DEPÓSITO EN LA BIBLIOTECA

Yo, D/Dña. JUAN MANUEL GÁLVEZ GÓMEZ con DNI 71559059-X, autor del Trabajo Fin de Carrera titulado “EFECTO DEL TAMAÑO DE VENTANA EN EL PROCESO DE RECONOCIMIENTO DE LA ACTIVIDAD DIARIA: ESTUDIO Y MÉTODOS” realizado en la Universidad de Granada

AUTORIZO: al depósito de dicho Trabajo en la Biblioteca de la Universidad de Granada, y de la visualización a través de Internet.

Granada,

Fdo. D/Dña.

JUAN MANUEL GÁLVEZ GÓMEZ

A mis padres Manuel y María Teresa,
y a mi hermana Tere

Prefacio

“Fuertes razones hacen fuertes acciones”

William Shakespeare

Agradecimientos

Quería mostrar mi agradecimiento al Departamento de Arquitectura y Tecnología de Computadores por haberme ofrecido la posibilidad de realizar este proyecto. En particular, a mis tutores y directores Oresti Baños Legrán y Héctor Pomares Cintas, que me mostraron su absoluta confianza para llevar a cabo el estudio y desarrollo de este proyecto.

En especial, deseo manifestar mi satisfacción por haber vuelto a trabajar con mi tutor y compañero Oresti Baños Legrán. Me ha brindado todo su apoyo, disposición y parte de su tiempo para conseguir sacar adelante este trabajo. Espero que este proyecto esté a la altura de la confianza y las expectativas que ha depositado en mí.

No puedo olvidarme bajo ningún concepto de mi familia, en especial a mis padres, pues gracias a ellos he podido alcanzar todas mis metas con su opinión, esfuerzo y apoyo incondicional. Por valorar mi esfuerzo, por comprenderme, y sobre todo, por arroparme y animarme en los peores momentos. Por darme la oportunidad de ser una persona con formación y con estudios. Gracias por vuestra paciencia, os estoy y estaré eternamente agradecido.

También me acuerdo de muchas personas que han pasado por mi vida y me ofrecieron siempre su apoyo. Desde amigos de la infancia hasta amigos y compañeros de carrera, les debo grandes momentos vividos y tiempos duros de trabajo compartido. A todos ellos, les dedico este proyecto fin de carrera, que cierra un capítulo de mi vida y espero que abra otro de sorpresas, mucha ilusión, proyectos y avances.

Índice general

Índice de figuras	ix
Índice de tablas	xv
1. Capítulo I: Introducción	1
1.1. Visión General	1
1.2. Motivación	2
1.3. Objetivos	3
1.4. Estado del arte	4
1.5. Estructura de la memoria	12
2. Capítulo II: Reconocimiento de la actividad	13
2.1. Introducción	13
2.2. Modalidades de sensado: sensores para el reconocimiento de actividades. Limitaciones y dificultades	13
2.3. La influencia de los factores exógenos	14
2.4. La cadena de reconocimiento de actividades	15
2.5. Configuración, métodos y datos empleados en el estudio	17
2.5.1. Dataset para el reconocimiento de actividades	18
2.5.2. Modelos y configuración experimental	20
2.5.2.1. Adquisición de datos	20
2.5.2.2. Preprocesamiento	20
2.5.2.3. Segmentación	20
2.5.2.4. Extracción de características	21
2.5.2.5. Clasificación	21

3. Capítulo III: Estudio del efecto del tamaño de ventana en el reconocimiento de la actividad	23
3.1. Introducción	23
3.2. Explicación del estudio	24
3.3. Evaluación del estudio	24
3.4. Resultados y discusión	28
3.5. Conclusiones	67
4. Capítulo IV: Fusión multiventana	69
4.1. Introducción	69
4.2. Descripción del algoritmo	70
4.3. Descripción del estudio. Análisis	73
4.4. Resultados y discusión	76
4.5. Conclusiones	86
5. Capítulo V: Conclusiones generales y trabajo futuro	87
5.1. Introducción	87
5.2. Conclusiones	87
5.3. Aplicabilidad y viabilidad	90
5.4. Ventajas e inconvenientes	92
5.5. Posibles mejoras y trabajo futuro	92
Referencias	95

Índice de figuras

2.1.	Cadena múltiple de reconocimiento de la actividad (M-ARC). M sensores proporcionan señales en bruto (μ_j) que son subsecuencialmente procesadas (p_j). Las señales son k -particionadas (s_{jk}) y un conjunto de características (habitualmente definidas como f) son extraídas de ellas, posiblemente de manera diferente para cada cadena. El vector de características es usado como entrada al clasificador de entidades. Cada clasificador produce una clase en un problema con N -clases que puede ser combinado a través de un método de decisión basado en fusión. Los índices son respectivamente definidos por $j = 1, \dots, M$; $k = 1, \dots, K$; $i = 1, \dots, N$.	15
2.2.	Ubicación de los 9 sensores en diferentes posiciones del cuerpo	18
3.1.	Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función del tamaño de ventana (en segundos)	28
3.2.	Tamaño de ventana mínimo necesario para un F1-score determinado según la metodología seleccionada (algoritmo de clasificación-características empleadas): a) DT-FS1 b) DT-FS2 c) DT-FS3 d) NB-FS1 e) NB-FS2 f) NB-FS3	39
3.3.	Tamaño de ventana mínimo necesario para un F1-score determinado según la metodología seleccionada (algoritmo de clasificación-características empleadas): g) NCC-FS1 h) NCC-FS2 i) NCC-FS3 j) KNN-FS1 k) KNN-FS2 l) KNN-FS3	40
3.4.	Tamaño mínimo de ventana para diversas tolerancias de F_1 - score sobre las metodologías de DT: a) DT-FS1 b) DT-FS2 c) DT-FS3. En la fila superior se indican los F_1 - score máximos alcanzados para cada metodología	43
3.5.	Tamaño mínimo de ventana para diversas tolerancias de F_1 - score sobre las metodologías de NB: d) NB-FS1 e) NB-FS2 f) NB-FS3. En la fila superior se indican los F_1 - score máximos alcanzados para cada metodología	44

3.6. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ sobre las metodologías de NCC: g) NCC-FS1 h) NCC-FS2 i) NCC-FS3. En la fila superior se indican los $F_1 - score$ máximos alcanzados para cada metodología	45
3.7. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ sobre las metodologías de KNN: j) KNN-FS1 k) KNN-FS2 l) KNN-FS3. En la fila superior se indican los $F_1 - score$ máximos alcanzados para cada metodología	46
3.8. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>andar, trotar y correr</i> en orden descendente	51
3.9. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>saltar hacia arriba, saltar adelante y atrás y saltar lateralmente</i> en orden descendente	52
3.10. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>saltar abriendo y cerrando piernas y brazos, saltar la cuerda y girar el tronco con los brazos extendidos</i> en orden descendente	53
3.11. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>girar el tronco con los codos doblados, doblar la cintura hacia delante y rotación de cintura</i> en orden descendente	54
3.12. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>alcanzar un pie con la mano contraria y la cintura doblada, alcanzar los talones flexionando las rodillas y doblarse lateralmente</i> en orden descendente	55

3.13. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>doblarse lateralmente con un brazo arriba, doblar la cintura hacia delante repetidamente y giro de extremidades superiores e inferiores en oposición</i> en orden descendente	56
3.14. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>elevación lateral de brazos, elevación frontal de brazos y palmadas frontales</i> en orden descendente	57
3.15. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>cruzar brazos frontalmente, rotación intensa de hombros y rotación moderada de hombros</i> en orden descendente	58
3.16. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>rotación interna de brazos, rodillas al pecho (alternativamente) y talones a la espalda (alternativamente)</i> en orden descendente	59
3.17. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>doblar las rodillas agachándose, doblar las rodillas hacia delante(alternativamente) y rotación de rodillas</i> en orden descendente	60

3.18. Tamaño mínimo de ventana para diversas tolerancias de $F_1 - score$ de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de $F_1 - score$ exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: <i>remar</i> , <i>hacer bicicleta elíptica</i> y <i>hacer bicicleta estática</i> en orden descendente	61
4.1. Estructura del clasificador HWC. Problema con N actividades y M sensores . . .	70
4.2. Esquema <i>ARC</i> adaptado al multiventanado. Se realiza una adecuación de la señal adquirida para conseguir una partición correcta en la etapa de segmentación de las K ventanas seleccionadas (múltiplos de la ventana de mayor tamaño, de tamaño N) y se multiplican índices de control para cada tamaño de ventana seleccionado. Se ponderan las K decisiones con una métrica seleccionada y se realiza la fusión de decisiones tras la etapa de clasificación, obteniendo una decisión única que es aceptada como actividad reconocida	71
4.3. Ejemplo de funcionamiento del algoritmo de fusión a 1 nivel y 2 niveles. Se muestran las decisiones parciales adoptadas por $K = 5$ ventanas en la parte superior para las actividades reconocidas A, B y C. En la parte inferior se muestra una tabla de $F_1 - score$ correspondiente al reconocimiento de cada actividad para un tamaño de ventana específico $W1-W5$. La decisión final de la fusión a 2 niveles muestra la clasificación de la actividad como A, mientras que el ventanado tradicional reconocería la actividad como B. Este ejemplo está adaptado para la utilización de fusión completa expuesta en este proyecto: fusión simultánea de las ventanas de 0.25, 0.75, 1.5, 3 y 6 segundos	75
4.4. Eficiencia de clasificación (en%) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de las combinaciones C1-C10 escogidas para el testeo del algoritmo de fusión a 2 niveles. Todas las combinaciones utilizan el tamaño de ventana de 6 segundos (ver tabla 4.1)	78
4.5. Eficiencia de clasificación (en%) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de las combinaciones C11-C15 escogidas para el testeo del algoritmo de fusión a 2 niveles. Todas las combinaciones utilizan el tamaño de ventana de 3 segundos (ver tabla 4.1)	80

4.6. Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de 8 ventanas seleccionadas para el testeo del algoritmo de fusión a 1 nivel.	83
---	----

Índice de tablas

3.1. Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando ventanado variable en el rango 0-7 segundos. En cada fila se incluyen los resultados para un tamaño de ventana específico. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, el mayor rendimiento obtenido para cada tamaño de ventana	29
3.2. Obtención de la variación (en %) entre la aplicación de un tamaño de ventana de 6 segundos y el tamaño óptimo que maximiza la eficiencia de clasificación para cada metodología planteada. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados. En las filas se indica el tamaño de ventana óptimo para maximizar el rendimiento con cada metodología planteada y el porcentaje de variación existente con un tamaño de ventana de 6 segundos. El signo (+) indica que se obtiene una mejora con respecto al tamaño de ventana de 6 segundos	32
3.3. Obtención de la variación (en %) entre la aplicación de un tamaño de ventana de 6 segundos y el tamaño de ventana mínimo que supera la eficiencia del tamaño de ventana de referencia (6 segundos) para cada metodología planteada. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En las filas se indica el tamaño de ventana mínimo para superar la eficiencia de clasificación conseguida con 6 segundos y el porcentaje de variación existente con un tamaño de ventana de 6 segundos. El signo (+) indica que se obtiene una mejora con respecto al tamaño de ventana de 6 segundos	33

3.4.	$F_{1,máx} - score$ de las 12 metodologías planteadas para un tamaño de ventana especificado en la tabla 3.5. En cada fila se incluyen los resultados para una actividad específica del dataset estudiado. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3)	35
3.5.	Tamaños de ventana asociados a los $F_{1,máx} - score$ de las 12 metodologías planteadas en la tabla 3.4. En cada fila se incluyen los resultados para una actividad específica del dataset estudiado. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3)	36
3.6.	$F_1 - score$ mínimo obtenido para cualquier actividad del dataset para las 12 metodologías planteadas aplicando un tamaño de ventana específico. En cada fila se incluyen los resultados para un tamaño de ventana del rango de estudio (0-7 segundos). En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el $F_1 - score$ mínimo que permite seleccionar el menor tamaño de ventana posible optimizando el $F_1 - score$. En cursiva, el $F_1 - score$ mínimo que permite seleccionar el menor tamaño de ventana posible considerando una penalización máxima ~ 0.05 al $F_1 - score$ optimizado	66
4.1.	Tabla de combinaciones de ventanas aplicadas para el testeo del algoritmo de fusión multiventana a 2 niveles	76
4.2.	Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento cada 6 segundos. Las ventanas seleccionadas son múltiplos de la ventana de mayor tamaño (6 segundos). En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada combinación aplicada	77

- 4.3. Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento cada 3 segundos. Las ventanas seleccionadas son múltiplos de la ventana de mayor tamaño (3 segundos). En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En **negrita**, el mayor rendimiento posible para cada metodología. En *cursiva*, la metodología con el mayor rendimiento para cada combinación aplicada 79
- 4.4. Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento en tiempos reducidos. Se aplican ventanas de 1.5, 0.75 y 0.25 segundos para estas combinaciones. En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En **negrita**, el mayor rendimiento posible para cada metodología. En *cursiva*, la metodología con el mayor rendimiento para cada combinación aplicada 81
- 4.5. Eficiencia de clasificación (en %) de las 12 metodologías planteadas sin aplicar fusión. Las ventanas seleccionadas son múltiplos del tamaño de ventana de referencia (6 segundos). En cada fila se incluyen los resultados de la aplicación de un tamaño de ventana específico en la etapa de segmentación. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En **negrita**, el mayor rendimiento posible para cada metodología. En *cursiva*, la metodología con el mayor rendimiento para cada tamaño de ventana aplicado 82
- 4.6. Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 1 nivel. Las ventanas seleccionadas son múltiplos del tamaño de ventana de referencia (6 segundos). En cada fila se incluyen los resultados de la aplicación de un tamaño de ventana específico en la fusión. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En **negrita**, el mayor rendimiento posible para cada metodología. En *cursiva*, la metodología con el mayor rendimiento para cada tamaño de ventana aplicado . . 82

4.7. Comparación en términos de eficiencia de clasificación (en %) de los mejores resultados alcanzados para las 12 metodologías planteadas entre la aplicación de ventanado tradicional (T), la aplicación del algoritmo de fusión a 1 nivel (F1) y la aplicación del algoritmo de fusión a 2 niveles (F2) con decisión final cada 6 segundos. En cada fila se incluyen los mejores resultados de cada opción. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada opción considerada 85

1

Capítulo I: Introducción

1.1. Visión General

La tendencia social en la búsqueda del mayor bienestar posible implica la inserción de mecanismos de control, alerta o aviso normalmente basados en el análisis del comportamiento humano. Los sistemas de reconocimiento de la actividad pueden ser utilizados con tal propósito. No obstante, cada persona se comporta de forma diferente en su vida diaria.

En términos cinemáticos, cada actividad se ejecuta de diferente forma (energía empleada, posición corporal, velocidad, etc.) e incluso una misma persona puede variar el modo de realizar sus actividades rutinarias y cotidianas a lo largo de su vida. En este punto, se puede pensar que la existencia de factores de diferente índole como la edad, el peso, la altura u otros como padecer alguna enfermedad o sufrir algún tipo de discapacidad, son motivos suficientes para que exista ese cambio en la realización de nuestras actividades. Adicionalmente, cada actividad se encuentra definida por diferentes características que la identifican unívocamente por los mecanismos de ejecución instintivos que la persona realiza para su ejecución.

El colapso de los servicios sociales y médicos en las sociedades de bienestar invitan a la inserción y aceptación de mecanismos que monitoricen nuestras acciones y permitan adoptar decisiones frente a eventualidades de forma controlada, sin requerir una asistencia presencial continua. Más allá de las propias necesidades que las personas pueden terminar creándose con la irrupción e innovación tecnológica, realmente existen necesidades que deben ser cubiertas.

Desde esta perspectiva, los sistemas de reconocimiento de actividades se están posicionando como una solución idónea a la problemática planteada.

1.2. Motivación

El reconocimiento de la actividad ha irrumpido como una novedosa línea de investigación tecnológica. Los sistemas *AR* (Activity Recognition) están siendo utilizados como una potente herramienta de monitorización en múltiples campos como la sanidad, la asistencia, el deporte, etc. y otros relacionados directamente con el ámbito tecnológico como la seguridad, la robótica, la industria de los videojuegos, etc. De especial interés, tanto personal como social, resulta su aplicabilidad en tiempo real para la determinación del *health-status* de una persona. Un abanico de múltiples posibilidades que permite la realización de estudios de prevención, pro-actividad o alerta tanto a nivel médico como deportivo.

El envejecimiento de la población o la congestión de los centros médicos han convertido la salud en un problema insostenible. Esta tendencia hace cada vez más necesaria la disposición de sistemas de monitorización de parámetros biomédicos que permitan una supervisión asistencial del estado de salud de cualquier paciente por parte de su equipo médico. De igual forma, la monitorización de parámetros fisiológicos continúa siendo un problema por diferentes motivos. El paciente padece cierta incomodidad causada por mecanismos invasivos en su reconocimiento, aunque ya existan muchos avances en aplicación de métodos y técnicas no invasivas, además del coste tanto económico como temporal que requiere dicho proceso. En este sentido, la adquisición de datos biomédicos a partir de dispositivos inalámbricos aportan muchas ventajas (reducción del tiempo de actuación en caso de urgencia, descentralización del control del centro hospitalario, seguimiento médico a distancia, disminución de las listas de espera, etc.).

Sin duda, gran parte de los avances tecnológicos actuales en la medicina han sido gracias a la puesta en práctica de sistemas de reconocimiento de actividades. Para una disciplina como la teleasistencia ha supuesto un impulso en diversos aspectos. La teleasistencia tiene como finalidad ayudar a que las personas mayores o discapacitadas permanezcan el mayor tiempo posible en sus hogares de manera independiente. Al valor añadido de la autonomía, la incorporación de los sistemas *AR* permite mejorar el control de la persona a partir de la monitorización de sus actividades rutinarias e identificar alguna anomalía puntual (caídas). También ha encontrado un gran aliado en su implementación como parte de la rehabilitación de una persona. El buen reconocimiento de actividades concretas (andar, correr, saltar, empujar, etc.) permite su aplicación en la recuperación de la movilidad de la extremidad dañada (brazo, tobillo, muñeca, muslo, etc.). De manera extensible, puede ayudar en el ámbito deportivo a la supervisión, control y corrección, así como recuperación del gesto deportivo.

A las motivaciones generales que despierta la utilización de sistemas de reconocimiento de la actividad se unen las motivaciones técnicas. En concreto, la consideración e investigación

de posibles mejoras en cualquiera de las fases de un sistema de reconocimiento. Hasta ahora, los sistemas *AR* han funcionado perfectamente bajo condiciones ideales con la utilización de sensores fijos o la realización de actividades muy concretas bajo un patrón de ejecución específico e invariante. En una aplicación real, cualquier resultado extrapolado de dichos estudios podría ser erróneo o al menos impreciso.

Una motivación específica surge con la revisión de la etapa de segmentación del sistema de reconocimiento de actividades: *windowing* (ventanado) o *segmentation* (segmentación). Gran parte de los trabajos existentes sobre esta temática indican que 6 segundos es un tiempo suficiente para reconocer con cierta eficiencia cualquier actividad humana en un dataset amplio (15-20 o más actividades diferentes). Por otra parte, existen múltiples estudios que tratan un número reducido de actividades y que demuestran que se pueden conseguir altas tasas de reconocimiento aplicando tamaños de ventana pequeños (0.25-2 segundos). En este punto, se plantea la posibilidad de realizar un estudio más profundo sobre el efecto del tamaño de ventana en la eficiencia de reconocimiento final a partir de un conjunto de actividades extenso. Realmente, no existe un estudio previo de esta magnitud que ofrezca una optimización del reconocimiento basándose en el tamaño de ventana ni tampoco se ha comprobado qué influencia implícita se produce sobre el rendimiento. Una hipotética influencia de esta consideración marcaría el desarrollo de este proyecto.

1.3. Objetivos

A continuación, se presentan una serie de objetivos a alcanzar a la finalización del presente proyecto. En términos generales, se pretende el estudio e identificación de posibles mejoras sobre el procesamiento de la señal de un sistema de reconocimiento de la actividad humana centrandose toda la atención en conseguir mejorar el tratamiento de los datos en la etapa de segmentación de un sistema *AR*.

En este punto, se plantean una serie de objetivos observando el contexto de los temas tratados:

- ★ Estudiar el estado del arte en lo referente a diferentes tendencias actuales en el reconocimiento de actividades
- ★ Indicar las principales limitaciones, dificultades y factores de influencia en el desarrollo de un proyecto que aborda la temática del reconocimiento de la actividad
- ★ Seleccionar los datos biomecánicos que serán objeto de análisis, estudio y procesamiento en el presente estudio

- ★ Estudiar la influencia del tamaño de la ventana sobre la eficiencia del sistema de reconocimiento de la actividad diaria
- ★ Investigar y seleccionar alguna métrica que permita evaluar la eficiencia del reconocimiento de una actividad entre el conjunto de actividades planteadas en el estudio
- ★ Definir e implementar alguna estrategia o algoritmo que utilice el particionamiento de datos en diferentes tamaños de ventana para una toma de decisiones
- ★ Analizar la eficiencia de reconocimiento en función del tamaño de ventana y la aplicación de la estrategia implementada
- ★ Analizar e interpretar los resultados obtenidos
- ★ Motivar trabajo futuro a partir de las conclusiones obtenidas de la elaboración de este proyecto

1.4. Estado del arte

En el reconocimiento de actividades, existen muchos trabajos que han conseguido fructificar en la discriminación de un subconjunto de actividades, realizadas bajo patrones concretos y cuyos datos eran recogidos mediante una serie de sensores específicos, bien localizados y colocados en ciertas partes del cuerpo humano. Todos estos condicionantes han convertido la monitorización de las actividades humanas en una supervisión que, en la realidad, no permite recoger lo imprevisible y natural del comportamiento humano. Las tendencias más actuales en el reconocimiento de actividades tienen en cuenta estas observaciones, en los que se habla de *conceptdrift*, *adaptation*, *self-calibration*, etc.

Uno de los primeros trabajos en este campo es el presentado por Bao e Intille (1) en 2004. En dicha contribución se consigue reconocer 20 actividades (desde andar o correr hasta subir las escaleras o cepillarse los dientes) utilizando 5 acelerómetros biaxiales (colocados en cadera, muñeca, brazo, muslo y tobillo) en un total de 20 sujetos. El estudio de Ermes et al.(2) de 2008 utiliza únicamente 2 acelerómetros (en cadera y muñeca), aplicado sobre 12 sujetos y 9 actividades (como permanecer tumbado o jugar al fútbol). Maurer et al. (3) prueban en 2006 la utilización exclusiva de 1 acelerómetro en la muñeca para monitorizar 6 actividades básicas (sentarse, andar, correr, permanecer de pie, subir y bajar escaleras) a partir de 6 sujetos, alcanzando una eficiencia del 87% con árboles de decisión. 3 algoritmos de clasificación basados en energía y otras características son probados por Pham y Abdelzaher (4) en 2008. Su autor emplea los datos recogidos por Ganti et al. (5) (2006), en cuyo trabajo monitoriza 4 actividades (montar en bicicleta, mecanografiar, andar y escribir) a partir de 5 acelerómetros biaxiales sobre

2 sujetos. El trabajo de Yang et al. (6) (de 2008) alcanza eficiencias de clasificación cercanas al 95 % a partir de redes neuronales multicapa adaptadas como clasificadores para el reconocimiento de actividades humanas y empleando la idea de 'divide y vencerás' para discernir entre actividades dinámicas y estáticas. Dalton y O'Laighin (7) presentaron un estudio en 2009 sobre 19 actividades diferentes clasificadas como *básicas*, *transitorias* y *utilizando instrumentos*, con la ayuda de 25 sujetos y evaluando 14 clasificadores diferentes, obteniendo la mayor eficiencia con AdaBoostM1 (C4.5 Graft) de un 95 % mediante la técnica de eximir a un sujeto del entrenamiento.

No obstante, la mayoría de los sistemas empleados en los estudios anteriores no están preparados para ser usados en condiciones realistas. Los sensores han sido utilizados sobre un conjunto de sujetos bajo supervisión, de manera que no se puede asegurar que los resultados obtenidos para dicho subconjunto sean extrapolables a diferentes personas y bajo condiciones naturalistas. Estos dispositivos, bajo tales condiciones, han sido empleados con una configuración determinada buscando la mayor similitud posible entre diferentes sesiones de experimentación. La repercusión inmediata es la posibilidad de ser probados bajo diferentes configuraciones. Además, lo más probable es que los sensores sufran una serie de problemas (descalibraciones, desconexiones, etc.) que deben ser tenidos en cuenta para no afectar el reconocimiento de la actividad monitorizada. También debe considerarse, en una monitorización totalmente libre, que el sujeto pueda encontrarse con obstáculos durante la realización de una actividad (puertas cerradas, vallas, etc.) que en una monitorización de laboratorio se elimina. Todas estas observaciones influyen directamente sobre la obtención de los datos monitorizados y en una desviación, cambio o evolución de dichos datos que en el panorama científico se conoce como *concept drift*.

Algunos de estos factores ya han sido considerados en trabajos anteriores. Baños et al. (8) detectan anomalías técnicas que se obtienen a partir de las medidas de los sensores (offset, ausencia de señal, ruido electrónico, etc.) en 2012. La influencia de la orientación en la colocación de los propios sensores sobre la eficiencia del reconocimiento de actividades es demostrada por Kunze y Lukowicz (9) en 2008. Otros trabajos desarrollan un algoritmo adaptativo que recalcula el espacio de características alterado por desplazamientos y rotaciones en los sensores. Destacan las aportaciones realizadas por Chavarriaga et al. (10) y Bayati et al. (11), ambos en 2011. A través del proyecto OPPORTUNITY, Roggen et al. (12) presentan en 2009 la necesidad de adaptar el sistema de reconocimiento a partir de la interpretación de los datos colectados de diferentes usuarios. Wang et al. (13) plantean una red de sensores inalámbricos para el reconocimiento de actividades humanas basándose en la coincidencia de patrones de movimiento en el año 2005. Se trata de un estudio muy específico que permite reconocer 3 actividades

concretas (llamar por teléfono, beber y escribir) utilizando una colocación de los sensores exacta (bajo el teléfono, bajo una copa y alrededor de un lápiz).

Para conseguir contrarrestar estos hechos (en relación al *concept drift*), existe una vertiente actual y novedosa que propone conceptos relacionados con la *adaptación* y *auto-calibración* de los sistemas de reconocimiento de actividades. Todos estos conceptos ya han sido aplicados con anterioridad en otras áreas científicas. Zliobaite y Pechenizkiy (14) ofrecieron en 2010 una visión general sobre las principales áreas de aplicación y detección de *concept drift*, que se ha convertido en un concepto recurrente y considerado en algunos ámbitos como la minería de datos. En este campo, Carmona y Gavaldá (15) en 2012 abordan dicho paradigma ampliamente. Para su detección y gestión, presenta un mecanismo online basado en interpretaciones abstractas y muestreo secuencial en conjunción con técnicas de aprendizaje novedosas. Su detección en tiempo real resulta crítica si se trata sobre señales vitales (ECG, EEG, SpO2, BP, etc.), tal y como apuntan Patil et al. (16) en 2012. En los últimos tiempos, ha cobrado también interés en tareas como la clasificación de imágenes. Mediante la aplicación de aprendizaje profundo (*deep learning*), Calandra et al. (17) en 2012 indican que los análisis de datos deben realizarse directamente sobre los datos recogidos in-situ, permitiendo ver cambios temporales y naturales.

A nivel de *adaptación*, varios ámbitos han tenido este factor muy presente. La irrupción de las tecnologías de la información se han convertido en otro agente que interactúa con el ser humano e implica una adaptación evolutiva en su entorno según comentan Smit y Wandel (18) en 2006. Se trata de un reto que pone a prueba las capacidades adaptativas de las poblaciones humanas. En lo que respecta al campo tecnológico, se ha convertido en un tópico recurrente sobre sistemas autoadaptativos (como el tratamiento y evolución del software tratado por Ghezzi (19) en 2012 o computación en paralelo sobre sistemas embebidos tratado por Guang et al. (20) ese mismo año), robótica (como estrategias de control de algoritmos presentados por Jimenez-Fabian y Verlinden (21) en 2012) e incluso tecnología reflectiva (como los sistemas orientados y basados en *context awareness* y *pervasive adaptation* por Serbedzija y Beyer (22) en 2011). También se ha utilizado este concepto para definir sistemas inteligentes que evalúan su impacto sobre el propio sistema, como la propuesta de Magnisalis et al. (23) en 2011.

Se pueden enumerar diferentes aportaciones que se han realizado para las distintas fases de la cadena en un sistema de reconocimiento. Ahn y Lee (24) proponen un umbral adaptativo durante la detección de voz como mecanismo de pre-procesamiento en el año 2006. Las simulaciones demuestran la mejora del rendimiento con el algoritmo propuesto en medios ruidosos frente a otros. En 2011, una técnica para estimar la densidad espectral de potencia en reconocimiento de voz es planteada por Ravan y Beheshti (25). Implementa un método con ventana adaptativa, que tradicionalmente se ha obtenido mediante aproximaciones con *MelFrequency*

Cepstral Coefficients (MFCC). La eficiencia empleando el clasificador de los K-Vecinos (KNN) mejora en un 12 %. Así mismo, Kan et al. (26) indican una adaptación del tamaño de ventana para la extracción de características del espectro óptico durante diagnósticos médicos en el año 2010. La técnica del ventanado adaptativo mejora el rendimiento frente al ventanado de tamaño fijo. Este mismo año, Ros et al. (27) analizan los comportamientos humanos mediante el aprendizaje de autómatas y propone adaptaciones a los cambios existentes en el medio basándose en el conocimiento adquirido.

En relación al reconocimiento de actividades, ya existen algunos estudios que consideran este aspecto. Misra y Lim (28) demuestran en 2011 que cambiando la frecuencia de adquisición de los datos recogidos por los sensores de un teléfono móvil, se puede reducir el gasto de su batería hasta en un 70 %. En esta misma línea, Rachuri et al. (29) indican una metodología adaptativa para muestrear según el historial de adquisición de datos, consiguiendo reducir el gasto de batería y penalizando el muestreo agresivo durante la obtención de datos en el año 2012. Ese mismo año y sobre un smartphone, Siirtola y Rönning (30) abordan todas las fases de la cadena del reconocimiento de actividades, además de plantear el debate entre reconocimiento offline y en tiempo real. Consigue eficiencias de clasificación con *Quadratic Discriminant Analysis* (QDA) superiores a KNN (95.8 % frente a 93.9 %) Otros autores formulan el problema de la adaptación en el reconocimiento de actividades humanas como un aprendizaje bajo cambio covariable (*covariate shift*), hasta el punto de proponer un método de clasificación probabilística eficiente computacionalmente basado en un muestreo adaptativo. Un claro ejemplo es el trabajo presentado por Hachiya et al. (31) en 2012. De igual forma, en 2012, Anguita et al. (32) implementan una adaptación del clasificador SVM para reducir costes computacionales manteniendo la eficiencia de clasificación. Esta aproximación viene motivada por las limitaciones de los teléfonos móviles en términos de energía y potencia de computación. En 2011, Augustyniak (33) ha desarrollado un estudio que permite reconocer patrones inusuales en el comportamiento humano a partir de la adaptación o compensación métrica basada en desplazamientos temporales entre patrones. Utiliza parámetros como el ritmo cardíaco, la postura del cuerpo, la aceleración de la muñeca o sonidos recogidos de la realización de cualquier actividad planificada. El algoritmo de similaridad empleado es DTW (*Dynamic Time Warping*).

A nivel de *auto-calibración*, existen precedentes en algunos campos de investigación y su aplicación también se ha convertido en primordial sobre cualquier sistema. Lotters et al. (34) indican en 1998 que la calibración es un procedimiento fundamental y recoge las primeras iniciativas en el campo de la biometría para conseguirlo. Lackner y DiZio (35) hablan en dicho trabajo, del año 2000, sobre la adaptación del cuerpo a partir de las fuerzas transitorias de Coriolis, de manera que debe entenderse la *auto-calibración* como un proceso de *adaptación*.

En las últimas 2 décadas, su consideración ha sido de suma importancia en el tratamiento con cámaras, tal y como presenta Sturm (36) en 1997 y sigue siendo un tema de actualidad a día de hoy, como demuestran los estudios de Espuny (37) (2012), Ackermann et al. (38) (2012), y Jin y Li (39) este mismo año. Diferentes tipos de sistemas han tenido en cuenta este factor en sus diseños: sistemas autónomos genéricos (como presentan Javed et al. (40) en 2012) u orientados a sistemas de navegación inercial (como el trabajo de Zhang et al. (41) de 2012), sistemas sensoriales (por ejemplo, el trabajo de Van de Ven et al. (42) de 2012), sistemas de control (el estudio de Bartolini et al. (43) este mismo año), etc.

Este mecanismo ya ha sido considerado también en distintas fases de la cadena del reconocimiento de actividades. En 2008, Kunze y Lukowicz (9) introducen una serie de heurísticas que influyen en el rendimiento del clasificador. Su estudio demuestra que el desplazamiento del sensor puede suponer hasta un 14% en el descenso de la tasa de eficiencia, de manera que la calibración pasa por ser un problema importante a resolver. Para conseguir una calibración automática de los sensores durante su operación, Gietzelt et al. (44) desarrollan en 2008 una técnica sin requerir un conocimiento específico sobre la orientación de los propios sensores. Su funcionamiento está basado en el reconocimiento de la actividad que se está ejecutando y se consiguen errores similares a los que introducen las técnicas tradicionales para calibración. Forster et al. (45) implementan en 2009 un algoritmo de autocalibración aplicado sobre la fase de clasificación. Las mejoras alcanzadas resultan relevantes según los escenarios presentados por el autor: un aumento del 33.3% sobre una interfaz de ordenador y un 13.4% con ejercicios de gimnasio. La deriva introducida en los acelerómetros es considerada por Gao et al. (46) en 2011 como una desviación lineal, de manera que la complejidad de la computación para su calibración se ve reducida. En 2011, Roggen (47) presenta estudios basados en la robótica para interpretar y alcanzar la mayor robustez posible durante el reconocimiento de actividades humanas. Entre otros aspectos, trata la *auto-calibración* como un factor importante que debe ser considerado y tratado en esta temática. Incluso, la descalibración es considerada como una de las anomalías más críticas y presentes en el tratamiento con sensores inerciales. Para reducir su impacto en el reconocimiento final de las actividades, Baños et al. (48) aplican la técnica de fusión de sensores en el año 2012.

Nuestra contribución principal se relaciona con *Windowing / Segmentation*, una metodología que ya ha sido empleada en diferentes ámbitos científicos y que guarda cierta relación con los conceptos introducidos (*concept drift*, *adaptación* y *auto-calibración*) En el año 2000, Klinkenberg and Joachims (49) consiguen detectar *concept drift* mediante la utilización de SVM para filtrados de información. Este hecho permite ajustar el tamaño de la ventana conforme se minimiza el error estimado. En lo referente al tratamiento de señales acústicas, Sayed-Mouchaweh

et al. (50) tratan en su estudio de 2011 el problema de *concept drift*. Para la mejora en el rendimiento del clasificador, propone un tamaño de ventana adaptativo según el cambio detectado (lento, moderado o rápido). Silic y Basic (51) basan su trabajo de 2012 en observar el cambio existente en fuentes de datos a partir del estudio de 248.000 textos durante 7 años. Utiliza diferentes técnicas de ventanado para detectar y cuantificar *concept drift*. En 2010, Zhu et al. (52) implementan un algoritmo de ventana doble (DWCDS, *Double Window Concept Drifting Data Streams*) para resolver el problema de *concept drift*, mejorando la robustez al ruido y la eficiencia de clasificación. Este estudio también se enfoca desde los principios de la minería de datos.

En relación al reconocimiento de actividades, algunos autores también han hecho uso de esta idea e incluso con el concepto de *concept drift*. Bifet y Gavaldá (53) desarrollan, en 2007, un algoritmo que permite adecuar el tamaño de la ventana según la tasa de cambio de los datos observados. De esta forma, ajusta la ventana a una longitud óptima estadísticamente. Gu et al. (54) introducen en 2009 la idea de que las actividades no tienen porqué ser solamente secuenciales, si no que pueden verse intercaladas dentro de otras o realizadas concurrentemente. De esta forma, utiliza como discriminador o diferenciador de actividades el concepto de *patrones emergentes*. Mediante la técnica MFI-CBSW (*Mining Frequent Itemsets over Circular Block Sliding Window*), un tipo de ventana deslizante en minería de datos, Memar et al. (55) plantean que dicho algoritmo actúa más rápido que otros propuestos recientemente para abordar *concept drift* en el año 2012. Patist (56) propone la obtención de un algoritmo en 2007 que es capaz de indicar un tamaño de ventana óptimo basado en el test de Kolmogorov-Smirnov para detectar cambio. En este sentido, hay que interpretar la detección de cambio como la lectura de *concept drift*.

Existen otros trabajos que han abordado el reconocimiento de actividades aplicando el concepto de ventanado. En 2008, Amft y Tröster (57) realizan un estudio de 3 actividades (tragar, masticar y mover brazos) con la orientación de observar la viabilidad para monitorizar enfermedades crónicas como la obesidad. Una de las decisiones tomadas consiste en adoptar un tamaño de ventana adaptado dinámicamente en consonancia con un algoritmo de búsqueda de similitud de características. La idea de reconocer las actividades conforme se producen nuevos eventos es planteada por Krishnan y Cook (58) en 2012. Las actividades se encuentran mejor caracterizadas aplicando tamaños de ventana diferentes a partir de 2 parámetros: tiempo de decaimiento e información mutua. En 2012, Okeyo et al. (59) plantean un sistema de reconocimiento de actividades continuo basado en conocimiento y ontologías usando ventanado temporal variable, sustentándose en principios de expansión y compresión.

En cuanto a la selección del tamaño de ventana, existe una gran diversidad entre los trabajos presentados según el estudio particular que pretenden. En el año 2001, Mantyjarvi et al. (60) indican que la frecuencia de muestreo es de 256 Hz, con una decimación de factor 2 y 256 muestras por ventana. La equivalencia supone tomar una ventana de 2 segundos de duración. Un único usuario utiliza 2 acelerómetros en la cadera (uno a cada lado) en un ambiente de oficina repitiendo una rutina de ejercicios predefinidos. Se busca el reconocimiento de la actividad 'andar' en conjunción con subida y bajada de escaleras así como la identificación de los eventos 'parar' y 'iniciar' dicha actividad. Se consiguen eficiencias de clasificación del 83-90 % con el clasificador *MLP* (Multi-Layer Perceptron). Posteriormente, Kern et al. (61) en 2003 emplean un tamaño de ventana de 1 segundo para reconocer hasta 6 actividades (andar, permanecer quieto, sentarse, escribir usando un teclado, escribir sobre una pizarra y darse la mano) y 12 sensores en diferentes configuraciones colocados en partes concretas (tobillo, cadera, codo, rodilla, muñeca, hombro, etc.). El muestreo de los datos de aceleración se lleva a cabo a 92 Hz y el algoritmo de clasificación testeado es *NB* (Naive Bayes) utilizando como características la media y la varianza de la señal. Se obtienen resultados dispares según la combinación y utilización de los diferentes sensores, alcanzando en algunos casos porcentajes superiores al 90 %. Ese mismo año, Krause et al. (62) justifican en su trabajo que 8 segundos por ventana supone una buena resolución para discernir entre varios patrones de movimiento con una granularidad fina. Utiliza técnicas de aprendizaje máquina, algoritmos de grafos y análisis estadísticos para tratar con parámetros fisiológicos como la temperatura o la frecuencia cardíaca junto a 2 acelerómetros biaxiales. El estudio permite reconocer cambios de actividad a partir de la consideración de las magnitudes mencionadas. Un dataset muy reconocido por la comunidad científica es el presentado por Bao e Intille (1) (2004), ya comentado con anterioridad. Se indica que el particionamiento de datos se realiza en ventanas de 6.7 segundos para reconocer hasta 20 actividades diferentes. Seguidamente, Ravi et al. (63) en 2005 destacan que actividades tan cotidianas como 'pasar la aspiradora', 'cepillarse los dientes' o 'subir escaleras' pueden ser muy bien reconocidas empleando tamaños de ventana grandes. En su estudio emplea una ventana de 5.12 segundos de duración (256 muestras a razón de 50 Hz) con un solapamiento del 50 % para reconocer un total de 8 actividades. A su vez, durante el año 2005, Huynh y Schiele (64) se plantean la posibilidad de que las actividades sean reconocidas con tamaños de ventana diferentes para maximizar la eficiencia. A partir de la consideración de magnitudes muy utilizadas en estos estudios (media, entropía, energía, varianza o coeficientes de la FFT) en combinación con hasta 19 parejas de características, concluye que los tamaños óptimos de reconocimiento para las 6 actividades estudiadas son 'saltar con un pie' (4 segundos), 'saltar a la comba' (2 segundos), 'trotar' (1 segundo), 'andar' (1 segundo), 'coger el autobús' (1 segundo) o 'permanecer

de pie' (0.5 segundos). Emplea una tasa de muestreo de 512 Hz y alcanza tasas de eficiencia que rondan el 80% empleando el clasificador *NCC* (Nearest Centroid Classifier), tomando como métricas determinantes la sensibilidad y la precisión para la obtención de conclusiones. Al año siguiente, Maurer et al. (3) consiguen clasificar actividades cada 0.5 segundos a partir del almacenamiento previo en un buffer de 4 segundos de duración. Con una tasa de muestreo de 50 Hz, su principal enfoque se destina al análisis de la influencia de la frecuencia de muestreo en la eficiencia de clasificación, por lo que aplica submuestreo en un rango de 1-30 Hz. Los algoritmos de clasificación utilizados son *DT* (Decision Tree), *KNN* (K-Nearest Neighbour) y la red de Bayes. A su vez, ese mismo año, Ward et al. (65) aplican voto mayoritario para realizar la clasificación con frames de 0.1 segundos tras la fragmentación e identificación de eventos en 1-2 segundos. Su trabajo considera sonidos recogidos por micrófonos junto a datos colectados por acelerómetros. Utiliza 2 algoritmos de clasificación fusionados para sus experimentaciones - *LDA* (Linear Discriminant Analysis) y *HMMs* (Hidden Markov Models) -, junto a 4 métodos diferentes (entrenamiento con un usuario dependiente, entrenamiento con usuario independiente, entrenamiento con usuarios adaptados y aislamiento). Pirttikangas et al. (66) consiguen en su estudio del año 2006 ver que con tamaños de ventana muy pequeños (0.1-1.5 segundos) se consiguen buenas tasas de reconocimiento. Utiliza una tasa de muestreo de 10 Hz, 17 sujetos y validación cruzada 4-fold sobre los algoritmos de clasificación *MLP* y *KNN*. Obtiene eficiencias de reconocimiento del 89.76% y 92.89%, respectivamente. En 2010, para el reconocimiento de la actividad a través de acelerómetros embebidos en un teléfono móvil, Sun et al. (67) testean diferentes tamaños de ventana (1, 2, 3, 4, 5 y 6 segundos) realizando un particionamiento de los datos en ventanas de 1 segundo. Su estudio se centra en el reconocimiento de 7 actividades de la vida diaria (andar, correr, permanecer de pie quieto, subir escaleras, bajar escaleras, conducir y montar en bicicleta) a partir de la monitorización de 7 sujetos y empleando un modelo de clasificación *SVM* (Support Vector Machine). El autor indica que el mejor reconocimiento se alcanza en 4 segundos sin utilizar datos inerciales, mientras que el tamaño de ventana óptimo se desplaza a 5 segundos si no se consideran dichos datos.

A la vista de las diferentes técnicas, planteamientos y enfoques realizados en investigaciones anteriores, queda de manifiesto que existe una influencia del tamaño de ventana sobre la eficiencia de reconocimiento de una actividad. Adicionalmente, los trabajos presentados reflejan que cada actividad requiere un tamaño de ventana diferente para alcanzar su mejor tasa de reconocimiento. Sin embargo, todos estos estudios no consiguen ofrecer una visión general sobre el reconocimiento de actividades ya que se tratan de investigaciones particulares centradas en alcanzar unos objetivos muy concretos. Algunos muestran el tamaño de ventana óptimo para un reducido número de actividades, mientras que otros experimentan con un número de sujetos

escaso (en algunos casos un único usuario, el propio investigador) o utilizan un tamaño de ventana fijo para el desarrollo de todo su estudio. Otros factores a tener en cuenta son la selección de la tasa de muestreo o la consideración de métricas que permitan evaluar con criterio los datos colectados. Ciertos trabajos aplican frecuencias de muestreo relativamente altas (256 o 512 Hz) en comparación con otros (10-50 Hz), un hecho que influye en una mayor o menor disposición de información para una adecuada clasificación. En lo que respecta a la aplicación de métricas, se convierte en una decisión crucial si se trata de tomar alguna decisión crítica en el devenir del estudio, que permita ofrecer una visión general sobre los resultados obtenidos y evitando su aplicación únicamente para observar resultados particulares.

1.5. Estructura de la memoria

Este documento está conformado por cinco capítulos. La memoria consta de: un capítulo de introducción, tres capítulos referentes a los estudios abordados para la elaboración de este proyecto, y un capítulo final en el que se incluyen las conclusiones e impresiones, así como una serie de reflexiones sobre posibles mejoras y el desarrollo futuro del tema tratado e investigado. En el primer capítulo, denominado “Capítulo I: Introducción” 1, se ha ofrecido una idea sobre la motivación existente para la realización de este estudio, con una visión general del panorama científico y los precedentes en este ámbito, marcando una serie de objetivos para este proyecto. Con el comienzo del segundo capítulo, titulado “Capítulo II: Reconocimiento de la actividad” 2, se ofrece una perspectiva más cercana a la temática planteada en lo que respecta a la familiarización con los términos más relevantes del reconocimiento de la actividad y las principales consideraciones para ser abordado. El tercer capítulo, al que se le ha llamado “Capítulo III: Estudio del efecto del tamaño de ventana en el reconocimiento de la actividad” 3, supone una toma de contacto con la herramienta para el tratamiento de los datos biomecánicos seleccionados, así como el análisis y estudio de la principal motivación del proyecto: comprobar la influencia del tamaño de ventana en el reconocimiento de la actividad. El siguiente capítulo está dedicado a la innovación y aplicación de los conocimientos adquiridos sobre el estudio realizado, buscando una mejora de las prestaciones del sistema de reconocimiento de la actividad a partir de la aplicación de un algoritmo de fusión de ventanas. Esta tarea se recoge en el capítulo como “Capítulo IV: Fusión multiventana” 4. El capítulo final recoge varios apartados dedicados a indicar las conclusiones extraídas tras la finalización del proyecto, además de alguna opinión personal sobre el futuro en este ámbito, entre otras valoraciones. Se ha denominado “Conclusiones generales y trabajo futuro” 5.

2

Capítulo II: Reconocimiento de la actividad

2.1. Introducción

El reconocimiento de la actividad se ha convertido en un reto continuo del ámbito de la biomecánica. El surgimiento de esta disciplina ha traído consigo un interés del panorama científico que va más allá de la resolución del problema general que plantea. Muchos investigadores dedican su esfuerzo para identificar, considerar y modelar diferentes parámetros de influencia y determinación en las distintas fases del sistema de reconocimiento de la actividad.

En las secciones posteriores se ofrece una visión sobre diferentes aspectos de consideración a la hora de adentrarse en el tratamiento de datos biomecánicos, así como una descripción tanto sobre la cadena de reconocimiento de la actividad como sobre la disposición de un dataset que contenga este tipo de datos para su procesamiento y estudio.

2.2. Modalidades de sensado: sensores para el reconocimiento de actividades. Limitaciones y dificultades

La complejidad del reconocimiento de actividades de la vida diaria reside en la diversidad de ejecuciones posibles y situaciones contextuales que pueden hacer referencia a la misma actividad. Dado que muchas de ellas están relacionadas al movimiento del cuerpo (o bien, a la ausencia de movimiento), la utilización de sensores de movimiento se han convertido en una de las alternativas más recurrentes en la literatura de este ámbito.

Para el análisis del movimiento, existen 2 tipos de sensores empleados dependiendo del lugar en el que se encuentren colocados (sensores portátiles o en la ropa) o en los alrededores (sensores ambientales) del sujeto. En principio, el uso de sensores ambientales como cámaras

o micrófonos se encuentra restringido a escenarios particulares donde su despliegue es factible. Aún siendo posible, existen algunas restricciones adicionales que pueden dificultar su uso como la privacidad del sujeto, oclusiones, ruido ambiente, etc. Por consiguiente, ha aparecido una tendencia que va en incremento y orientada al uso de sensores en el cuerpo que carecen de este tipo de limitaciones.

Del abanico de sensores que pueden ser sujetos al cuerpo, la mayoría emplean exhaustivamente los sensores inerciales debido a que proporcionan buenos resultados para diferentes configuraciones y actividades. Sin embargo, uno de los principales inconvenientes reside en lo importuno o molesto que supone portarlos. Varios dispositivos de un tamaño considerable atados a diferentes partes del cuerpo del sujeto resulta una manera incómoda que puede ser descartada como una solución realista para adoptarla diariamente. Afortunadamente la miniaturización de los sensores, la reducción del consumo de batería y la producción a bajo coste nos permite vislumbrar una nueva generación de sensores diminutos que lleguen a ser integrados en *cosas portables* como la ropa o accesorios del sujeto. Estas ideas fueron presentadas por Amft y Lukowicz (68) en 2009.

2.3. La influencia de los factores exógenos

Hay algunas características bien conocidas de las principales actividades de la rutina diaria que teóricamente pueden ser adecuadamente discriminadas. Por ejemplo, dependiendo de la intensidad de los movimientos, se debería ser capaz de distinguir entre algunos ejercicios como andar o correr, aún compartiendo un estilo de ejecución común. La orientación del cuerpo puede aportar información útil sobre la postura cuando se realiza una actividad de baja intensidad o cuasi-estática (como por ejemplo, diferenciar entre estar acostado o permanecer de pie parado). Sin embargo, existen factores exógenos que dificultan la tarea del reconocimiento. La edad, el peso, la altura u otras características relacionadas con el sujeto, así como los factores ambientales y contextuales (por ejemplo, cuando el sujeto lleva objetos, un suelo inestable, etc.) pueden determinar adquisiciones de datos con diferencias notables y que podrían referirse a una actividad similar. Un ejemplo ilustrativo: uno no puede esperar registrar el mismo tipo de datos cuando un adulto está haciendo bicicleta que cuando lo hace una persona mayor. De manera similar, el paso de una persona puede diferir cuando camina sobre tierra, hierba o incluso una superficie helada.

2.4. La cadena de reconocimiento de actividades

Para hacer frente a la mayoría de estos problemas, se propone una metodología general. Tradicionalmente reconocida como *activity-recognition chain* (de manera abreviada ARC, y que significa cadena de reconocimiento de la actividad), se presenta con un esquema general como el mostrado en la figura 2.1.

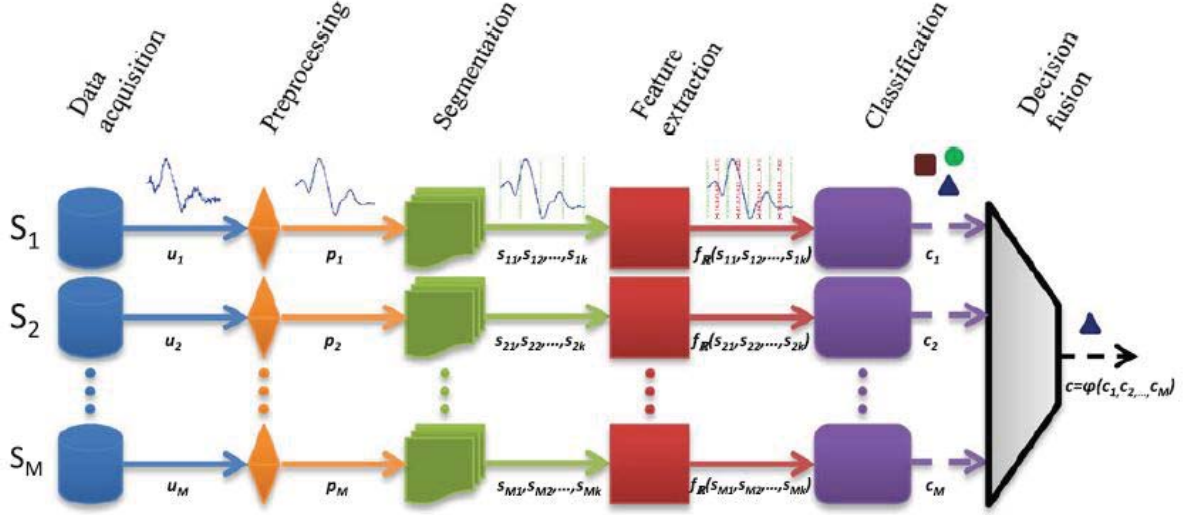


Figura 2.1: Cadena múltiple de reconocimiento de la actividad (M-ARC). M sensores proporcionan señales en bruto (μ_j) que son subsecuencialmente procesadas (p_j). Las señales son k -particionadas (s_{jk}) y un conjunto de características (habitualmente definidas como f) son extraídas de ellas, posiblemente de manera diferente para cada cadena. El vector de características es usado como entrada al clasificador de entidades. Cada clasificador produce una clase en un problema con N -clases que puede ser combinado a través de un método de decisión basado en fusión. Los índices son respectivamente definidos por $j = 1, \dots, M$; $k = 1, \dots, K$; $i = 1, \dots, N$.

De izquierda a derecha, un total de M fuentes (sensores) proporcionan señales en bruto sin procesar (μ_j) que representan la magnitud medida (por ejemplo, aceleración). Las señales son usualmente preprocesadas (p_j) típicamente a través de un proceso de filtrado para evitar la interferencia de ruido y diversos artefactos naturales. Con el fin de capturar la dinámica de las señales, éstas son particionadas en segmentos de una cierta longitud (s_{jk}). Se han diseñado diferentes técnicas para este propósito, principalmente basadas en *windowing* (ventanado) o *event-activity-based segmentation* (segmentación de eventos basada en actividades). Subsecuencialmente, se lleva a cabo un proceso de extracción de características para proporcionar una representación útil de las señales durante la etapa de reconocimiento de patrones. Se emplean para esta tarea un amplio rango de heurísticas, tanto en el dominio temporal como espectral, así como otras funciones matemáticas y estadísticas muy sofisticadas. El vector de características

se entrega a la entrada del clasificador ($f(s_{jk})$), que termina el proceso asignando la actividad o clase reconocida (c_j) a uno de los N considerados para el problema particular. La última etapa opcional corresponde a un modelo de fusión que puede combinar las decisiones de cada *ARC* individual para mejorar la fiabilidad del sistema de reconocimiento. Se puede consultar una revisión extensa de las principales etapas del sistema en el trabajo de Preece et al. (69) (2009).

Cada etapa del *ARC* puede verse afectada por diferentes tipos de dificultades. A nivel de señal, y más allá de las características especiales de cada sensor, los datos pueden contener ruido y artefactos. Dependiendo del tipo de señales y de las actividades consideradas, las técnicas de preprocesamiento clásicas pueden ser más o menos apropiadas. Por ejemplo, algunos procesos de filtrado pueden implicar pérdida de información. Dado que la eliminación eventual de información puede incluso ser información crucial para determinar algunas de las actividades consideradas, la selección de una técnica de preprocesamiento adecuada debería ser cuidadosamente estudiada. Con respecto al proceso de segmentación, no está claramente definido que tamaño de ventana de datos debe considerarse. En general, depende de la complejidad, la duración y la granularidad de las actividades entre otras consideraciones. El proceso de extracción de características constituye usualmente el cuello de botella computacional del sistema *ARC*. Puesto que el objetivo final podría ser definir un sistema de reconocimiento de la actividad en tiempo real, es importante buscar características que no sean demasiado costosas en términos de recursos o reducir el número de características requeridas al mínimo posible (optimización de la eficiencia). La reducción puede en general ayudar a definir clasificadores más simples. Otro problema es la búsqueda de un vector de características óptimo (optimización del rendimiento). Idealmente, todas las combinaciones posibles sobre las características extraídas deberían ser evaluadas, lo que implica un problema de búsqueda con una complejidad exponencial (O^n). Un filtro que aplique métodos de selección de características nos permite reducir la búsqueda, pero en general, no aseguran la selección del mejor subconjunto de características posible. Por un lado, los *wrapper methods* (métodos contenedor) evalúan las capacidades de las características cuando se emplean en los clasificadores, así proporcionan una representación más fiable de la característica, pero desgraciadamente requieren una enorme cantidad de recursos computacionales y tiempo. Por otra parte, los métodos de aprendizaje máquina se ven específicamente más afectados por problemas relacionados a su uso práctico, normalmente debido a datos no balanceados o límites de convergencia. La generalidad empleada para la definición de un sistema *ARC* permite que todos estos problemas puedan ser afrontados de una forma diferente dependiendo de las modalidades de sensores considerados. Sin embargo, no todas las soluciones están en línea

con la definición de un modelo utilizable con independencia de una configuración y un contexto particular.

Incluso con una cadena de reconocimiento de la actividad optimizada, resulta evidente que algunos sensores puedan estar más especializados en el reconocimiento de ciertas actividades. Consecuentemente, para incrementar las posibilidades del reconocimiento, se sugiere la utilización y combinación de sensores. Esta combinación o fusión puede ser implementada en cada nivel del *ARC*, como sugieren Sharma et al. (70) (1998). Por ejemplo, la fusión de sensores puede implementarse en el nivel de extracción de características, definiendo así un vector de características simple compuesto por las características independientes extraídas por cada sensor. Esta proposición se puede observar en el trabajo de Mantyjarvi et al. (60) (2001). El problema que surge es que habría que lidiar con un problema de espacio de características de alta dimensionalidad, cuestión que dificultaría la selección de características y el proceso de clasificación tal y como se ha descrito anteriormente. Esta aproximación no es escalable para un aprendizaje online, puesto que la inclusión de un nuevo sensor requiere redefinir el vector de características y reentrenar el sistema completo.

Más interesante resulta definir un esquema de fusión actuando a nivel de clasificación (como se presenta en la figura 2.1). La idea es básicamente combinar las decisiones desarrolladas por cada *ARC* individual en una decisión única reforzada. En general, el sistema basado en fusión será más preciso y robusto que los *ARC* individuales. Además, a diferencia de la fusión a nivel de características, la introducción o eliminación de sensores no supondría tener que reentrenar los sistemas iniciales, pero sí actualizar la estructura de parámetros de decisión. El principal problema la fusión de decisiones es definir un modelo eficiente que considere la robustez y la escalabilidad. Esto significa que el modelo debería ser suficientemente eficiente con independencia de la topología o el número de sensores considerados, como exponen Baños et al. (71) (2013).

2.5. Configuración, métodos y datos empleados en el estudio

Para la realización de cualquier tipo de estudio relacionado con esta disciplina, tener acceso a un dataset que considere los aspectos mencionados en epígrafes anteriores y que reúna diferentes requisitos en lo que respecta a los datos almacenados, se trata de un punto de partida muy importante para su abordaje. Este proyecto ha tomado como base fundamental para su desarrollo el dataset presentado por Baños et al. (8) (2012). En la siguiente sección se exponen todos los detalles relacionados con esta base de datos y los motivos principales para su elección.

2.5.1. Dataset para el reconocimiento de actividades

La base de datos seleccionada para la elaboración de este proyecto es muy reciente, ya que fue presentada justo hace un año, en Septiembre de 2012. Los autores han elaborado un dataset¹ que está conformado por diferentes datos biomecánicos recogidos por sensores capacitados para medir señales de aceleración, medidas de giróscopo y magnetómetro en las 3 dimensiones espaciales junto a una estimación de la orientación en formato cuaternión en 4 dimensiones. Un total de 33 actividades han sido monitorizados con 9 sensores inerciales colocados en la espalda y las extremidades de hasta 17 sujetos 2.2. Cada sensor recoge 13 valores de cada medida, confeccionando un conjunto de 117 señales, grabadas a una frecuencia de muestreo de 50 Hz. Este trabajo presenta 3 escenarios que dependen del desplazamiento de los sensores: ideal, calibrado y mutuo.

Para este proyecto, dado el volumen de datos almacenados y la especificación del estudio que se pretende, se ha optado únicamente por utilizar las señales de aceleración. Estudios previos han demostrado que empleando únicamente medidas inerciales se consiguen buenos resultados. De igual forma, y acotando un poco más este estudio, se han seleccionado los datos recogidos para el primer escenario (datos ideales), en el que se entiende que los sensores son colocados por un instructor en un ambiente supervisado.

Los sujetos realizan una serie de ejercicios de gimnasio que pueden ser clasificados según la parte del cuerpo implicada específicamente en la actividad:

★ Todo el cuerpo

- a) **Movimientos de traslación** (1, 2 y 3)
- b) **Saltos** (4, 5, 6, 7 y 8)
- c) **Fitness** (31, 32 y 33)

★ Parte específica

- a) **Tronco** (9, 10, 11, 12, 13, 14, 15, 16, 17 y 18)
- b) **Extremidades superiores** (19, 20, 21, 22, 23, 24 y 25)
- c) **Extremidades inferiores** (26, 27, 28, 29 y 30)

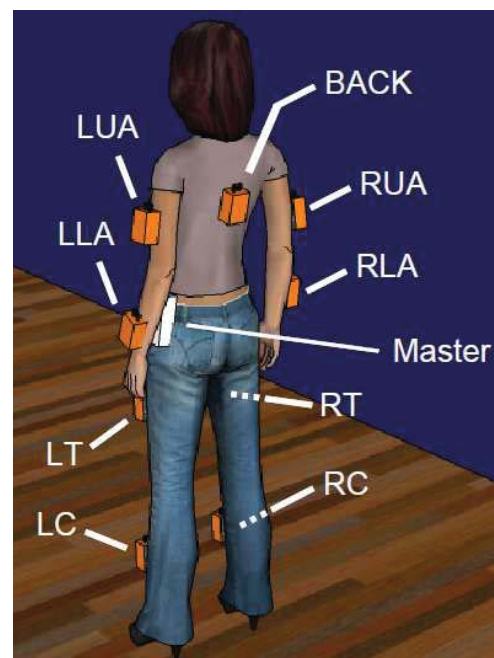


Figura 2.2: Ubicación de los 9 sensores en diferentes posiciones del cuerpo

¹El dataset está disponible en <http://www.actlab.ele.tue.nl/datasets>

A continuación, se presenta el conjunto de actividades tratadas especificando el grupo al que pertenecen entre los expuestos en la clasificación anterior:

1. **Andar** (1 min)
2. **Trotar** (1 min)
3. **Correr** (1 min)
4. **Saltar hacia arriba** (20x)
5. **Saltar adelante y atrás** (20x)
6. **Saltar lateralmente** (20x)
7. **Saltar abriendo y cerrando piernas y brazos** (20x)
8. **Saltar la cuerda** (20x)
9. **Girar el tronco con los brazos extendidos** (20x)
10. **Girar el tronco con los codos doblados** (20x)
11. **Doblar la cintura hacia delante** (20x)
12. **Rotación de cintura** (20x)
13. **Alcanzar un pie con la mano contraria y la cintura doblada** (20x)
14. **Alcanzar los talones flexionando las rodillas** (20x)
15. **Doblarse lateralmente** (10x izquierda + 10x derecha)
16. **Doblarse lateralmente con un brazo arriba** (10x izquierda + 10x derecha)
17. **Doblar la cintura hacia delante repetidamente** (20x)
18. **Giro de extremidades superiores e inferiores en oposición** (20x)
19. **Elevación lateral de brazos** (20x)
20. **Elevación frontal de brazos** (20x)
21. **Palmadas frontales** (20x)
22. **Cruzar brazos frontalmente** (20x)
23. **Rotación intensa de hombros** (20x)
24. **Rotación moderada de hombros** (20x)
25. **Rotación interna de brazos** (20x)
26. **Rodillas al pecho (alternativamente)** (20x)
27. **Talones a la espalda (alternativamente)** (20x)
28. **Doblar las rodillas agachándose** (20x)
29. **Doblar las rodillas hacia delante (alternativamente)** (20x)
30. **Rotación de rodillas** (20x)
31. **Remar** (1 min)
32. **Hacer bicicleta elíptica** (1 min)
33. **Hacer bicicleta estática** (1 min)

Los datos monitorizados para cada participante y escenario específico se recogen en ficheros que constan de 120 columnas que almacenan la siguiente información: *marcas temporales* (columnas 1 y 2), *datos biomecánicos* (columnas 3-119) y *etiqueta de identificación de la actividad* (columna 120). Las marcas temporales indican el instante de muestreo en segundos y microsegundos, así como la etiqueta de identificación puede indicar cualquier actividad (entre 1 y 33), incluyendo 0 para mostrar que no hay actividad registrada en ese momento. De mayor interés es la identificación de los datos de aceleración. Cada sensor recoge 13 medidas en el siguiente orden: acelerómetro (columnas 1-3 para ejes X, Y, Z), giróscopo (columnas 4-6 para ejes X, Y, Z), magnetómetro (columnas 7-9 para ejes X, Y, Z) y cuaternión (columnas 10-13).

2.5.2. Modelos y configuración experimental

Una vez el dataset ha sido presentado, hay que seleccionar una serie de parámetros que determinarán unívocamente los resultados finales. En el epígrafe 2.4 se indica con detalle cada una de las decisiones necesarias que hay que adoptar en las diferentes fases de un *ARC*. A continuación se presentan brevemente todos los detalles específicos para la configuración de nuestro sistema.

2.5.2.1. Adquisición de datos

Se han empleado un total de $M = 9$ sensores para la monitorización de los datos biomecánicos. Dado que únicamente se desean tomar los datos recogidos por los acelerómetros (es decir, la aceleración es nuestra magnitud μ_j - con $j = 1, \dots, 9$ -), se extraen las columnas 3, 4 y 5 de cada fichero, así como las columnas con un desfase de 13 columnas para el siguiente sensor. En total, $9 \times 3 = 27$ columnas de datos de aceleración se emplean en el estudio.

2.5.2.2. Preprocesamiento

Se han tomado los datos inerciales bajo supervisión en bruto, sin ningún tipo de preprocesamiento previo a la etapa de segmentación.

2.5.2.3. Segmentación

Esta fase requiere la selección de un tamaño de ventana para la segmentación de los datos en K -particiones denominadas *ventanas*. En la sección 1.2 ya se introdujo el interés sobre el estudio de esta etapa: comprobar el efecto del tamaño de ventana sobre la eficiencia del reconocimiento de la actividad. Bao e Intille (1) demostraron que un tamaño de ventana en torno a los 6 segundos es suficiente para reconocer cualquier actividad con ciertas garantías. Este valor se ha tomado como referencia a lo largo del proyecto para la comparación de resultados. Además, indicaba

que se trataba de una duración idónea en un dataset amplio. En su estudio, se trataban de 20 actividades y en este caso son un total de 33 actividades. Para comprobar la influencia existente con la variación del ventanado, se ha optado por realizar un barrido entre 0 y 7 segundos, ya que en líneas generales, el principal interés recae en comprobar si es posible reducir el tamaño de ventana en cualquier caso y la investigación de Bao e Intille (1) en 2004 invita a tomar su demostración como una acotación para este estudio. Una reducción del tamaño de ventana supondría una ventaja importante como aumentar la velocidad de respuesta del sistema de reconocimiento o, lo que es lo mismo, disponer de la decisión del reconocimiento de la actividad en menor tiempo. Tras la selección del rango de interés, se ha elegido un paso de ventana de 0.25 segundos. Esta elección ya ha sido utilizada anteriormente por Huynh y Schiele (64) (2005). De esta forma, el particionamiento supone la aplicación de 28 tamaños de ventana diferentes.

2.5.2.4. Extracción de características

En la sección 2.4 se indica que es recomendable la selección de características que no supongan demasiado coste computacional y la reducción del repertorio en la medida de lo posible. Las 5 características seleccionadas han sido utilizadas ampliamente en diferentes estudios sobre el reconocimiento de la actividad: media, desviación estándar, máximo, mínimo y la tasa de cruce medio de la señal inercial analizada. Así lo respaldan los estudios de Ravi et al. (63) (2005), Lester et al. (72) (2005) y (73) (2006), Maurer et al. (3) (2006), Wu et al. (74) (2009), Yang (75) (2009) y Sun et al. (67) (2010). En total, se han elaborado 3 conjuntos de características para las fases de experimentación como combinación de las 5 características mencionadas. El primer conjunto selecciona únicamente la media de la señal, el segundo añade la desviación estándar de la señal y el tercero agrega adicionalmente las 3 características restantes. A lo largo del proyecto serán nombrados como *FS1*, *FS2* y *FS3* (que son las siglas de *feature set* 1, 2 ó 3).

2.5.2.5. Clasificación

4 algoritmos de clasificación han sido seleccionados para este proyecto teniendo en cuenta las indicaciones presentadas en 2.4:

- ★ **DT** (Decision Tree): Se trata de un algoritmo cuya estructura es un *árbol de decisión*. Suelen emplearse en procesos de búsqueda binaria, sistemas expertos o árboles de juego. Yang (75) (2009) lo ha testeado en su trabajo. Su implementación puede consultarse en el trabajo de Duda (76) (2000).
- ★ **NB** (Naive Bayes): El clasificador bayesiano ingenuo se fundamenta en la utilización del teorema de Bayes con suposiciones de independencia fuertes, aplicando una probabilística simple

y haciendo uso del aprendizaje. La aplicación de ciertas hipótesis simplificadoras adicionales, que se suelen resumir en la hipótesis de independencia entre las variables predictoras, le han supuesto la adopción del apelativo *ingenuo*. Estudios como los presentados por Bao e Intille (1) (2004), Ravi et al. (63) (2005), Maurer et al. (3) (2006) o Yang (75) (2009) han hecho uso de dicho algoritmo. Se ha adoptado la aproximación presentada por Theodoridis (77) del año 2008.

- ★ **NCC** (Nearest Centroid Classifier): Su idea consiste en localizar *el centroide más cercano*. En el panorama científico también es conocido con esas siglas como *Nearest Class Center*. Requiere un proceso de entrenamiento previo y aplica una función de predicción que evalúa la distancia mínima entre el centroide y la observación asignada a una clase. Ha sido empleado en importantes estudios sobre el reconocimiento de la actividad como los estudios de Roggen et al. (12) y Forster et al. (45), ambos en 2009.
- ★ **KNN** (K-Nearest Neighbors): Conocido como el algoritmo de los K vecinos más cercanos, basa su estrategia de clasificación en comparar la muestra con las K clases más próximas. Se toma la clase que obtenga el mayor número de votos. Normalmente, se selecciona un K 'par' con el fin de evitar hipotéticos empates en clasificación. Este algoritmo también se ha aplicado en diversos trabajos como los presentados por Yang (75) (2009) y Baños et al. (8) (2012). Para este proyecto, se ha tomado $\mathbf{K} = \mathbf{3}$ ya que ofrece buenos resultados además de evitar sobrecarga computacional. El algoritmo se encuentra explicado por Cover y Hart (78) (1967).

3

Capítulo III: Estudio del efecto del tamaño de ventana en el reconocimiento de la actividad

3.1. Introducción

El capítulo 1 ha permitido contextualizar y presentar la metodología habitual en el reconocimiento de actividades, y que se pretende aplicar en este proyecto. A partir del dataset elaborado por Baños et al. (71) este mismo año y la selección de parámetros específicos en *ARC*, las condiciones de partida del presente estudio han quedado definidas.

En las siguientes secciones se da un paso adelante en el estudio del efecto del tamaño de ventana sobre el reconocimiento de la actividad al tratarse de la principal motivación de este trabajo

En primer lugar, se realiza una explicación del estudio del ventanado que se desea llevar a cabo.

Posteriormente, durante la evaluación del estudio, se presenta la herramienta empleada para el tratamiento de los datos, así como la obtención de una serie de parámetros que permiten extraer diversas conclusiones con la presentación de tablas de resultados y análisis gráficos.

Seguidamente, se ofrece la interpretación y un análisis crítico sobre los resultados obtenidos en la sección anterior.

El capítulo termina con la presentación de las conclusiones derivadas de dicho estudio y que son la base fundamental para la continuación del proyecto.

3.2. Explicación del estudio

Los estudios que se pretenden abordar durante este capítulo son los siguientes:

- ★ **Estudio de la eficiencia global de clasificación en función del tamaño de ventana:** Se estudia la influencia del particionamiento de los datos en la etapa de segmentación sobre el reconocimiento del clasificador a partir de la consideración de un parámetro que permite medir el rendimiento del clasificador como es el *accuracy*.
- ★ **Minimización del tamaño de ventana sobre cada actividad:** Se analiza el efecto del tamaño de ventana sobre cada actividad en particular a partir de una métrica concreta que se presenta en la sección 3.3.
- ★ **Penalización del reconocimiento de la actividad a costa de optimizar el tamaño de ventana:** Se aplican una serie de reducciones de la magnitud métrica seleccionada con el fin de minimizar el tamaño de ventana para el reconocimiento de cada actividad.
- ★ **Obtención del tamaño de ventana que minimiza la penalización en el reconocimiento global de actividades:** Se realiza un estudio específico sobre la influencia de la métrica empleada en el reconocimiento particular de cada actividad y se obtiene un compromiso de diseño.

Tras presentar algunas consideraciones necesarias para la evaluación del estudio en el siguiente epígrafe, estos estudios se desarrollan en la sección 3.4.

3.3. Evaluación del estudio

Para el desarrollo y análisis de las distintas fases de experimentación del proyecto se ha utilizado la herramienta MATLAB. Este programa permite tratar los datos inerciales de manera sencilla, así como ofrecer múltiples posibilidades para su estructuración, identificación y localización. En el epígrafe 2.5.1 se ofrece toda la información de interés sobre el dataset abordado, y en concreto, sobre la extracción de los datos inerciales y la estructura en la que se encuentran. De la misma forma, se indican los parámetros particulares de cada fase del *ARC*.

Las diferentes combinaciones (tamaño de ventana, conjunto de características y paradigma de clasificación) han sido testeadas tomando las 33 actividades del dataset y los 17 sujetos participantes. Para asegurar la robustez estadística de las simulaciones, se han considerado 100 iteraciones con una validación cruzada de 10-fold por iteración.

La aplicación de diferentes tamaños de ventana sobre los datos disponibles permite obtener una serie de parámetros que serán objeto de un posterior análisis:

- ★ **Specificity** (especificidad): Indica la capacidad del predictor para identificar resultados negativos con respecto a la prueba sometida. Su valor está comprendido entre 0 y 1.
- ★ **Sensitivity** (sensibilidad): Indica la capacidad del predictor para identificar resultados positivos con respecto a la prueba sometida. También es conocido como *recall*. Su valor está comprendido entre 0 y 1.
- ★ **Precision** (precisión): Indica la eficiencia con la que una clase específica es predicha. También es conocido como *positive predictive value*. Su valor está comprendido entre 0 y 1.
- ★ **Accuracy** (eficiencia): Determina la exactitud global de la prueba realizada. También es conocido como *correct rate*. Se puede presentar entre 0 y 1, o bien en porcentajes (%)
- ★ **Confusion Matrix** (matriz de confusión): También llamada 'matriz de error' o 'matriz de contingencia', se trata de una herramienta muy usada para evaluar la exactitud de una clasificación y parámetros como los mencionados anteriormente se pueden obtener a partir de ella. En nuestro caso, será una matriz cuadrada de 33x33 dado que disponemos de 33 actividades. Las filas indican las etiquetas de las clases conocidas (datos de referencia), mientras que las columnas responden a las etiquetas de las clases predichas (datos categorizados por el clasificador).

Una vez se han obtenido los resultados de simulación pertenecientes a las 336 combinaciones posibles (a partir de los 28 tamaños de ventana, 4 clasificadores y 3 conjuntos de características seleccionados), se han considerado diferentes estudios a partir de dichos parámetros.

Uno de los principales objetivos de este estudio es comprobar el efecto del tamaño de ventana para las distintas metodologías planteadas sobre el rendimiento del clasificador. Dichos resultados se presentan en una tabla 3.1 y una representación gráfica que permite comprobar la influencia del tamaño de ventana a partir de la observación de las tendencias para cada metodología 3.1. Otros detalles particulares, como comparaciones de resultados frente al tamaño de ventana de 6 segundos, se recogen en las tablas 3.2 y 3.3.

En este punto, se ha obtenido una optimización de la eficiencia de clasificación con la variación del tamaño de ventana, consiguiendo incluso una reducción de dicho parámetro. Sin embargo, la optimización contempla una mejora a nivel global del reconocimiento de las actividades procesadas, sin considerar los reconocimientos particulares de cada actividad. Un parámetro como la eficiencia no permite discernir ni cuantificar la evolución en el reconocimiento de una actividad a partir de la variación del tamaño de ventana.

Para evaluar con más precisión el impacto o efecto del tamaño de ventana sobre cada actividad del dataset, se requiere la identificación de una métrica que considere tanto los fallos

cometidos al tratar de reconocerla como los fallos cometidos al tratar de predecirla. Una estructura como la matriz de confusión de actividades recoge exactamente ambos requisitos. Sin embargo, contrastar cada combinación acierto-fallo en el reconocimiento entre cada par de actividades aumenta la complejidad del análisis y resulta inabordable para el desarrollo de este proyecto. Dos métricas concretas permiten resumir los fallos y los aciertos en el reconocimiento de una actividad: la *precisión* y la *sensibilidad*.

La precisión es una magnitud que se penaliza cuando existen falsos positivos. Cuanto más cercano sea su valor a 1, mayor será la probabilidad de que cualquier otra actividad no sea confundida con la realización de la actividad predicha. Por ejemplo, si la precisión es igual a 1 para la actividad *andar*, quiere decir que cuando el clasificador detecte una actividad que no sea *andar*, nunca se confundirá clasificándola como *andar*. Podrá confundirse con otra actividad, pero en ningún caso diría que se ha detectado la actividad *andar*. Conforme el valor de la precisión disminuye y más cerca está de 0, mayor probabilidad existirá de que el clasificador la asigne como clase predicha. En este sentido, se puede decir que la precisión es interpretada como una medida de la eficiencia con la que una clase específica es reconocida por el clasificador.

Por su parte, la sensibilidad es una magnitud que se penaliza cuando existen falsos negativos. Cuanto más cercano sea su valor a 1, mayor será la probabilidad de que la actividad que se pretende reconocer sea reconocida y no se reconozca como otra. Por ejemplo, si la sensibilidad es igual a 1 para la actividad *correr*, quiere decir que cuando el clasificador reciba una característica que define a la actividad *correr*, no cometerá fallo en el reconocimiento y siempre será predicha correctamente. Conforme el valor de la sensibilidad disminuye y más cerca está de 0, mayor probabilidad existirá de que el clasificador sea incapaz de reconocerla y la confunda con cualquier otra. Se puede decir que la sensibilidad es interpretada como una medida de la capacidad del clasificador para seleccionar instancias de una determinada actividad de un conjunto de actividades.

Para considerar la influencia de ambos parámetros, existe una métrica que se denomina F_1 - *score* (que procede de la medida F_β con $\beta = 1$), interpretada como la media ponderada de los 2 parámetros implicados: precisión (P) y sensibilidad (o *recall* (R)):

$$F_\beta = (\beta^2 + 1) \frac{PR}{\beta^2 P + R} \xrightarrow{\beta=1} F_1 = \frac{2PR}{P + R} \quad (3.1)$$

Esta métrica también ha sido empleada en trabajos que tratan el reconocimiento de la actividad (Minnen et al. (79) (2006), Tapia et al. (80) (2007), Fan et al. (81) (2009), Sun et al. (67) (2010) y Bouma et al. (82) (2012)). También ha sido empleada en otras disciplinas científicas que la consideran un buen parámetro de medida para algoritmos de aprendizaje

(Sokolova et al. (83) (2006)) o para clasificación de correos en su variante con $\beta = 3$ (Dredze et al. (84) (2006)).

Una vez que se dispone de una única métrica, se puede calcular $F_1 - score$ para las 336 combinaciones planteadas anteriormente (28 ventanas x 4 clasificadores x 3 conjuntos de características). Sin embargo, más allá de presentar 12 tablas correspondientes a cada metodología con los 28 tamaños de ventana y 33 actividades, se utilizan dichos datos para realizar una serie de análisis relacionados con el reconocimiento específico de cada actividad incluida en este estudio. Se han considerado 2 estudios particulares para comprobar la influencia del tamaño de ventana sobre cada actividad:

- ★ **Minimización del tamaño de ventana:** Permite observar qué $F_1 - score$ se alcanza según el tamaño de ventana seleccionado. El criterio impuesto supone la selección del tamaño de ventana mínimo que garantiza un $F_1 - score$ determinado para el reconocimiento de la actividad.
- ★ **Penalización de $F_1 - score$:** Permite aplicar penalizaciones específicas al $F_{1,máx} - score$ alcanzado con el fin de reducir el tamaño de ventana para el reconocimiento de cada actividad. Este análisis presenta 2 variantes gráficas: *análisis por actividad* (agrupando las 12 metodologías) y *análisis por metodología* (agrupando las 33 actividades).

Para estos análisis, se han obtenido 2 tablas resumen que recogen los $F_1 - scores$ máximos de cada metodología y actividad 3.4 y los tamaños de ventana asociados a los $F_{1,máx} - score$ anteriores 3.5. Estas tablas determinan unívocamente cuál es el tamaño de ventana máximo que se debe aplicar para el reconocimiento específico de una actividad con una metodología seleccionada para conseguir su mayor eficiencia de reconocimiento posible. Esta observación permite descartar como tamaños de ventana válidos (o recomendables para su aplicación) aquellos que superen el tamaño de ventana obtenido en la elaboración de la tabla 3.5, simplemente por el mero hecho de que aumentar el tamaño de ventana no permite reconocer la actividad con una eficiencia mayor.

Por último, se presenta un último estudio que permite responder a la siguiente pregunta: *¿qué tamaño de ventana se podría emplear en la etapa de segmentación para penalizar lo mínimo posible el reconocimiento global de todas las actividades?*

En la siguiente sección se presentan toda la información anticipada en este epígrafe, con diferentes tablas y diversas representaciones gráficas. Además, se exponen las interpretaciones derivadas de cada elemento presentado, indicando todos los detalles extraídos de su análisis y discutiendo la validez de los resultados presentados.

3.4. Resultados y discusión

En la sección anterior se han presentado las diferentes representaciones consideradas para abordar el estudio del efecto del tamaño de ventana sobre el reconocimiento de la actividad.

La primera inquietud reside en comprobar qué eficiencia de clasificación se obtiene tras la aplicación de un ventanado variable. La tabla 3.1 muestra las combinaciones existentes entre algoritmos de clasificación y conjuntos de características para los tamaños de ventana indicados, así como la figura 3.1 refleja las tendencias de cada metodología:

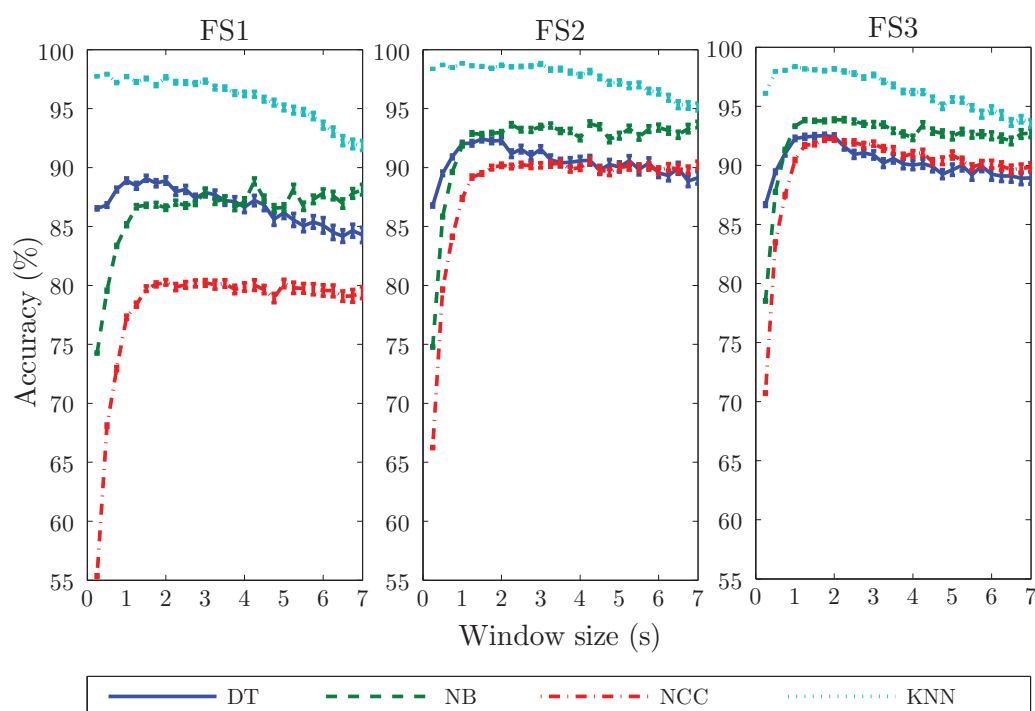


Figura 3.1: Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función del tamaño de ventana (en segundos)

La lectura inmediata que se puede extraer es que queda de manifiesto la influencia del tamaño de ventana sobre la eficiencia de reconocimiento en clasificación. La metodología que presenta los mejores resultados para cualquier tamaño de ventana es el algoritmo de clasificación KNN a través de la extracción de las características media y desviación estándar de la señal (de manera abreviada, *KNN2*). Dicho resultado se ha resaltado en cursiva. En cuanto al tamaño de ventana que permite alcanzar la mayor eficiencia de clasificación para cada metodología, se ha destacado en negrita.

	DT			NB			NCC			KNN		
0.25	86.54	86.77	86.68	74.26	74.79	78.53	55.35	66.25	70.73	97.74	<i>98.37</i>	96.09
0.5	86.81	89.51	89.45	79.56	85.83	87.75	68.10	79.65	83.48	97.92	<i>98.70</i>	97.95
0.75	88.14	90.91	90.99	83.35	89.63	91.31	72.94	84.12	87.48	97.21	<i>98.48</i>	98.04
1	88.88	92.03	92.27	85.14	91.89	93.32	77.30	87.38	90.44	97.71	98.84	98.35
1.25	88.45	92.06	92.41	86.68	92.89	93.83	78.36	89.19	91.69	97.25	<i>98.63</i>	98.17
1.5	89.08	92.43	92.47	86.80	92.81	93.77	79.72	89.48	91.81	97.55	<i>98.59</i>	98.08
1.75	88.66	92.25	92.51	86.84	92.92	93.77	80.09	89.99	92.19	96.99	<i>98.42</i>	98.02
2	88.90	92.29	92.42	86.60	92.97	93.85	80.26	90.18	92.21	97.65	<i>98.68</i>	98.16
2.25	87.96	91.17	91.60	87.02	93.63	93.86	79.87	90.09	92.00	97.23	<i>98.54</i>	97.95
2.5	88.15	91.57	90.92	86.78	93.20	93.72	80.04	90.24	91.94	97.21	<i>98.57</i>	97.76
2.75	87.40	90.98	91.05	86.98	93.19	93.51	80.17	90.15	91.71	97.13	<i>98.60</i>	97.43
3	87.89	91.56	90.88	87.94	93.46	93.46	80.22	90.19	91.83	97.33	<i>98.77</i>	97.65
3.25	87.69	90.67	90.20	87.14	93.56	93.45	80.07	90.30	91.44	96.77	<i>98.31</i>	97.10
3.5	87.20	90.41	90.57	87.31	93.10	92.93	80.17	90.36	91.30	96.73	<i>98.35</i>	96.76
3.75	87.11	90.44	90.12	86.65	93.11	92.66	79.66	89.86	90.78	96.31	<i>98.09</i>	96.25
4	86.66	90.57	90.03	87.10	92.50	92.33	79.86	89.99	90.98	96.24	<i>97.83</i>	96.20
4.25	87.21	90.61	90.15	88.83	93.74	93.43	80.04	90.54	91.06	96.21	<i>98.14</i>	96.20
4.5	86.83	89.79	89.91	87.36	93.46	92.90	79.65	89.79	90.30	95.79	<i>97.63</i>	95.63
4.75	85.61	90.26	89.26	86.56	92.39	92.70	78.93	89.63	90.38	95.51	<i>97.18</i>	94.99
5	86.15	90.04	89.62	86.60	92.70	92.46	80.17	90.01	90.87	95.10	<i>97.22</i>	95.57
5.25	85.55	90.57	90.00	88.20	93.31	92.86	79.79	90.10	90.40	94.87	<i>96.94</i>	95.51
5.5	85.06	89.83	89.19	86.71	92.60	92.58	79.75	89.99	89.85	94.72	<i>97.01</i>	94.87
5.75	85.38	90.36	89.93	87.31	93.27	92.64	79.68	90.10	90.11	94.37	<i>96.50</i>	94.33
6	85.09	89.55	89.26	87.82	93.39	92.50	79.58	89.76	90.10	93.59	<i>96.41</i>	94.72
6.25	84.50	89.32	89.11	87.62	93.13	92.31	79.54	90.00	89.94	93.04	<i>95.92</i>	94.27
6.5	84.19	89.94	89.10	86.94	92.78	92.07	79.09	89.89	89.65	92.29	<i>95.33</i>	93.63
6.75	84.64	88.85	88.91	87.81	93.30	92.65	79.11	89.80	89.73	92.00	<i>95.26</i>	93.90
7	84.28	89.13	88.93	88.12	93.64	92.73	79.39	90.18	89.85	91.87	<i>95.13</i>	93.55

Tabla 3.1: Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando ventanado variable en el rango 0-7 segundos. En cada fila se incluyen los resultados para un tamaño de ventana específico. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, el mayor rendimiento obtenido para cada tamaño de ventana

La figura 3.1 permite observar con mayor claridad los resultados recogidos en la tabla 3.1. En este caso, la optimización del rendimiento refleja tendencias divergentes entre las metodologías, obteniendo tamaños de ventana de corta duración en el caso de *KNN* en todas sus variantes y de larga duración en algunos casos de *NB* (cuando se aplica la media de la señal exclusivamente o junto a su desviación estándar) o *NCC*, cuando se aplica el 2º conjunto de características seleccionado. El algoritmo *DT* muestra una tendencia similar a *KNN*, aunque con una diferencia muy importante: la eficiencia de clasificación disminuye para tamaños de ventana inferiores a 1.75-2 segundos, mientras que para *KNN* la eficiencia de clasificación permanece prácticamente constante para tamaños de ventana de corta duración a excepción de su variante *KNN3* que desciende en más de un 2% para el tamaño de ventana mínimo.

Intuitivamente, la tendencia que cabría esperar conforme disminuye el tamaño de ventana es una disminución de la eficiencia. Cuanto menor es el número de muestras de la ventana para definir una característica, más complicado resulta alcanzar una magnitud correcta. A priori, cuanto mayor sea el número de muestras, más fácil será la obtención exacta de la magnitud que define de dicha característica. Estas suposiciones pierden su validez observando las diferentes tendencias de las metodologías presentadas conforme varía el tamaño de ventana. El modo de funcionamiento de cada algoritmo está marcado por sus características e influye de manera determinante en este resultado.

DT basa sus decisiones de clasificación en tests sucesivos hasta adoptar una decisión final. Particularizado al reconocimiento de la actividad, y en concreto, a la interpretación de las características extraídas en la etapa de extracción de característica del *ARC*, su comportamiento ofrece indicios de que para tamaños de ventana excesivamente cortos no dispone de información suficiente para no confundirse entre las propiedades que definen a las actividades testeadas. Y, al contrario, una vez que dispone de información suficiente para el reconocimiento de la actividad, el aumento del tamaño de ventana perjudica por la adición de información que corrige sus estimaciones previas y confunde al clasificador. En nuestro estudio, se ha comprobado que, independientemente del conjunto de características seleccionadas, el tamaño de ventana óptimo para alcanzar la mayor eficiencia de clasificación global para el dataset de actividades estudiado se encuentra en torno a 1.5-1.75 segundos.

Como ya se explicó en la subsección 2.5.2.5 del capítulo 2, *NB* es un clasificador bayesiano ingenuo. La aceptación de la hipótesis de independencia entre las variables predictoras permite que las características que conforman los conjuntos de características aplicadas en este estudio sean independientes entre sí. El razonamiento inmediato con la asunción de esta simplificación permite explicar el comportamiento de las metodologías (*NB* considerando FS1, FS2 ó FS3). Para tamaños de ventana de corta duración (hasta 1-1.5 segundos), la eficiencia de clasificación

aumenta conforme se incrementa el tamaño de ventana. A partir de ese tamaño, el clasificador dispone de información suficiente para reconocer lo mejor posible cualquier actividad. Seguir aumentando el tamaño de ventana no aporta ningún beneficio y se observa una oscilación entre valores de rendimiento muy próximos. Esto quiere decir que no se alcanza un instante de saturación, sino que el valor óptimo de reconocimiento se puede encontrar en un tamaño de ventana cualquiera dentro del rango de oscilación. Las tendencias ofrecidas por las 3 variantes del algoritmo *NB* planteadas en este estudio presentan esta observación. La diferencia entre ellas radica únicamente en obtener una mayor eficiencia de clasificación para un mismo tamaño de ventana hasta el momento en el que comienzan las oscilaciones. Una vez que se sobrepasa el tamaño de ventana que cambia las tendencias, la adición de características no garantiza el aumento de la eficiencia, ya que las tendencias de *NB* con 2 y 5 características son similares.

El algoritmo *NCC* presenta, en términos generales, los peores resultados en eficiencia entre los paradigmas de clasificación planteados. Sin embargo, conforme se agregan características en el reconocimiento, el rendimiento mejora considerablemente para un mismo tamaño de ventana. De forma similar a *NB*, el algoritmo entra en una zona de oscilación en la que se localiza la eficiencia óptima de reconocimiento de cada variante considerada. Si bien para el reconocimiento con la utilización de la media de la señal apenas se alcanza el 80.26 % de eficiencia en 2 segundos, la introducción de más características supone un aumento hasta el 92.21 % con las 5 características seleccionadas y el mismo tamaño de ventana. Dado que este algoritmo se basa en la comparativa de la muestra con el centroide más cercano, cuanto más información disponga el clasificador mayor será la probabilidad de acercar el valor de la característica al valor del centroide. Aún así, y como ya se ha indicado, encuentra un punto de inflexión al igual que *NB*, puesto que la agregación de más información (alargar la ventana más de lo necesario) no supone beneficio alguno en la tasa de reconocimiento del clasificador.

Por último, se obtienen los mejores resultados para el algoritmo *KNN*. El comportamiento de este clasificador dista por completo del resto de paradigmas planteados aún guardando alguna semejanza con *DT* en tendencias y tamaños de ventana de corta duración. La reducción del tamaño de ventana no supone un perjuicio en la eficiencia de clasificación para *KNN*, sino que además alcanza sus mejores registros. Con una ventana de 0.5 segundos de duración empleando únicamente la media de la señal, consigue el valor óptimo de reconocimiento alcanzando el 97.92 %. Para las otras 2 variantes con 2 y 5 características, maximiza el reconocimiento en 1 segundo, con tasas de 98.84 % y 98.35 %, respectivamente. Una explicación sobre las tendencias obtenidas es la sensibilidad con la que *KNN* trata la granularidad de las actividades. Un reconocimiento global óptimo para un dataset amplio con una ventana de duración muy pequeña supone que el algoritmo se encuentra capacitado para extraer trazos de gestos, como si de una

	DT			NB			NCC			KNN		
Tamaño óptimo (s)	1.5	1.5	1.75	4.25	4.25	2.25	2	4.25	2	0.5	1	1
Variación (%)	+3.99	+2.88	+3.25	+1.01	+0.35	+1.36	+0.69	+0.78	+2.12	+4.33	+2.42	+3.64

Tabla 3.2: Obtención de la variación (en%) entre la aplicación de un tamaño de ventana de 6 segundos y el tamaño óptimo que maximiza la eficiencia de clasificación para cada metodología planteada. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados. En las filas se indica el tamaño de ventana óptimo para maximizar el rendimiento con cada metodología planteada y el porcentaje de variación existente con un tamaño de ventana de 6 segundos. El signo (+) indica que se obtiene una mejora con respecto al tamaño de ventana de 6 segundos

inferencia se tratase. El aumento del tamaño de ventana únicamente introduce más información para evaluar, contrastar y confundir las buenas decisiones que se pueden obtener con un tamaño de ventana mucho menor. Se trata de un resultado impactante y esperanzador, puesto que su aplicación supone tanto la reducción del tamaño de ventana (pudiendo tomar decisiones en un menor tiempo, a mayor velocidad) como una optimización del reconocimiento de la actividad (en este punto, una mejora en el reconocimiento global de las actividades recogidas en el dataset) aplicando un repertorio de características muy reducido (1 ó 2 características, ya que con 5 características no se mejoran los resultados de la 2^a variante considerada).

Si se comparan los diferentes algoritmos de clasificación considerados, *KNN* se desmarca por completo del resto de clasificadores hasta el punto de no existir igualdad para ningún tamaño de ventana utilizado. Por el contrario, el resto de clasificadores se comportan de manera distinta conforme varía el repertorio de características empleadas. *NCC* supera a *DT* para tamaños de ventana superiores a 2 segundos cuando se consideran 5 características. Por otro lado, las prestaciones de *NB* terminan siendo similares a *KNN* aplicando una ventana de larga duración (7 segundos), en gran medida debido a la penalización que sufre *KNN* conforme aumenta el tamaño de ventana que por mérito en el reconocimiento del algoritmo *NB*, que prácticamente oscila en sus 3 variantes una vez alcanzado el punto de inflexión. En cualquier caso, entre los 3 clasificadores (*DT*, *NB*, *NCC*), el que ofrece mejores resultados para tamaños de ventana de corta duración (0.25-1 segundo) es el algoritmo de árbol de decisión.

A lo largo de esta explicación, se ha apreciado que existen tamaños de ventana menores a 6 segundos que mejoran la eficiencia de clasificación para todas las metodologías planteadas. La tabla 3.2 muestra la mejora existente con respecto a la aplicación en la etapa de segmentación de una ventana de 6 segundos.

A partir de la tabla 3.1 se puede comprobar el porcentaje de variación existente entre la ventana referencia y la ventana de tamaño óptimo. En líneas generales, las mejoras obtenidas

	DT			NB			NCC			KNN		
Tamaño mínimo (s)	0.25	0.75	0.5	3	2.25	1	1.5	1.75	1	0.25	0.25	0.25
Variación (%)	+1.45	+1.36	+0.19	+0.12	+0.24	+0.82	+0.14	+0.23	+0.34	+4.15	+1.96	+1.37

Tabla 3.3: Obtención de la variación (en%) entre la aplicación de un tamaño de ventana de 6 segundos y el tamaño de ventana mínimo que supera la eficiencia del tamaño de ventana de referencia (6 segundos) para cada metodología planteada. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En las filas se indica el tamaño de ventana mínimo para superar la eficiencia de clasificación conseguida con 6 segundos y el porcentaje de variación existente con un tamaño de ventana de 6 segundos. El signo (+) indica que se obtiene una mejora con respecto al tamaño de ventana de 6 segundos

como consecuencia de la reducción del tamaño de ventana son considerables. Todas las metodologías consiguen maximizar su rendimiento a costa de disminuir la duración de la ventana, por lo que a priori ambas decisiones son ventajosas. Los paradigmas *NB* o *NCC* apenas consiguen aumentar la eficiencia de clasificación (aunque con selección de 5 características con *NCC* se supera el 2%), mientras que los clasificadores *DT* y *KNN* llegan a alcanzar mejoras en torno al 4%. Los tamaños de ventana óptimos para cada metodología son variados, especialmente con ventanas de corta duración para *DT* (1.5-1.75 segundos) y *KNN* (0.5-1 segundo), y tamaños dispares para *NB* (2.25 ó 4.25 segundos) o *NCC* (2 ó 4.25 segundos).

Adicionalmente, dado que se ha tomado como referencia el tamaño de ventana de 6 segundos, se presenta otra tabla 3.3 en la que se indica cuál es el tamaño de ventana mínimo que permite superar la eficiencia del tamaño de ventana de referencia.

A la vista de los resultados, se puede apreciar que se puede conseguir una eficiencia de clasificación superior para un tamaño de ventana muy inferior a 6 segundos. El caso óptimo se refleja para el clasificador *KNN* puesto que, con este criterio, consigue dicha mejora con el tamaño de ventana de menor duración considerado: 0.25 segundos. También consigue este hecho el algoritmo de árbol de decisión utilizando sólo la media de la señal, así como para 2 y 5 características lo consigue con 0.75 y 0.5 segundos, respectivamente. Aplicando los clasificadores *NB* y *NCC* con las 5 características se necesita un tamaño de ventana de 1 segundo. El tamaño de ventana de mayor duración aparece para *NB* con la media de la señal, que requiere 3 segundos. El resto de metodologías se encuentran en valores intermedios a los citados.

Una vez se ha obtenido una cuantificación global sobre la eficiencia en el reconocimiento del conjunto de actividades procesadas, el estudio se enfoca hacia el análisis y observación de la influencia que produce el tamaño de ventana sobre la identificación de cada actividad. En la sección 3.2 se presenta la métrica $F_1(F_1 - score)$ como una magnitud idónea para evaluar el reconocimiento de cada actividad monitorizada en el conjunto del dataset presentado. En las

tablas 3.4 y 3.5 se recogen los valores $F_1 - score$ máximos y tamaños de ventana asociados a dicho valor para cada metodología y actividad.

De las tablas 3.4 y 3.5 se extraen una serie de apreciaciones que permiten entender, a posteriori, los resultados gráficos que son mostrados en las figuras 3.2 y 3.3.

La tabla 3.4 recoge el $F_1 - score$ máximo que se puede alcanzar para una actividad aplicando las diferentes metodologías consideradas en este estudio, sin aplicar restricciones en la duración del tamaño de ventana. Posteriormente, la tabla 3.5 indica los tamaños de ventana asociados a la obtención de los $F_1 - score$ máximos de la tabla 3.4. Esta representación permite delimitar en el proceso experimental bajo qué condiciones se puede conseguir el mejor reconocimiento posible sobre cada actividad.

Analizando detenidamente las tablas, se observan muy buenos resultados (con $F_1 = 1$ o muy cerca) y otros nada halagüeños (con $F_1 \sim 0.5$ o incluso peor). En términos generales, algunas actividades que implican la realización de saltos presentan problemas de reconocimiento para ciertas metodologías. El peor reconocimiento se consigue con *saltar lateralmente* aplicando el algoritmo *NCC* con la media de la señal, que apenas alcanza un $F_{1,máximo} = 0.192$ con un tamaño de ventana de 0.5 segundos. Sin embargo, la misma actividad se puede reconocer con *KNN* y cualquier repertorio de características garantizando un $F_1 \geq 0.94$. Incluso utilizando únicamente la media de la señal, se puede considerar el tamaño mínimo de ventana aquí presentado: 0.25 segundos. Se tratan de resultados totalmente divergentes entre sí y que permiten realzar las propiedades específicas de funcionamiento del clasificador *KNN* para detectar trazos de actividades. Las actividades *saltar hacia arriba* y *saltar adelante y atrás* también presentan F_1 bajos para *NCC* con la media de la señal (0.510 y 0.554, respectivamente), y no consiguen alcanzar la mejora experimentada para la actividad *saltar lateralmente* con *KNN* con la media de la señal (0.805 y 0.857, respectivamente). Estas actividades alcanzan sus mejores reconocimientos para *KNN* empleando adicionalmente la desviación estándar de la señal, alcanzando para la métrica F_1 los valores 0.941 y 0.977, respectivamente. La otra diferencia entre ambos clasificadores se centra en el tamaño de ventana: una reducción desde los 5 segundos hasta 2 y 0.25 segundos para cada caso.

Casos similares a la actividad *saltar lateralmente* se presentan para actividades en las que se ejercitan ciertas extremidades: *elevación frontal de brazos*, *cruzar brazos frontalmente*, *rotación intensa de hombros*, *rotación interna de brazos y rodillas al pecho alternativamente*. Aplicando el clasificador *NCC* para reconocer estas actividades a partir únicamente de la media de la señal, los valores $F_1 - score$ son 0.609, 0.661, 0.665, 0.406 y 0.457, respectivamente. La mejora alcanzada con *KNN* aplicando sólo la media de la señal es importante (0.982, 0.989, 0.979, 0.970

	DT			NB			NCC			KNN		
1	0.965	0.970	0.965	0.971	0.984	0.982	0.900	0.965	0.977	0.979	0.985	0.984
2	0.935	0.968	0.958	0.919	0.960	0.926	0.802	0.934	0.902	0.987	0.993	0.986
3	0.985	0.987	0.975	0.989	0.971	0.937	0.930	0.951	0.923	0.996	0.997	0.995
4	0.586	0.763	0.775	0.697	0.780	0.787	0.510	0.637	0.681	0.805	0.941	0.900
5	0.599	0.813	0.838	0.637	0.744	0.819	0.554	0.671	0.730	0.857	0.977	0.964
6	0.629	0.821	0.837	0.498	0.780	0.860	0.191	0.695	0.775	0.942	0.982	0.971
7	0.896	0.934	0.934	0.999	1	0.995	0.953	0.969	0.973	1	1	0.998
8	0.741	0.843	0.799	0.906	0.877	0.904	0.865	0.806	0.786	0.969	0.990	0.980
9	0.932	0.949	0.955	0.842	0.974	0.979	0.738	0.850	0.959	0.994	0.999	0.993
10	0.971	0.979	0.977	0.963	0.997	0.985	0.874	0.929	0.946	1	1	0.999
11	0.956	0.950	0.966	0.977	0.996	0.993	0.879	0.989	0.981	0.992	0.999	0.992
12	0.961	0.969	0.971	0.961	1	0.994	0.895	0.923	0.902	0.999	1	0.996
13	0.971	0.986	0.971	0.979	0.987	0.985	0.894	0.994	0.995	0.994	1	0.995
14	0.921	0.926	0.920	0.900	0.942	0.932	0.820	0.921	0.917	0.986	0.994	0.984
15	0.974	0.976	0.970	0.984	0.988	0.987	0.821	0.999	0.977	0.998	1	0.999
16	0.965	0.950	0.963	0.980	0.959	0.951	0.947	0.995	0.967	0.999	1	0.999
17	0.933	0.947	0.972	0.968	0.975	0.973	0.859	0.971	0.976	0.999	1	0.988
18	0.781	0.868	0.859	0.868	0.857	0.861	0.733	0.855	0.914	0.984	0.983	0.989
19	0.923	0.949	0.944	0.863	0.901	0.949	0.764	0.831	0.923	0.987	0.993	0.987
20	0.889	0.930	0.933	0.719	0.913	0.915	0.609	0.791	0.898	0.982	0.989	0.976
21	0.893	0.951	0.959	0.876	0.949	0.956	0.813	0.916	0.909	0.982	0.994	0.992
22	0.897	0.930	0.920	0.806	0.916	0.952	0.661	0.841	0.948	0.989	0.998	0.988
23	0.863	0.897	0.897	0.779	0.848	0.880	0.665	0.802	0.868	0.979	0.996	0.975
24	0.796	0.853	0.858	0.490	0.824	0.907	0.406	0.643	0.772	0.970	0.982	0.963
25	0.959	0.967	0.962	0.903	0.967	0.990	0.819	0.874	0.970	0.999	0.997	0.991
26	0.701	0.755	0.832	0.734	0.901	0.947	0.457	0.884	0.921	0.944	0.982	0.976
27	0.907	0.847	0.877	0.881	0.939	0.950	0.892	0.892	0.923	0.961	0.975	0.968
28	0.911	0.924	0.921	0.882	0.942	0.928	0.831	0.922	0.930	0.998	0.998	0.991
29	0.900	0.920	0.927	0.920	0.977	0.967	0.854	0.965	0.948	0.993	0.996	0.995
30	0.947	0.945	0.936	0.992	0.994	0.994	0.876	1	0.984	1	1	1
31	1	1	1	1	1	1	1	1	1	1	1	1
32	0.961	0.967	0.966	0.967	0.994	0.992	0.914	0.977	0.981	0.999	0.999	0.999
33	0.987	0.989	0.987	0.999	0.991	0.990	0.996	1	0.990	1	1	0.998

Tabla 3.4: $F_{1,max}$ -score de las 12 metodologías planteadas para un tamaño de ventana especificado en la tabla 3.5. En cada fila se incluyen los resultados para una actividad específica del dataset estudiado. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3)

	DT			NB			NCC			KNN		
1	5.75	2.25	1.75	5.75	7	7	5	5.25	3	2.75	0.75	1.5
2	4.25	3	1.75	6.25	5	1.5	2.25	5	2.5	0.5	7	3
3	3.75	3.75	3.75	6	7	1.5	1.75	6	3.25	2	7	1
4	6.75	6.5	6.5	7	7	6.75	5	4.75	6.5	2	3	1.5
5	5	5.25	5	5.25	5.25	4.25	5.25	5.75	3.25	0.25	2	2
6	0.25	5.25	5	3.75	5.75	2.25	0.5	2.25	5.5	0.25	2	2
7	6.75	2	3	6	6	3	5.75	5.5	6	4.25	4.25	5.5
8	6.75	2	2	6	7	6.75	6.75	7	6.75	3	3	4
9	1.5	1.5	1.5	1	6.25	1.75	6.25	3.25	1.75	1.25	6.5	1
10	1.5	1.5	1.5	4.75	6	5.5	6	6.5	1.5	2.75	2.75	3
11	1.75	2.25	2	6	4.25	4.25	6.25	5.5	4.25	3.75	4.25	3.75
12	3.25	3.25	2	4.75	5.5	2.5	4.5	4	0.75	4.5	4	0.5
13	1	5	4.5	6.25	6.5	2.25	2	5.25	5	1.75	4.5	1.25
14	1.25	1.25	1.25	3	5	5	3.25	5.75	3	1.25	3.25	3
15	2.5	5.5	4	3.25	4.5	1.25	3.5	6.25	6.25	4	4.5	6.5
16	1.25	1.25	1.5	7	4.75	1	5	7	3	3.5	7	7
17	1	1.5	2.25	6	6.75	2.25	4.5	4.5	5.5	2.25	2.25	2.25
18	4.75	2	4	5	7	1.75	3	6.25	7	2.75	2.75	2.75
19	1.25	2.75	6.25	5	2.75	2.75	5.5	5.25	4.25	2.25	4.25	4.25
20	1	2.75	1	4.25	2.75	2.75	2	2.75	1.75	2.75	2.75	0.75
21	1	2.75	3.25	4.75	2.5	2.25	5.75	3.75	2.25	2.25	3.25	2.5
22	1.5	4.75	3.25	4.25	2.25	2.25	1	1	2	0.25	1.25	1
23	1	2.75	1.5	2.5	3	1.5	2.25	5.25	2.75	1.75	5.25	1.5
24	0.75	0.75	0.75	3.5	2.25	2.25	5.75	0.75	0.75	2	2.5	0.5
25	1	5.75	0.5	2.5	3.75	2	2.5	7	1.75	2.5	3.75	2
26	0.25	1.75	3.5	6.75	6.25	5.25	2	5.25	5.25	0.25	3.25	1.25
27	5.75	5.25	5.75	5.75	5.75	5.5	4.75	4	4	0.25	2.25	1
28	1.5	2	2	4	5	2	3.25	3.25	1.75	5	5	1.75
29	1	1.5	2.25	1.5	6	6	6.5	2.75	2.75	0.5	3	2.5
30	0.75	3.5	1.25	5.75	3.25	3	4	3.25	1.75	2.5	3.25	4.25
31	3.25	3.25	3.25	0.5	1	0.75	1	1	0.75	0.25	0.25	0.5
32	1.5	2	1.25	4	4.5	4.75	3.75	4.75	5.75	2.75	4	5.75
33	4.75	1	1	6.25	5.25	5.25	2.75	3.25	1	4	4.5	0.5

Tabla 3.5: Tamaños de ventana asociados a los $F_{1,máx}$ - score de las 12 metodologías planteadas en la tabla 3.4. En cada fila se incluyen los resultados para una actividad específica del dataset estudiado. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3)

y 0.944), consiguiendo incluso reducciones del tamaño de ventana hasta 0.25 segundos o en el peor de los casos, hasta 2.75 segundos.

Estos análisis ofrecen una perspectiva crítica sobre la elección adecuada de los factores de influencia y las consideraciones oportunas para el reconocimiento de la actividad. No se trata de catalogar *NCC* como un algoritmo deficiente, sino que para el reconocimiento de ciertas actividades no es el más adecuado. Tampoco de etiquetar *KNN* como el mejor algoritmo posible para el reconocimiento de la actividad, pero sí quizás el más sensible a los movimientos y el que ofrece en este estudio los mejores resultados. El análisis anterior refleja que incluso una actividad difícil de reconocer con un algoritmo puede verse beneficiada por otro. Hasta este punto, se han analizado los “peores casos” puesto que destacan en la tabla 3.4, pero también se pueden resaltar otros más positivos a priori.

Actividades como *rotación de rodillas* o *hacer bicicleta estática* destacan por conseguir maximizar su reconocimiento para algunas metodologías, aunque una actividad destaca por encima del resto de actividades del dataset: la actividad *remar*. Todos los clasificadores consiguen maximizar su reconocimiento empleando un tamaño de ventana inferior a 6 segundos. El clasificador *DT* necesita el mayor tamaño de ventana (3.25 segundos) empleando cualquiera de los conjuntos de características seleccionados. En cuanto al resto de clasificadores, se obtienen tamaños de ventana extremadamente cortos, requiriendo como máximo 1 segundo empleando el clasificador *NB* con 2 características y *NCC* con sólo la media de la señal o acompañada de la desviación estándar. *KNN* consigue reconocer la actividad con el menor tamaño de ventana planteado para este estudio (0.25 segundos) con 1 y 2 características, mientras que necesita 0.5 segundos para maximizar su reconocimiento con 5 características.

También se puede maximizar el reconocimiento de algunas actividades seleccionando adecuadamente ciertas metodologías evaluadas en este proyecto. Por ejemplo, la actividad *saltar abriendo y cerrando piernas y brazos* se puede reconocer perfectamente con cualquier metodología (el peor reconocimiento se obtiene con *DT* y 1 característica, con un $F_1 = 0.896$), pero maximiza su reconocimiento con *NB* y *KNN* con 1 y 2 características.

Las actividades en las que se trabaja específicamente el tronco pueden ser muy bien reconocidas según el F_1 máximo presentado. De las 10 actividades que conforman este grupo, el peor reconocimiento de una actividad con el clasificador *KNN* tiene lugar con *giro de extremidades superiores e inferiores en oposición* utilizando sólo la media de la señal y alcanza un $F_1 = 0.984$. Ésta actividad presenta los peores resultados, obteniendo $F_1 = 0.733$ para *NCC* y $F_1 = 0.781$ para *DT*, ambos utilizando únicamente la media de la señal.

En general, para todos los casos citados, la maximización del reconocimiento de la actividad se consigue disminuyendo el tamaño de ventana. Sin embargo, existen casos en los que se

consigue para el tamaño de ventana de referencia (6 segundos) como *saltar abriendo y cerrando piernas y brazos* (para *NB*) o que incluso requieren un tamaño de ventana mayor como *doblarse lateralmente* (6.25 segundos), *girar el tronco con los brazos extendidos* (6.5 segundos) o *doblarse lateralmente con un brazo arriba* (7 segundos). A priori, y dado que el estudio se ha centrado en comprobar el efecto del tamaño de ventana sobre el reconocimiento de las actividades intentando aplicarlo para la reducción de dicho parámetro, la optimización del rendimiento (y en este caso maximización) para tamaños de ventana mayores a 6 segundos podría suponer una contrariedad. Nada más lejos de la realidad, simplemente se indica que para reconocer con todas las garantías dichas actividades es recomendable utilizar tamaños de ventana de larga duración. En general, la interpretación inmediata que se debe hacer sobre una actividad para la que se maximice su reconocimiento en un tamaño de ventana específico es que no se extrae beneficio del aumento del tamaño de ventana a partir del tamaño de ventana óptimo. Si se optimiza (o en el mejor de los casos, se maximiza) el reconocimiento para una actividad con un tamaño de ventana de una duración muy corta (por ejemplo, inferior a 1 segundo), la interpretación inmediata que se debe hacer es que apenas existe margen de mejora en cuanto a su reconocimiento en menor tiempo. Por otro lado, si para una actividad se optimiza (o en el mejor de los casos se maximiza) su reconocimiento para un tamaño de ventana de larga duración (por ejemplo, superior a 5 segundos), la implicación directa que supone este resultado es que existe un margen amplio de mejora en cuanto a observar si es posible su reconocimiento en menor tiempo con ciertas garantías (un F_1 - score considerable).

Estos razonamientos permiten enfocar el estudio de cada actividad tomando los criterios mencionados en la sección 3.2: la **minimización del tamaño de ventana** y la **penalización del F_1 - score**.

En primer lugar, se presentan las figuras 3.2 y 3.3 que reflejan el tamaño de ventana mínimo que se puede seleccionar para un valor F_1 - score determinado para cada actividad del dataset y con una metodología concreta.

Gran parte de la tabla 3.5 se refleja en cada una de las gráficas debido a que la parte superior de las barras coincide, en muchos casos, con el tamaño de ventana del $F_{1,máx}$ - score. Que no coincida exactamente se debe al diseño realizado para las representaciones expuestas. Las gráficas representan la magnitud F_1 - score en incrementos de 0.01 en el rango [0 1]. Dado que el criterio seleccionado para la representación es la selección del mínimo tamaño de ventana que cumple la exigencia marcada por F_1 - score, en algunos casos existirá un tamaño de ventana inferior al obtenido para el $F_{1,máx}$ - score de la tabla 3.5. Esta apreciación podrá verse con mayor claridad en las figuras asociadas al estudio de la penalización del F_1 - score.

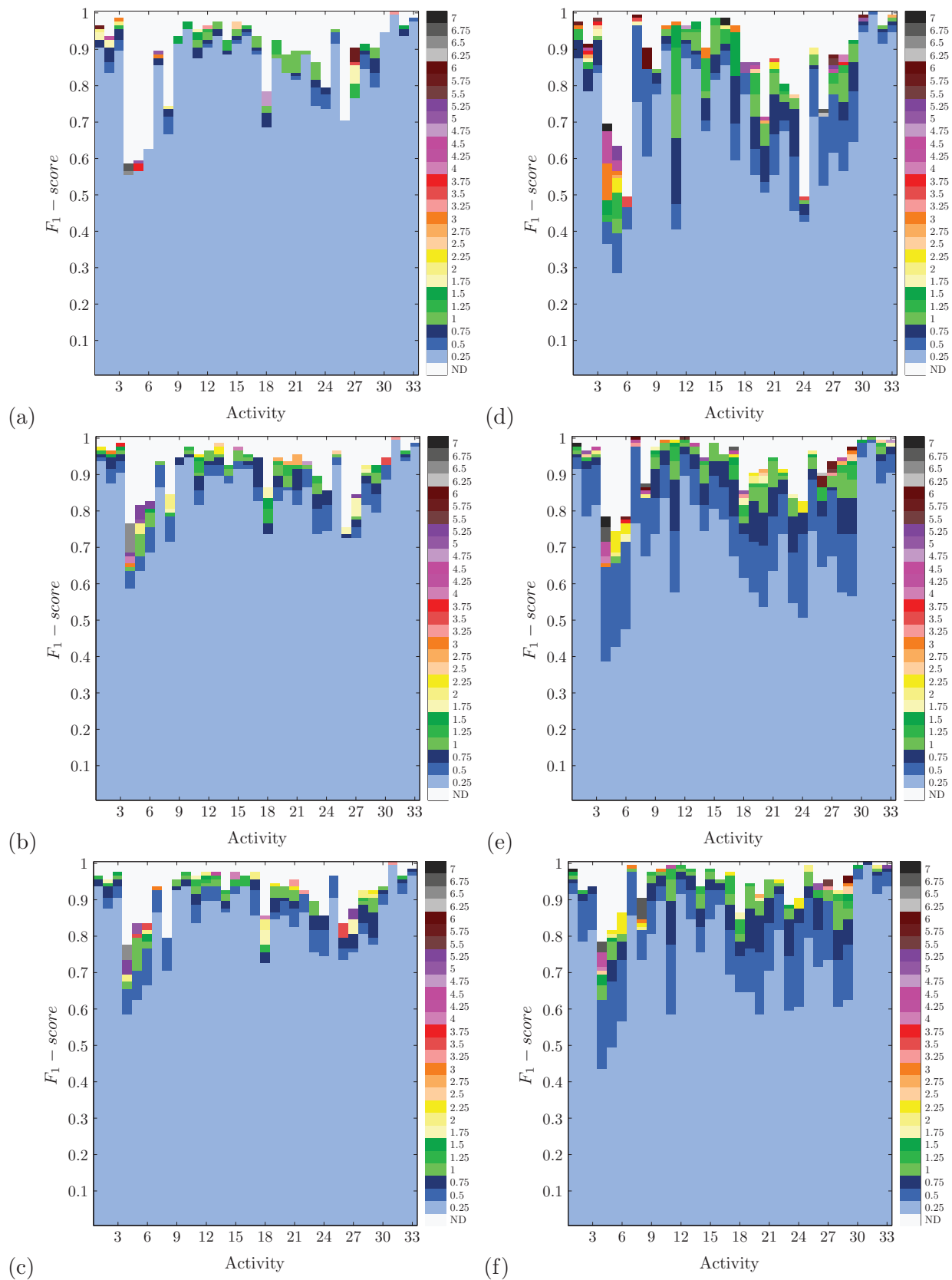


Figura 3.2: Tamaño de ventana mínimo necesario para un F_1 -score determinado según la metodología seleccionada (algoritmo de clasificación-características empleadas): a) DT-FS1 b) DT-FS2 c) DT-FS3 d) NB-FS1 e) NB-FS2 f) NB-FS3

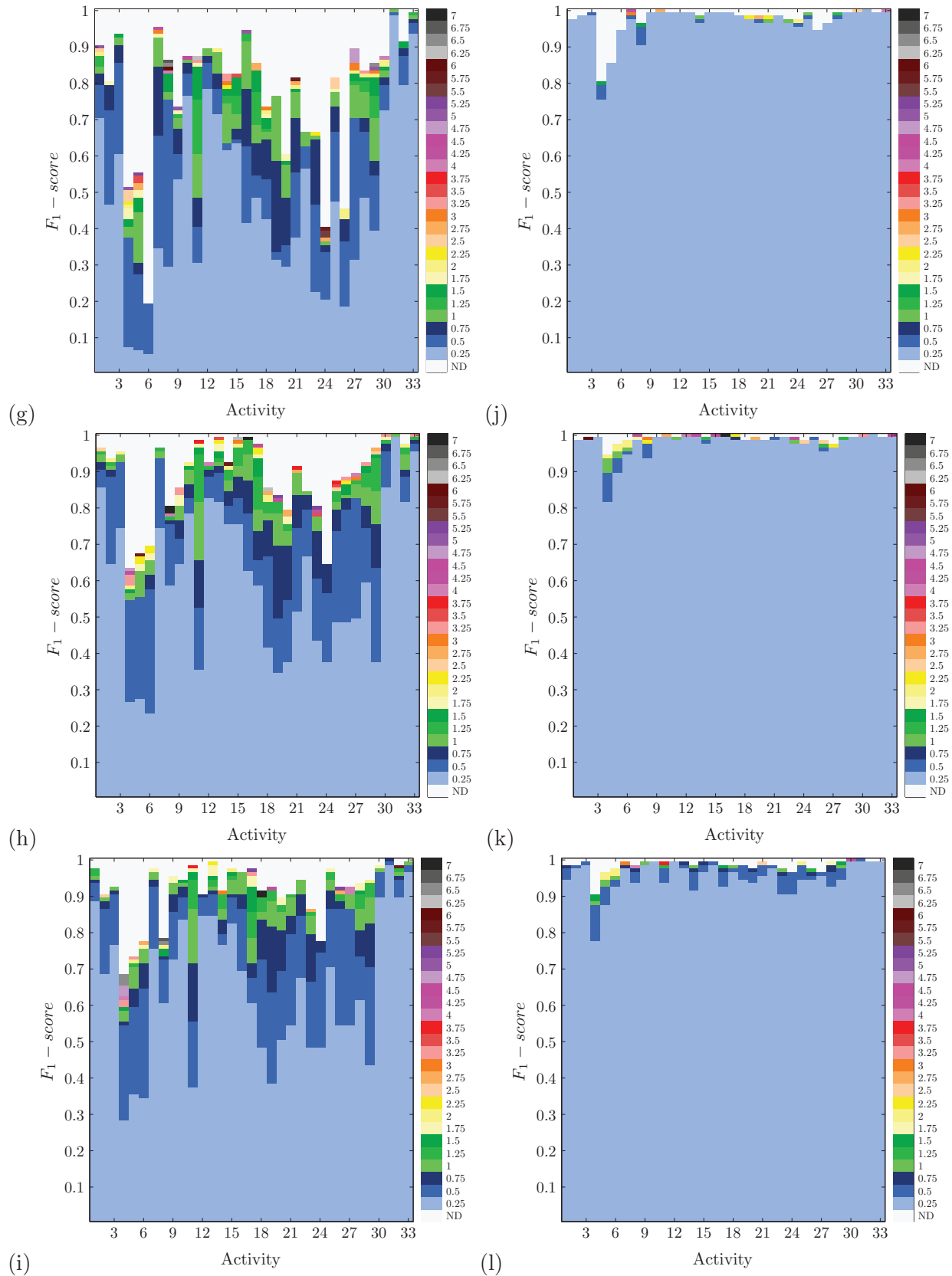


Figura 3.3: Tamaño de ventana mínimo necesario para un F1-score determinado según la metodología seleccionada (algoritmo de clasificación-características empleadas): g) NCC-FS1 h) NCC-FS2 i) NCC-FS3 j) KNN-FS1 k) KNN-FS2 l) KNN-FS3

En cuanto al análisis específico de las figuras 3.2 y 3.3, se pueden extraer diferentes observaciones. Dependiendo de la metodología empleada, se observa un dominio amplio del tamaño de ventana mínimo expuesto para este proyecto: 0.25 segundos. Principalmente destaca el clasificador *KNN* en este aspecto, ya que se consigue alcanzar $F_1 \geq 0.9$ para todas las actividades, a excepción de ejercicios problemáticos explicados anteriormente que suponen la realización de saltos como *saltar hacia arriba* y *saltar adelante y atrás*, rondando $F_1 \leq 0.85$. También se pueden comprobar casos especiales comentados sobre la tabla 3.4. Por ejemplo, para el clasificador *KNN* usando 2 características para el reconocimiento de la actividad *doblar lateralmente con un brazo arriba*, maximiza F_1 para un tamaño de ventana de 7 segundos. Efectivamente, observando la figura 3.3 (k) se comprueba dicho resultado, pero resulta que si se desea aplicar un tamaño de ventana inferior, inmediatamente se reduce drásticamente el tamaño de la ventana hasta 0.25 segundos asumiendo un F_1 comprendido entre 0.99 y 1 sin alcanzar en ningún momento la maximización.

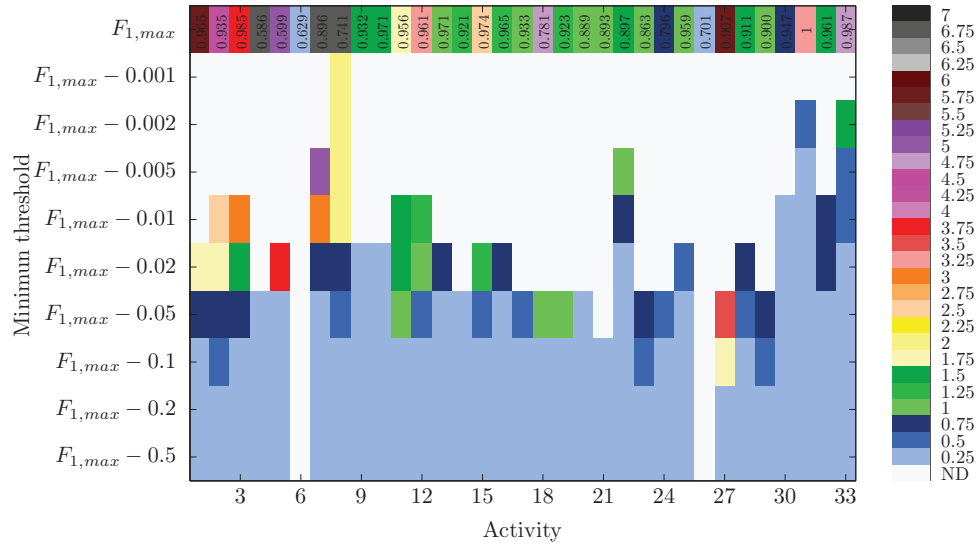
Al igual que este resultado, se pueden apreciar casos similares para las 12 metodologías expuestas. En la figura 3.2 (b) (que recoge el comportamiento del clasificador *DT* empleando 2 características de la señal) se observa cómo la actividad *remar* se maximiza para 3.25 segundos (ya se vió en la tabla 3.4). Asumiendo una reducción mínima de F_1 para el reconocimiento de dicha actividad (nuevamente superior a 0.99), se consigue reducir el tamaño de ventana hasta 0.25 segundos. Sobre el resto de clasificadores, se puede observar la evolución de *NB* y *NCC* a partir de la introducción de características. Actividades como *saltar hacia arriba*, *saltar adelante y atrás*, *saltar lateralmente* o *rotación moderada de hombros* mejoran paulatinamente el reconocimiento. En concreto, tomando las subfiguras (d), (e) y (f) de la figura 3.2 (correspondientes a *NB*), la actividad *rotación moderada de hombros* aumenta su reconocimiento si se observa el incremento de la magnitud $F_1 - score$: $F_1 \sim 0.5$ con un tamaño de ventana de 3.5 segundos y 1 característica, $F_1 \sim 0.82$ con un tamaño de ventana de 2.25 segundos y 2 características, y $F_1 \sim 0.91$ con un tamaño de ventana de 2.25 segundos y 5 características. De la misma forma, observando las subfiguras (g), (h) e (i) de la figura 3.3 (correspondientes a *NCC*), la actividad *saltar lateralmente* aumenta su reconocimiento si se observa la evolución de $F_1 - score$: $F_1 \sim 0.19$ con un tamaño de ventana de 0.5 segundos y 1 característica, $F_1 \sim 0.69$ con un tamaño de ventana de 2.25 segundos y 2 características, y $F_1 \sim 0.77$ con un tamaño de ventana de 2.75 segundos y 5 características. En el primer caso, existe una optimización tanto del $F_1 - score$ como una reducción del tamaño de ventana; mientras que para el segundo caso, la actividad consigue reconocerse un poco mejor a costa de incrementar el tamaño de ventana. Estos resultados puede observarse también entre las tablas 3.4 y 3.5. En términos generales, gran parte del abanico de actividades que integran el dataset pueden reconocerse con diferentes clasificadores aplicando

tamaños de ventana de corta duración. Este hecho se puede comprobar por inspección visual observando las tendencias cromáticas presentadas en las gráficas con tonos azules y verdes que suponen la aplicación de un tamaño de ventana máximo de 1.5 segundos.

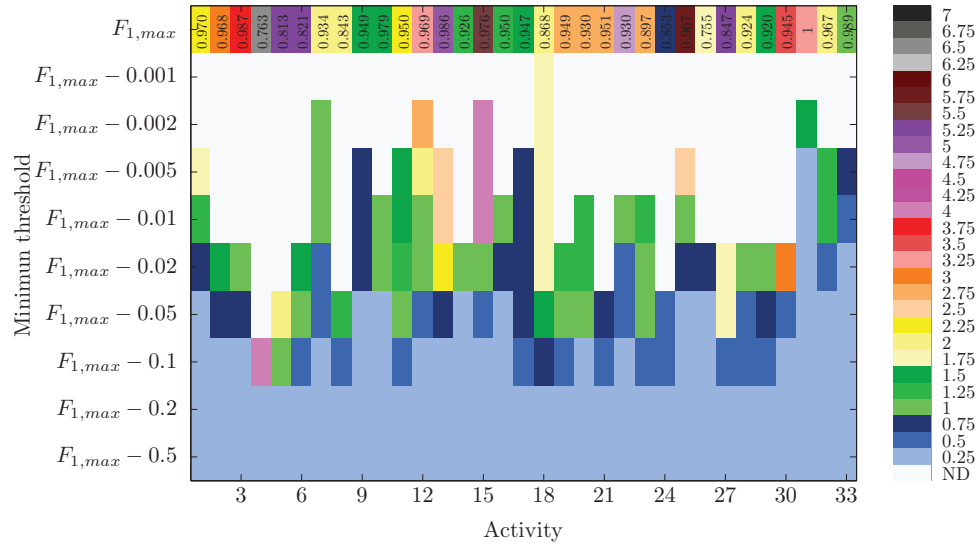
Una vez efectuado el estudio del tamaño de ventana mínimo para alcanzar un $F_1 - score$ determinado para cada actividad individualmente, se realiza un avance adicional en la comprobación del efecto del tamaño de ventana sobre el reconocimiento de cada actividad. La consideración de penalizaciones sobre el $F_1 - score$ permiten una reducción del tamaño de ventana necesario y ofrece una perspectiva más exacta sobre la influencia de este parámetro. Dos tipos de representaciones pueden realizarse dependiendo del interés particular en el análisis: *análisis por metodología* (agrupando las 33 actividades) y *análisis por actividad* (agrupando las 12 metodologías).

Las figuras 3.4, 3.5, 3.6 y 3.7 resumen las penalizaciones obtenidas en la aplicación de cada clasificador con los 3 conjuntos de características. En ellas se pueden observar todas las actividades del dataset y la reducción que se produce en el tamaño de ventana necesario conforme se aplica la umbralización con el parámetro $F_1 - score$. Se han considerado 3 tipos de penalizaciones o umbrales:

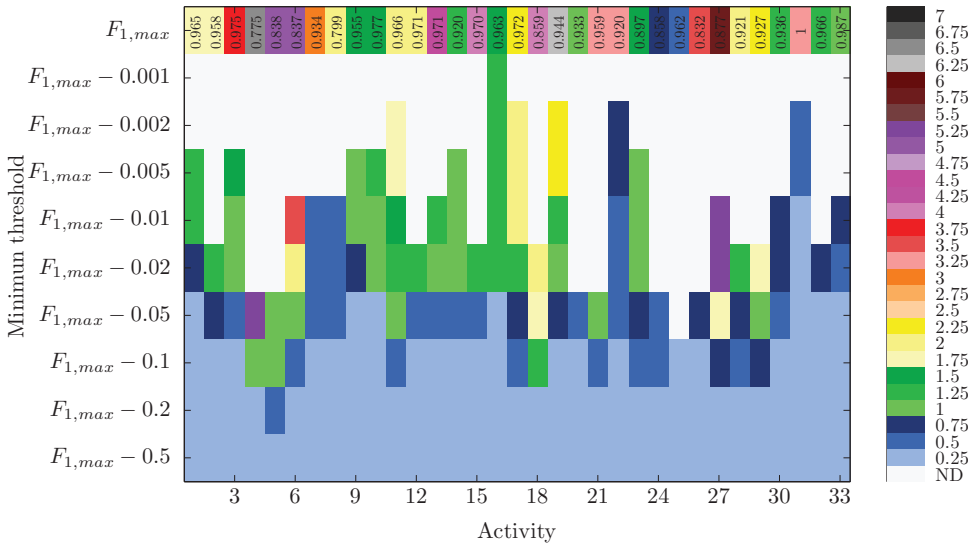
- ★ **Penalización mínima:** Permite comprobar qué actividades pueden reconocerse fielmente con una disminución del parámetro $F_1 - score$ de hasta 0.005. Concretamente, se consideran penalizaciones de 0.001, 0.002 y 0.005. Existen casos en los que la aplicación de un umbral mínimo modifica el tamaño de ventana del $F_1 - score$ máximo y supone una pérdida de reconocimiento despreciable.
- ★ **Penalización moderada:** Permite comprobar qué actividades pueden reconocerse fielmente con una disminución del parámetro $F_1 - score$ aplicando penalizaciones de 0.01, 0.02 y 0.05. Existen casos en los que la aplicación de este umbral supone la salida del mejor rango de reconocimiento ($F_1 \geq 0.85-0.9$).
- ★ **Fuerte penalización:** Permite comprobar qué actividades ofrecen un reconocimiento aceptable tras soportar una disminución del parámetro $F_1 - score$ de 0.1, 0.2 y 0.5. Se trata del caso más extremo y la única ventaja recae en la obtención de un tamaño de ventana de muy corta duración.



(a)

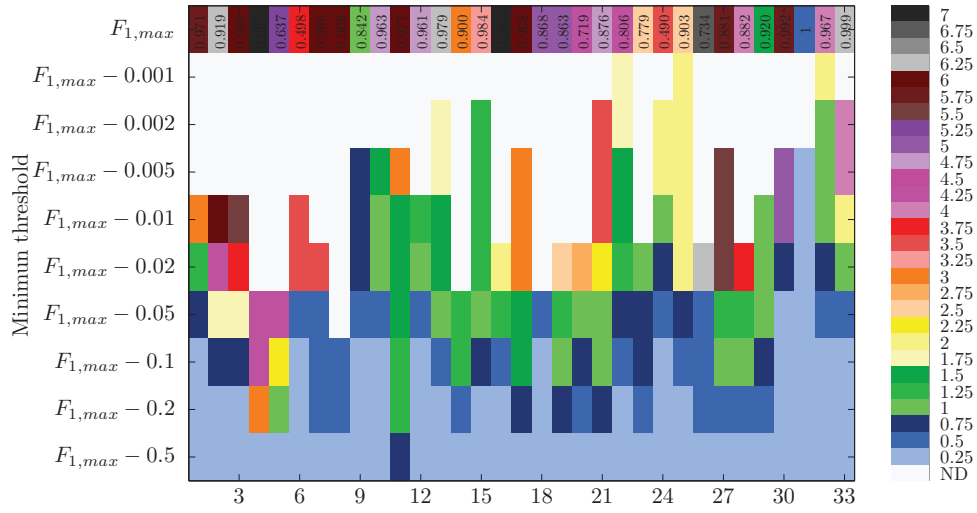


(b)

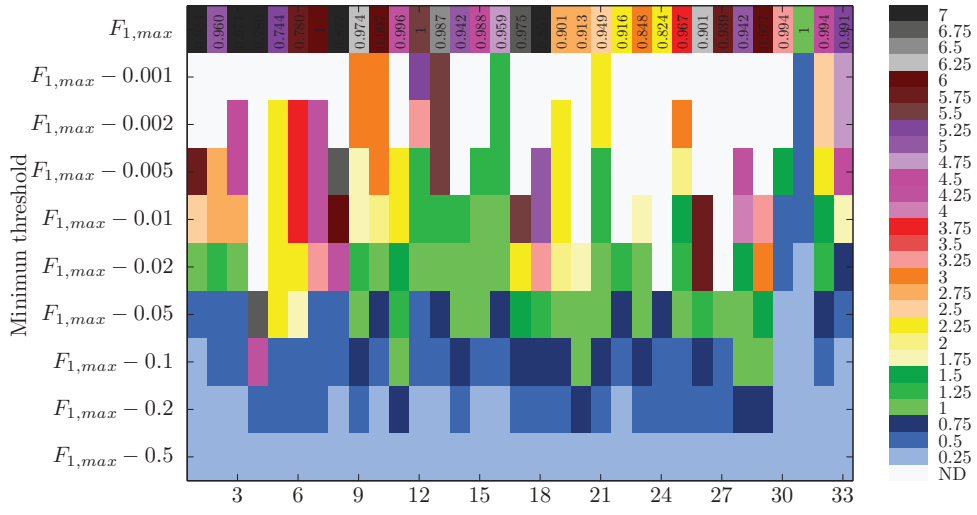


(c)

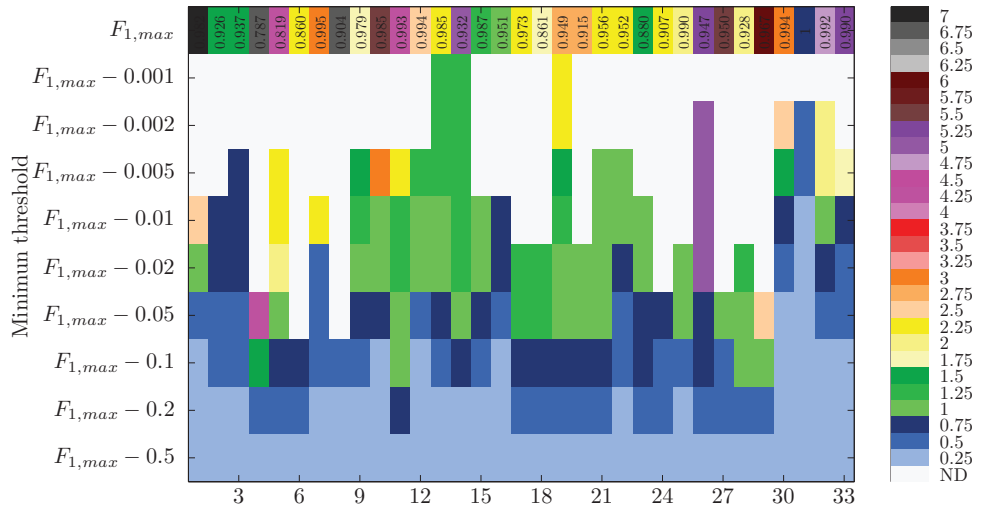
Figura 3.4: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score sobre las metodologías de DT: a) DT-FS1 b) DT-FS2 c) DT-FS3. En la fila superior se indican los F_1 - score máximos alcanzados para cada metodología



(d)

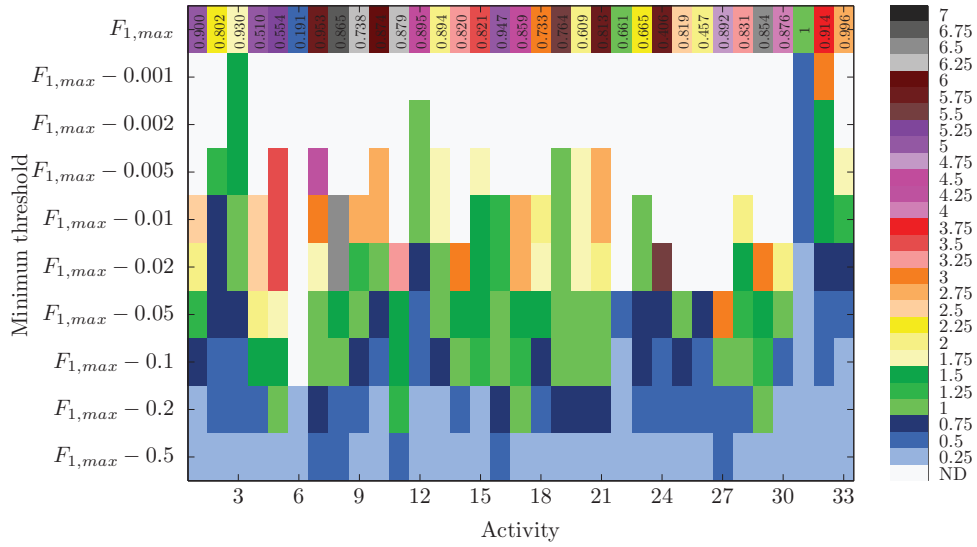


(e)

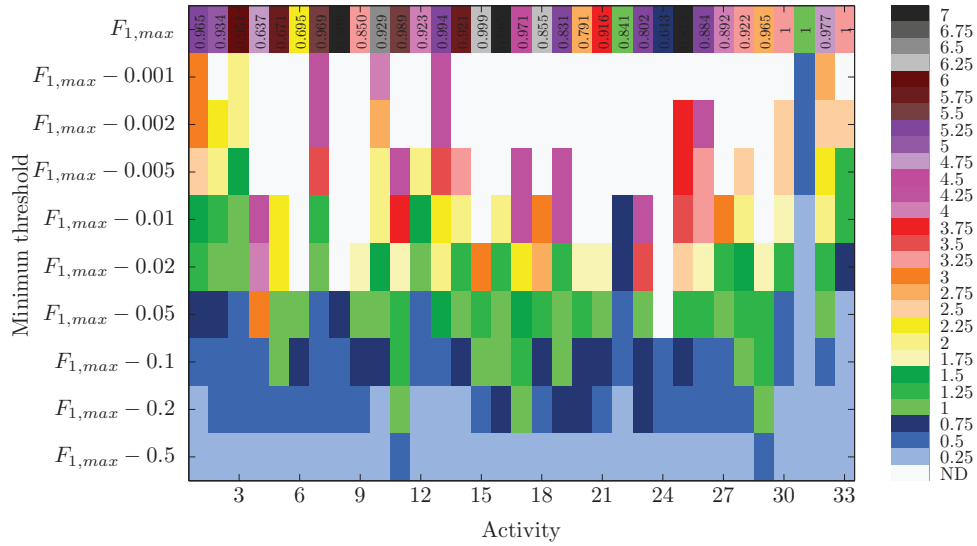


(f)

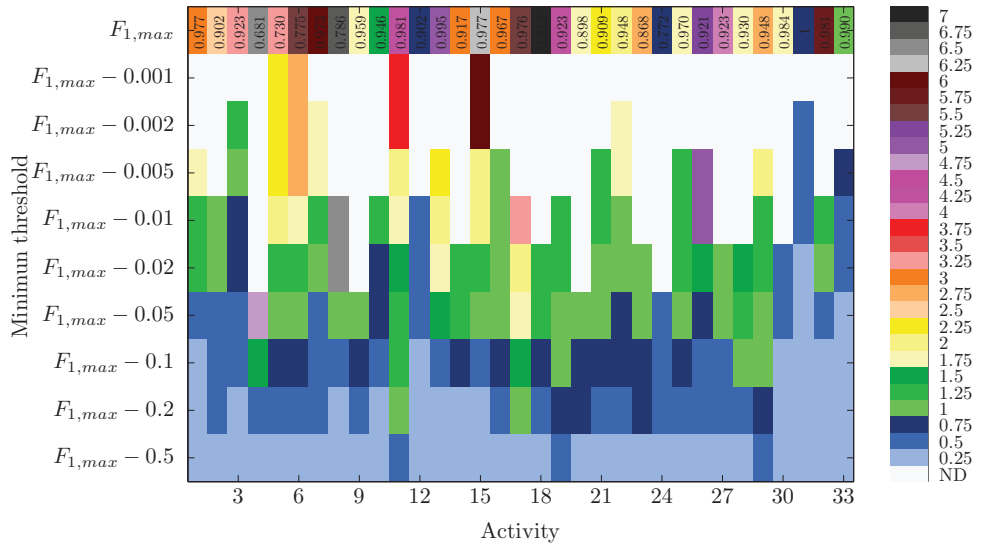
Figura 3.5: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score sobre las metodologías de NB: d) NB-FS1 e) NB-FS2 f) NB-FS3. En la fila superior se indican los F_1 - score máximos alcanzados para cada metodología



(g)



(h)



(i)

Figura 3.6: Tamaño mínimo de ventana para diversas tolerancias de F_1 – score sobre las metodologías de NCC: g) NCC-FS1 h) NCC-FS2 i) NCC-FS3. En la fila superior se indican los F_1 – score máximos alcanzados para cada metodología

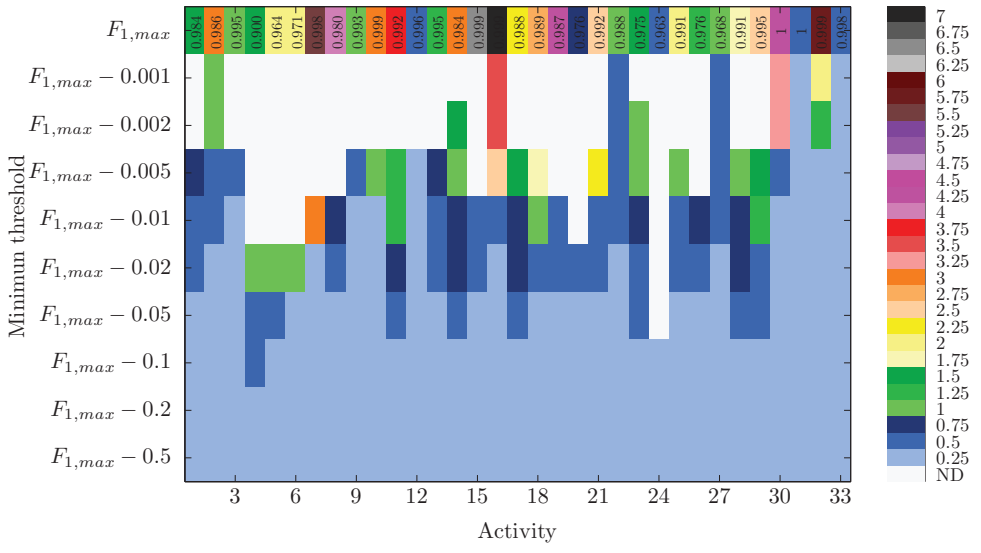
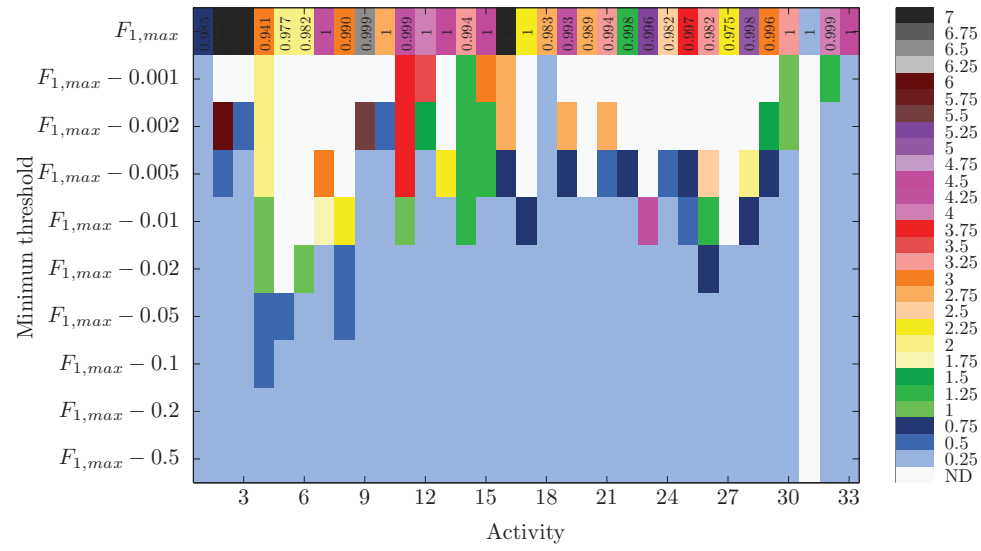
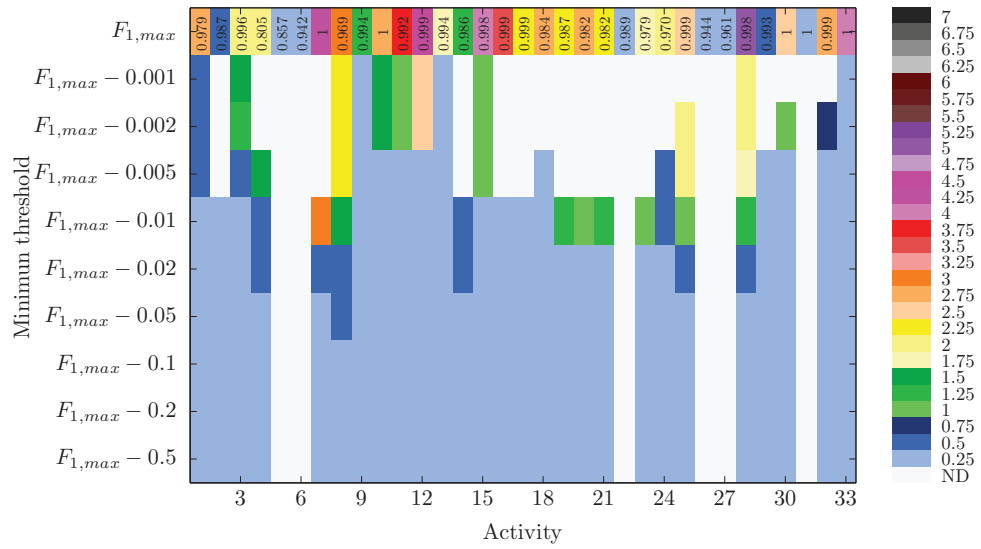


Figura 3.7: Tamaño mínimo de ventana para diversas tolerancias de F_1 – score sobre las metodologías de KNN: j) KNN-FS1 k) KNN-FS2 l) KNN-FS3. En la fila superior se indican los F_1 – score máximos alcanzados para cada metodología

Las figuras presentadas para el análisis por metodología destacan por la facilidad para interpretar sin un estudio profundo la información más relevante de cada clasificador y su comportamiento frente al tamaño de ventana. Cada figura, desde la 3.4 hasta 3.7, está dedicada a un clasificador concreto con los conjuntos de características seleccionadas.

En la figura 3.4 se observan las penalizaciones aplicadas a $F_{1,máx} - score$ para el clasificador *DT*. Cuando solamente se utiliza la media de la señal, se pueden hacer 3 observaciones principales. Las actividades en las que se trabaja específicamente el tronco y las extremidades superiores pueden ser reconocidas con tamaños de ventana de corta duración, destacando la tendencia cromática de tonos verdosos (duración de hasta 1.5 segundos) con $F_{1,máx} - score \geq 0.9$ salvo excepciones. Por otro lado, la actividad *saltar la cuerda* presenta un $F_{1,máx} = 0.741$ con una ventana de 6.75 segundos, pero con una penalización sobre $F_1 \leq 0.001$, se consigue reducir el tamaño de ventana a 2 segundos. Más drástico es el caso de la actividad *remar*, cuyo reconocimiento ha sido maximizado para una ventana de 3.25 segundos y que con una penalización máxima de 0.005 sobre $F_{1,máx}$, puede ser identificada con el tamaño de ventana mínimo de este estudio. Aplicando la misma penalización se puede reducir el tamaño de ventana para el reconocimiento de la actividad *hacer bicicleta estática* desde 4.75 a 0.5 segundos con un $F_1 - score \sim 0.98$. Por último, se comprueba que 26 de las 33 actividades pueden ser reconocidas con un tamaño de ventana de hasta 0.75 segundos aplicando una penalización sobre $F_{1,máx} \leq 0.05$. Para 15 de las de 33 actividades se tendría un $F_1 - score \geq 0.88$ tras la penalización indicada. Este resultado indica que, aplicando un tamaño de ventana bastante bajo, se obtendrían reconocimientos aceptables para casi la mitad de las actividades del dataset. En el caso de utilizar 2 características, hasta 16 actividades obtienen un $F_1 \sim 0.9$ tras la penalización de 0.05, pudiendo utilizar un tamaño de ventana de 0.75 segundos como máximo. Sobre la consideración de 5 características, destaca el reconocimiento de la actividad *saltar abriendo y cerrando piernas y brazos* que permite reducir el tamaño de ventana de 3 a 0.5 segundos aplicando una penalización de 0.01 sobre $F_{1,máx} - score = 0.934$. Otras actividades que destacan por una reducción considerable del tamaño de ventana son *elevación lateral de brazos* (de 6.25 a 2.25 con una penalización mínima de 0.002 sobre $F_{1,máx} = 0.944$ o incluso aplicando un umbral de 0.02 que permite reducir el tamaño hasta 1 segundo) y *cruzar brazos frontalmente* (de 3.25 a 0.5 segundos con una penalización de 0.01 sobre $F_{1,máx} = 0.92$)

Sobre los resultados reflejados por la figura 3.5 se extrae una percepción inmediata: la necesidad global de un tamaño de ventana de larga duración para identificar las actividades del dataset con ciertas garantías de reconocimiento (según las tendencia cromática de tonos marrones y violetas, se requieren ventanas de 4.5 segundos al menos) con el clasificador *NB*. A partir de la penalización y considerando únicamente la media de la señal, entre las actividades

que ofrecen un reconocimiento destacable se encuentran *hacer bicicleta elíptica* (para la que se puede reducir la ventana de 4 a 0.5 segundos asumiendo una penalización de 0.05 para ser reconocida con $F_1 \sim 0.92$) y *hacer bicicleta estática* que permite reducir también la ventana de 6.25 a 0.5 segundos para reconocerse con $F_1 \sim 0.95$. Si se emplean 2 características para el reconocimiento de las actividades, destaca principalmente la reducción del tamaño de ventana de un grupo específico de ejercicios: el tronco. En conjunto, se puede pasar de tamaños de ventana del intervalo 5-7 segundos hasta 0.5-1 segundo aplicando una penalización de 0.05, obteniendo $F_1 \geq 0.91$. También resalta la optimización de la actividad *saltar abriendo y cerrando brazos y piernas*, para la que se consigue máximo reconocimiento con 6 segundos y que asumiendo una penalización de 0.05 se puede reducir la ventana hasta 0.5 segundos. En cuanto a la consideración de 5 características para el reconocimiento, destaca la mejora sobre el tamaño de ventana de la actividad *alcanzar los talones flexionando las rodillas*, debido a que se puede reducir de 5 a 1.5 segundos asumiendo la mínima penalización con un $F_{1,máx} = 0.932$.

La figura 3.6 está dedicada al clasificador *NCC*. Empleando únicamente la media de la señal, la aplicación de penalizaciones sobre reconocimientos generalmente bajos se convierte en una utopía. Sin embargo, este clasificador consigue reconocer muy bien las actividades del grupo exclusivo de *fitness*, así como *correr* o *saltar abriendo y cerrando piernas y brazos*. En concreto, para esta última actividad se puede reducir el tamaño de ventana de 6 a 1.5 segundos garantizando un $F_1 \geq 0.9$. Con la introducción de la desviación estándar, 19 de las 33 actividades alcanzan un $F_{1,máx} - score \geq 0.9$, aunque con requerimientos sobre el tamaño de ventana estrictos en su duración: al menos 4.5 segundos salvo casos excepcionales. Además de la actividad *remar*, sobre la que todas las metodologías obtienen resultados excepcionales, destaca la actividad *girar el tronco con los codos doblados*, para la cual se puede reducir el tamaño de ventana de 6.5 a 1.5 segundos asumiendo una penalización de 0.02 y garantizando un $F_1 \sim 0.91$. Un resultado similar se obtiene para la actividad *saltar abriendo y cerrando piernas y brazos*, con una reducción de la ventana desde 5.5 hasta 0.5 segundos asumiendo una penalización máxima de 0.05 para establecer $F_1 \sim 0.92$. Las actividades de *fitness*, con el máximo reconocimiento posible, se pueden reconocer asumiendo penalizaciones moderadas con tamaños de ventana mínimos. La consideración de 5 características para el reconocimiento permite incrementar el número de actividades del dataset con $F_{1,máx} - score \geq 0.9$ hasta 26. A diferencia de las 2 metodologías anteriores, la zona de penalizaciones refleja tonos más azulados y verdosos incluso para penalizaciones moderadas, por lo que la interpretación inmediata es que esta metodología permite reconocer un gran número de actividades con tamaños de ventana reducidos. Por ejemplo, para *doblarse lateralmente* se puede reducir el tamaño de ventana de 6.25 a 1 segundo asumiendo una penalización moderada de 0.05 y garantizando $F_1 \sim 0.93$. Otro caso de interés es *talones*

a la espalda (alternativamente), para la que se puede reducir la ventana de 4 a 1 segundos considerando una penalización de 0.02 para establecer $F_1 \sim 0.9$.

Por último, la figura 3.7 recoge los resultados del clasificador *KNN*. En las representaciones dedicadas a este clasificador se pueden observar detalles novedosos con respecto a los clasificadores anteriores. Para algunas actividades se obtiene una optimización de su reconocimiento para el tamaño mínimo de ventana seleccionado para este estudio, lo que supone que la zona de penalizaciones se encuentre despejada (en blanco). Considerando sólo la media de la señal, este hecho se produce para 6 actividades: *saltar adelante y atrás*, *saltar lateralmente*, *cruzar brazos frontalmente*, *rodillas al pecho (alternativamente)*, *talones a la espalda (alternativamente)* y *remar*. El $F_{1,máx} - score$ obtenido para cada una es 0.857, 0.942, 0.989, 0.944, 0.961 y 1, respectivamente. Para alcanzarse en 0.25 segundos, se tratan de resultados considerables exceptuando *saltar adelante y atrás*, aunque no se puede conseguir un resultado mejor dadas las condiciones. Exceptuando dicha actividad y otro tipo de salto (*saltar hacia arriba*), el resto de actividades del dataset alcanzan $F_{1,máx} - score \geq 0.94$. Asumiendo una penalización moderada de hasta 0.02 sobre $F_{1,máx} - score$, 26 de las 31 actividades consideradas anteriormente pueden reducir el tamaño de ventana hasta 0.25 segundos. Las otras 5 actividades (*saltar abriendo y cerrando piernas y brazos*, *saltar la cuerda*, *alcanzar los talones flexionando las rodillas*, *rotación interna de brazos y doblar las rodillas agachándose*) consiguen reducir la ventana hasta 0.5 segundos considerando la misma penalización. Observando las representaciones dedicadas a 2 y 5 características, parece que la primera sensación observando las tendencias cromáticas es que el tamaño de ventana requerido para la optimización de $F_{1,máx} - score$ es inferior aplicando sólo la media de la señal. Ciertamente, ocurre algo similar, con la salvedad de que el incremento del tamaño de ventana ha permitido maximizar para muchas actividades su reconocimiento. Para 2 características, este hecho tiene lugar para 9 actividades (excluyendo la actividad *remar*). Por otro lado, 2 actividades que implican movimientos de traslación (*trotar* y *correr*) optimizan su reconocimiento con un tamaño de ventana de 7 segundos. Este hecho hace pensar en que la repetitividad de la ejecución permite afinar en la optimización del reconocimiento de dichas actividades y que si el tamaño de ventana fuera superior, quizás se podría maximizar con este clasificador. Sin embargo, asumiendo una penalización mínima de 0.01, ambas actividades garantizan un $F_1 \geq 0.98$. Partiendo de la base de que todas las actividades del dataset presentan un $F_{1,máx} - score \geq 0.94$, con una penalización de 0.02 sobre dicho parámetro se consigue reducir el tamaño de ventana hasta 0.25 segundos para 28 actividades. Con la introducción de 3 características adicionales, se producen efectos divergentes para las actividades observando $F_{1,máx} - score$: algunas reducen dicho parámetro a costa de reducir el tamaño de ventana (como

trotar y correr) y otras incluso requieren aumentar el tamaño de la ventana (como *saltar abriendo y cerrando piernas y brazos* o *hacer bicicleta elíptica*, que pasan de 4.5 y 4 segundos a 5.75 segundos ambas, respectivamente). De estas apreciaciones, junto a la observación de la zona de penalización, se extrae una conclusión: la introducción de más características no proporciona mayor robustez al clasificador *KNN*, si no que incluso supone confusión y requerimientos de mayores tamaños de ventana para mantener el reconocimiento. Ciertamente, esta observación ya se intuía al comienzo de esta sección con el análisis de la figura 3.1. La especificación y el análisis particular sobre cada actividad permite confirmar estas hipótesis.

Tras este extenso análisis, en el que se han ofrecido varios detalles para la interpretación adecuada de las figuras 3.4, 3.5, 3.6 y 3.7, se procede a la observación particularizada de cada actividad. Este análisis, denominado como *análisis por actividad* (considerando como parámetro de interés la penalización del $F_{1,máx} - score$ para la reducción del tamaño de ventana), supone la presentación de 33 figuras individuales correspondientes a las 33 actividades consideradas en este proyecto. Para ser mostradas, se han agrupado de 3 en 3, por lo que las figuras de interés van desde 3.8 hasta 3.18. Más allá de analizar pormenorizadamente cada actividad, se han seleccionado algunas que presenten ciertas curiosidades y detalles importantes que ilustren y permitan al lector su correcta interpretación.

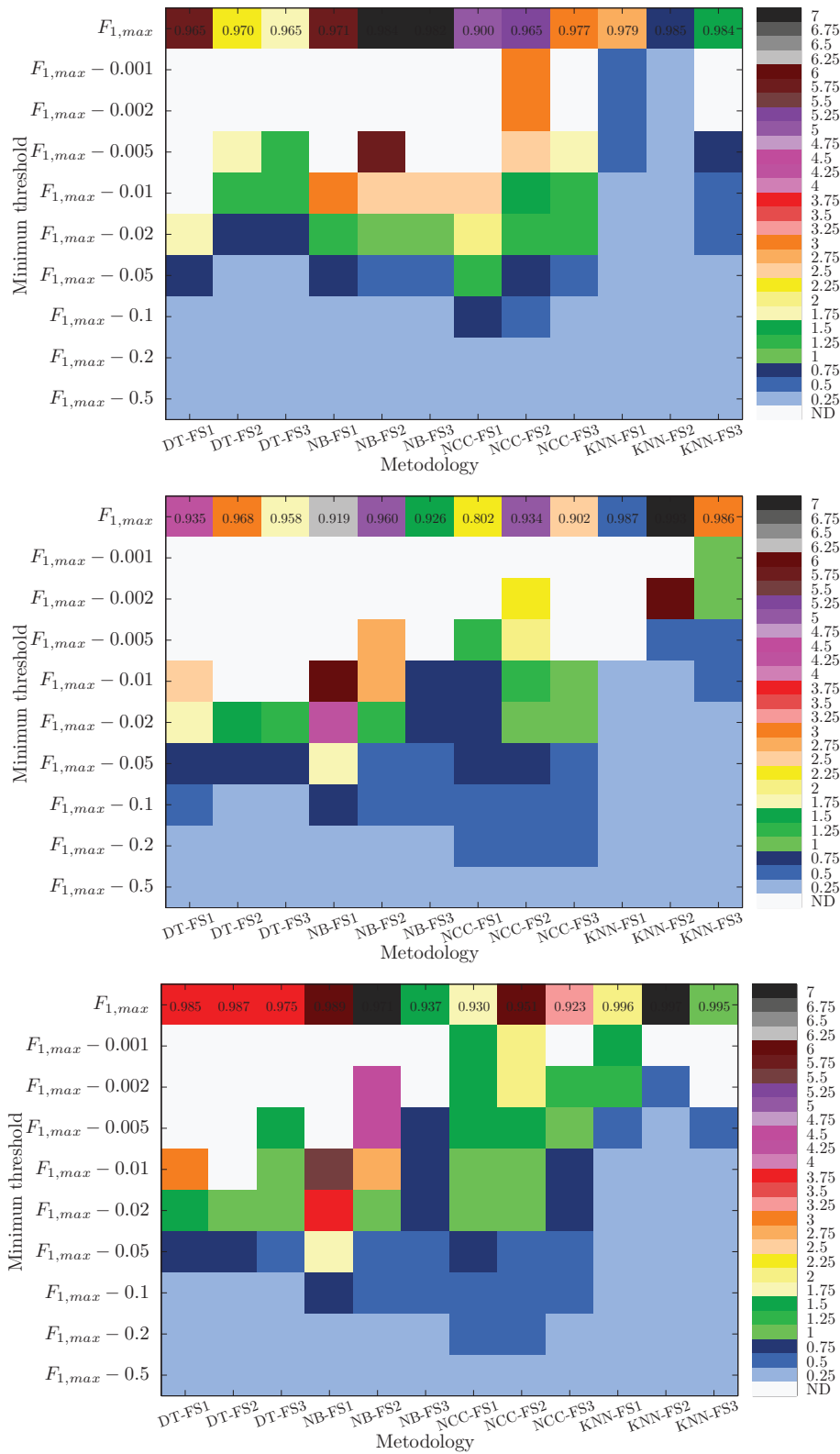


Figura 3.8: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *andar*, *trotar* y *correr* en orden descendente

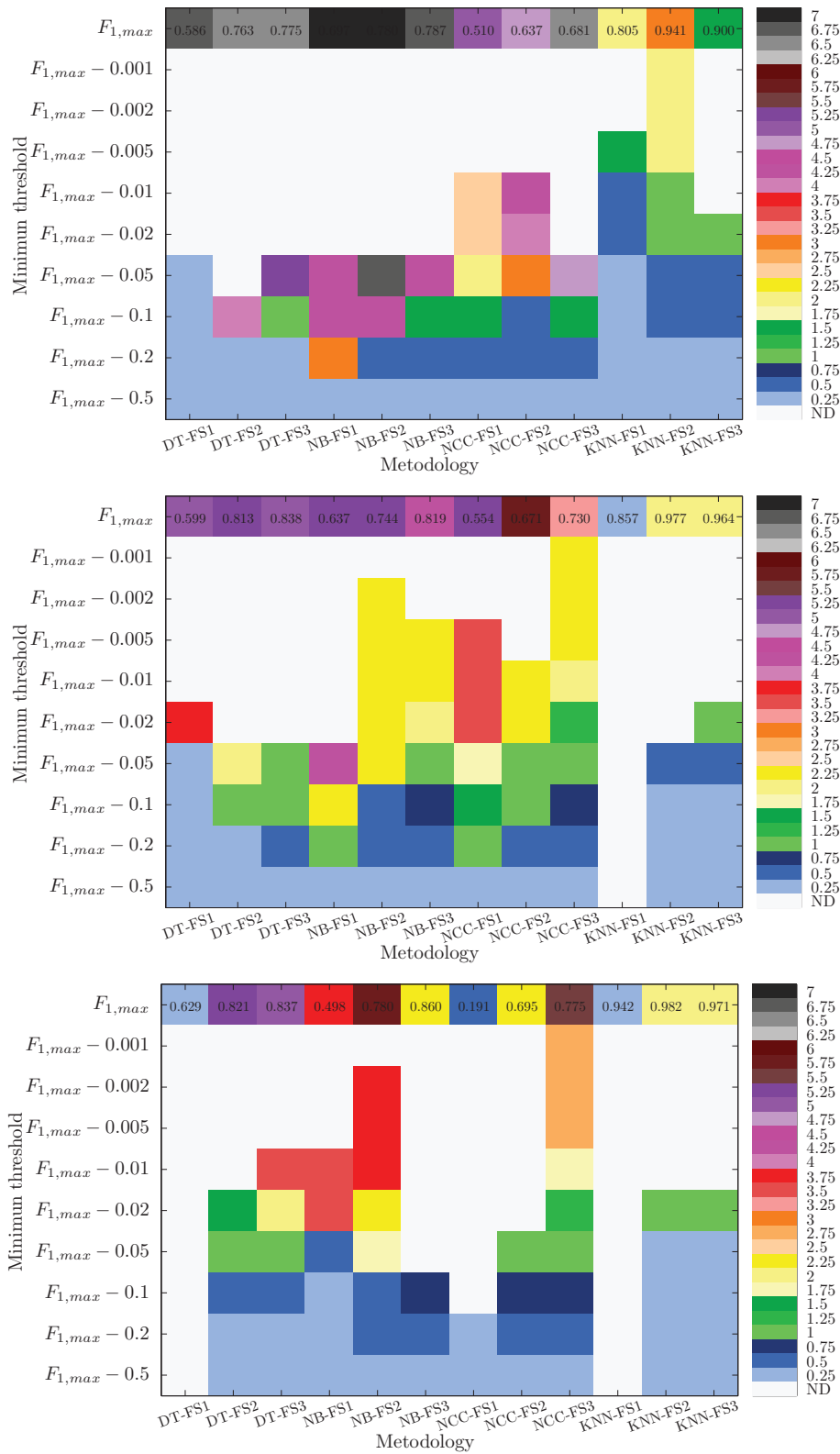


Figura 3.9: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *saltar hacia arriba*, *saltar adelante y atrás* y *saltar lateralmente* en orden descendente

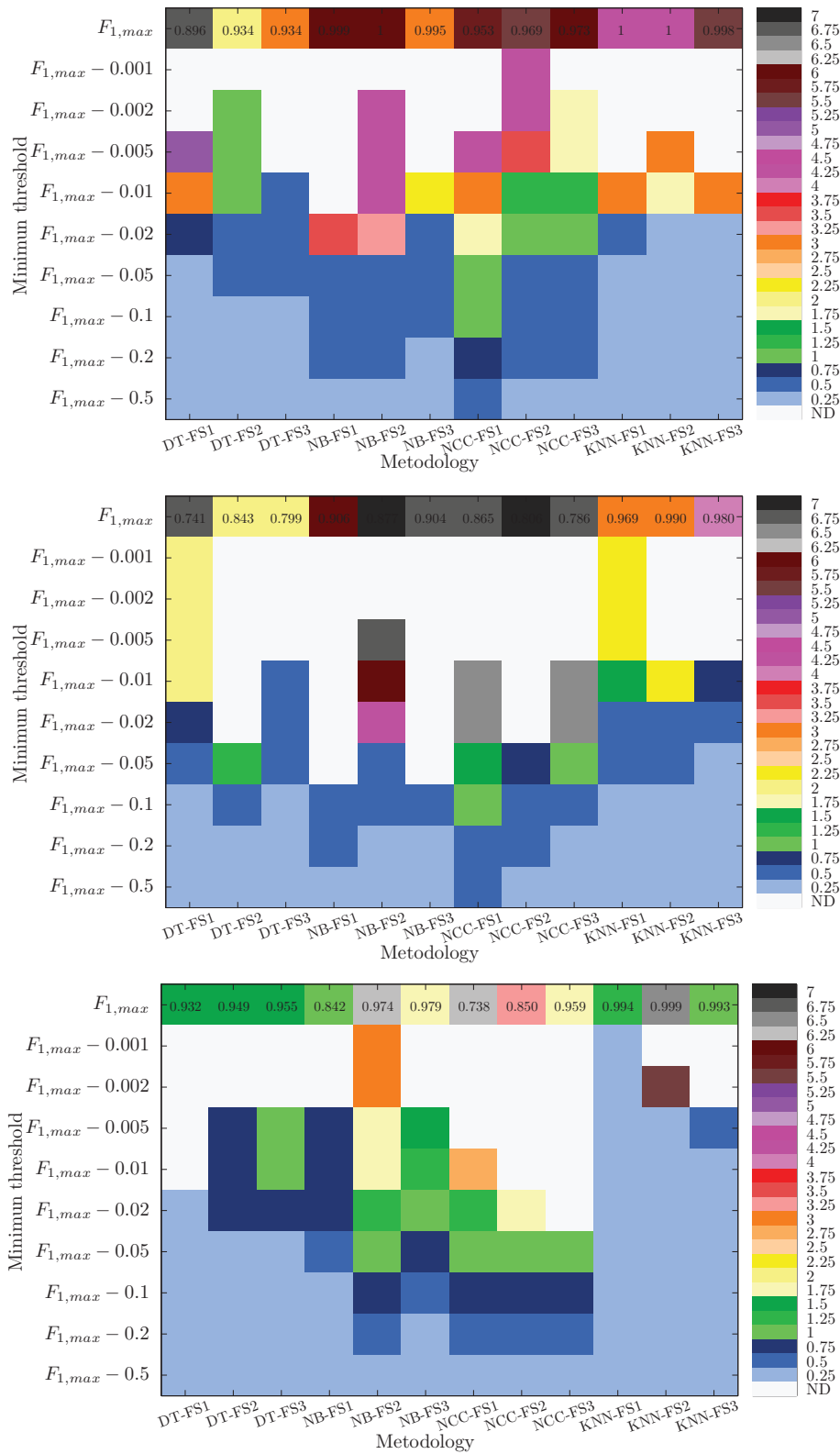


Figura 3.10: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *saltar abriendo y cerrando piernas y brazos*, *saltar la cuerda* y *girar el tronco con los brazos extendidos* en orden descendente

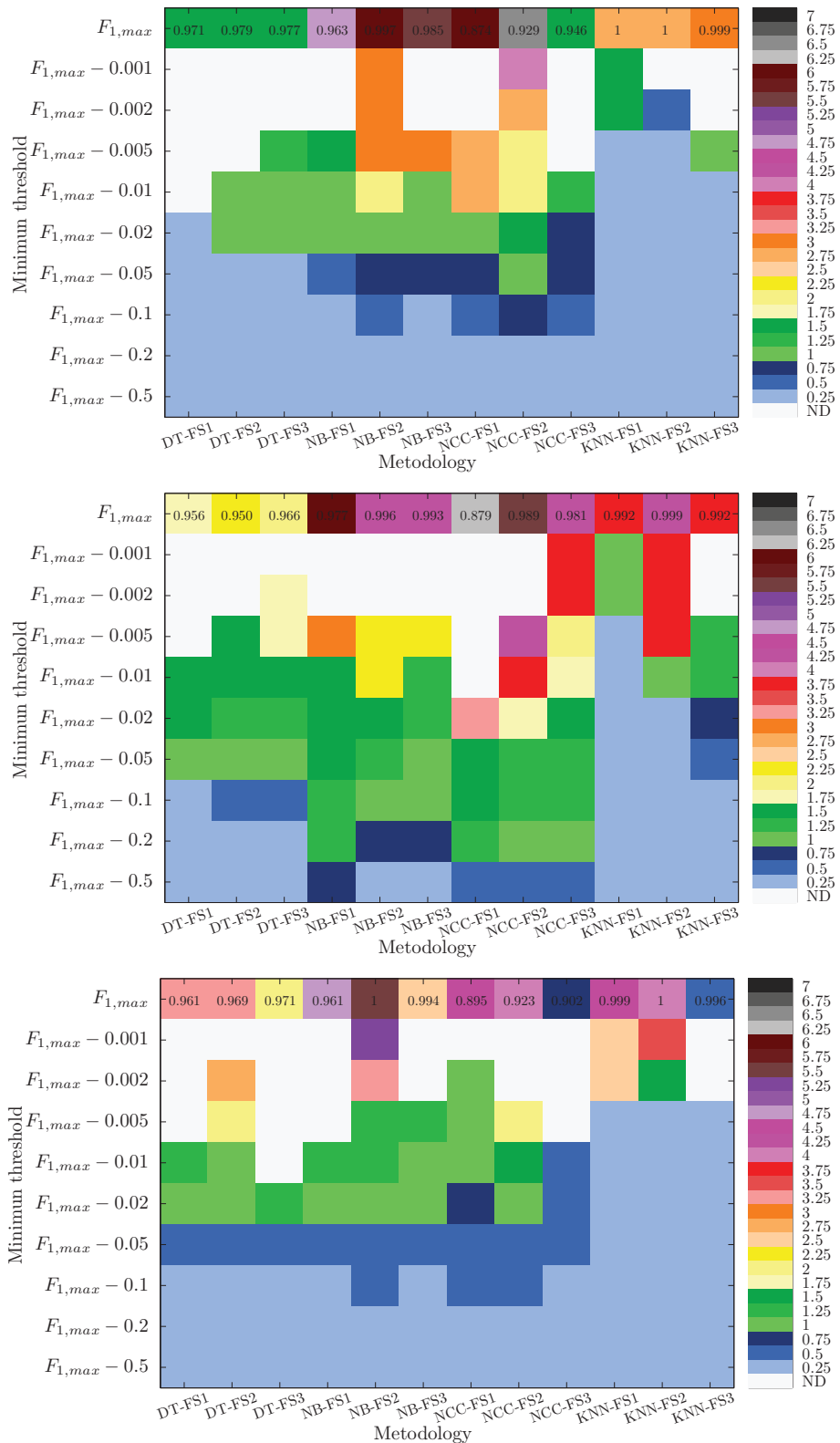


Figura 3.11: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: girar el tronco con los codos doblados, doblar la cintura hacia delante y rotación de cintura en orden descendente

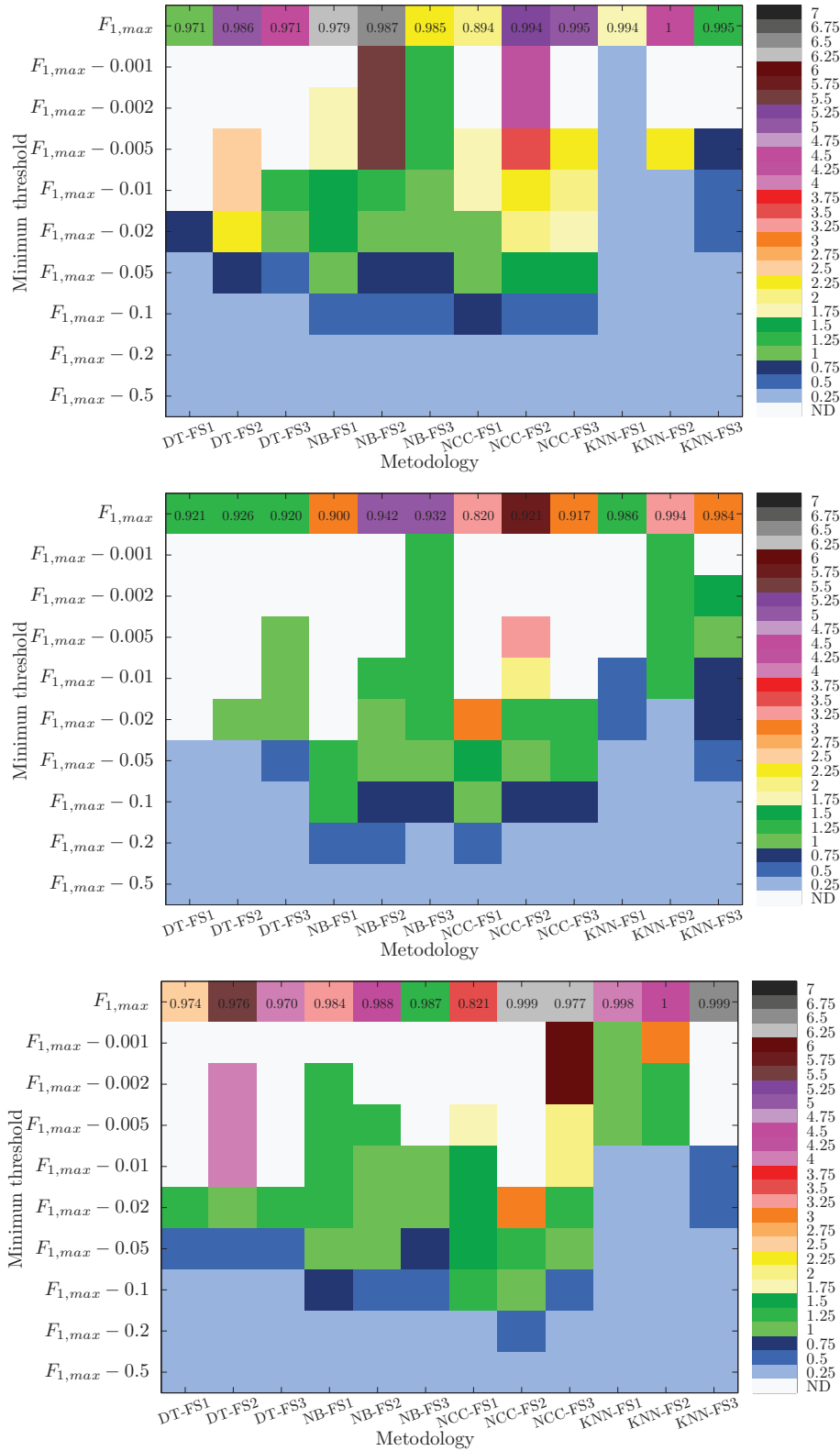


Figura 3.12: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *alcanzar un pie con la mano contraria y la cintura doblada*, *alcanzar los talones flexionando las rodillas* y *doblar lateralmente* en orden descendente

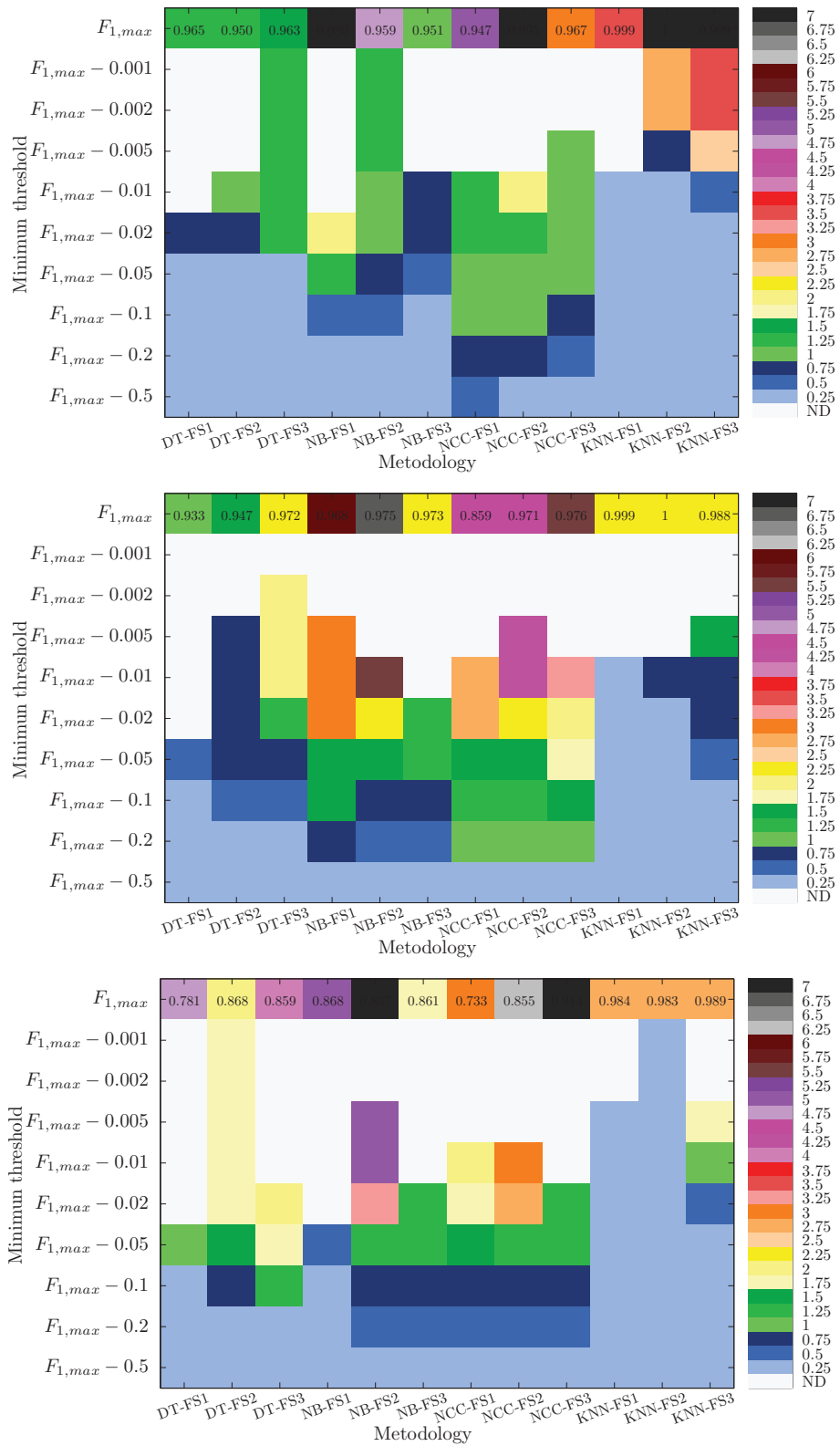


Figura 3.13: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *doblarse lateralmente con un brazo arriba*, *doblar la cintura hacia delante repetidamente* y *giro de extremidades superiores e inferiores en oposición* en orden descendente

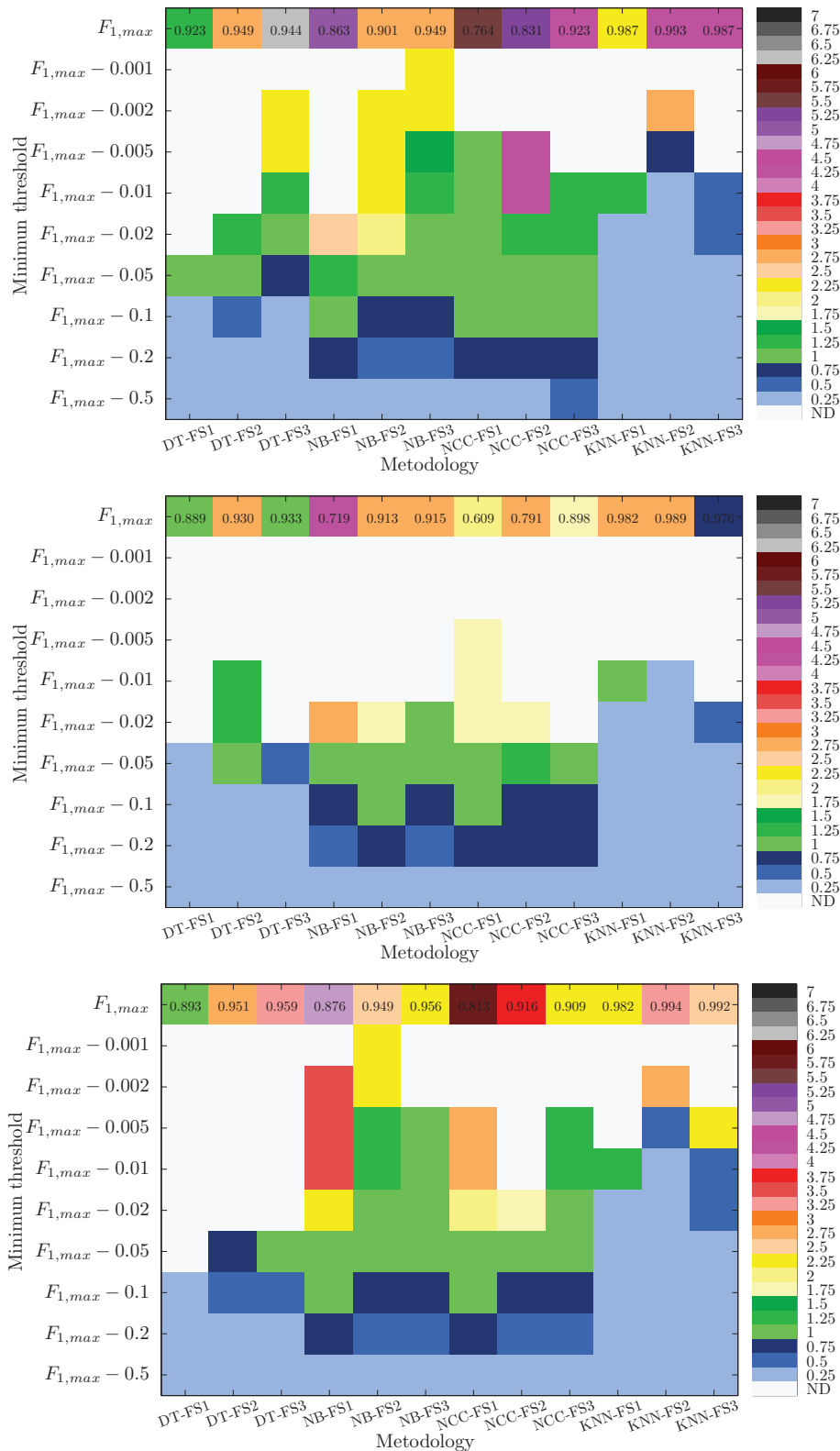


Figura 3.14: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *elevación lateral de brazos*, *elevación frontal de brazos* y *palmas frontales* en orden descendente

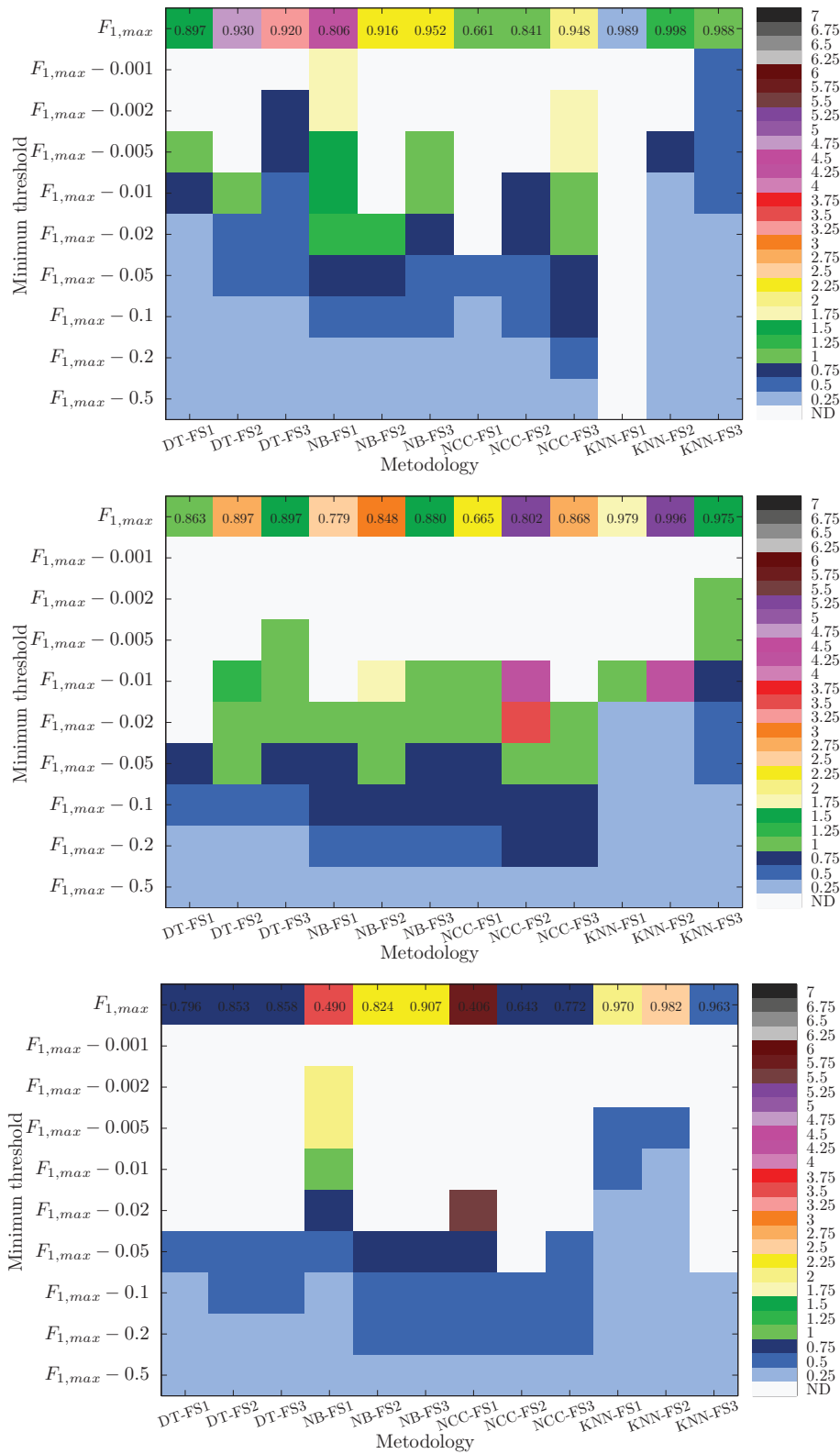


Figura 3.15: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *cruciar brazos frontalmente*, *rotación intensa de hombros* y *rotación moderada de hombros* en orden descendente

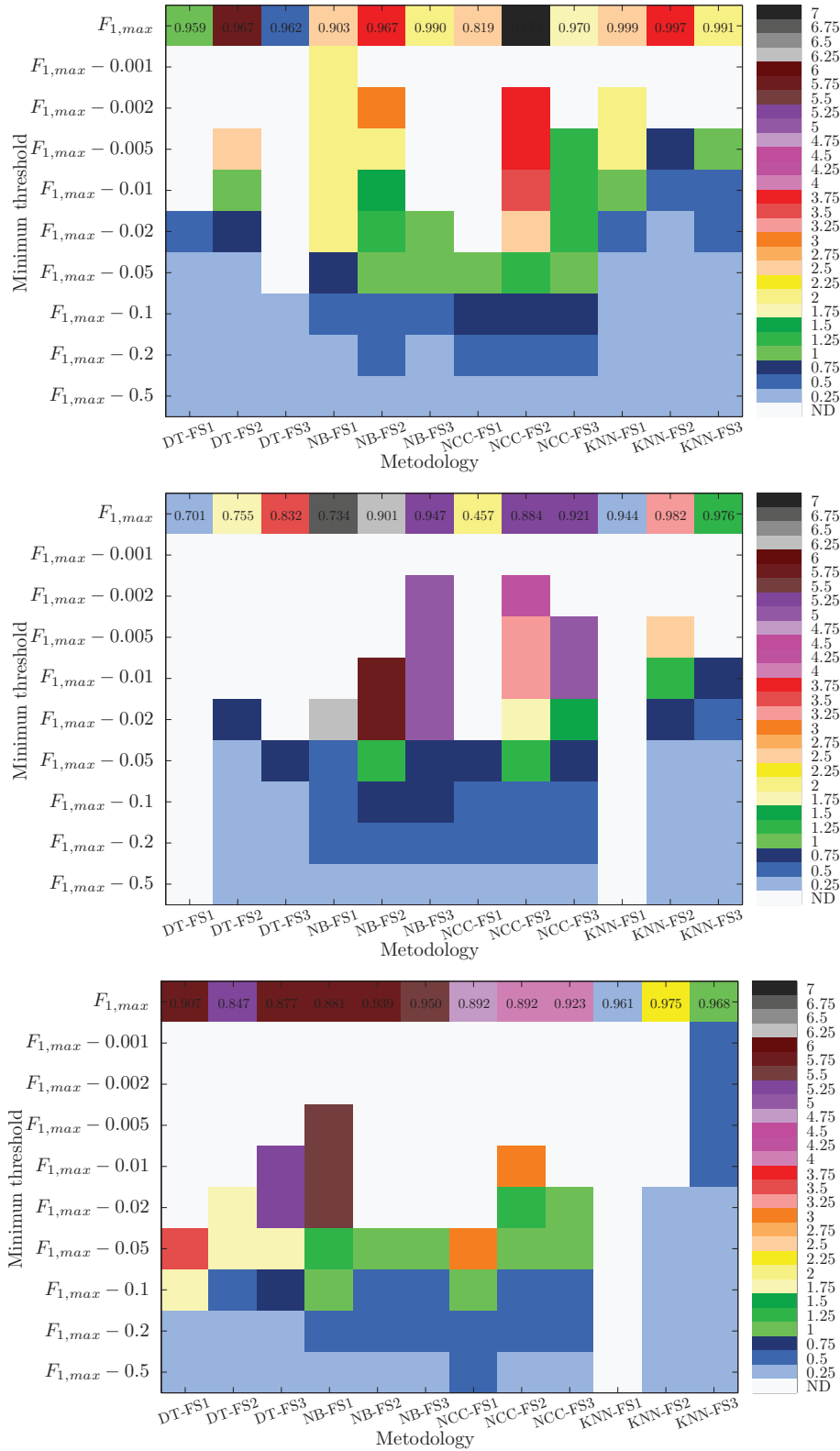


Figura 3.16: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *rotación interna de brazos*, *rodillas al pecho (alternativamente)* y *talones a la espalda (alternativamente)* en orden descendente

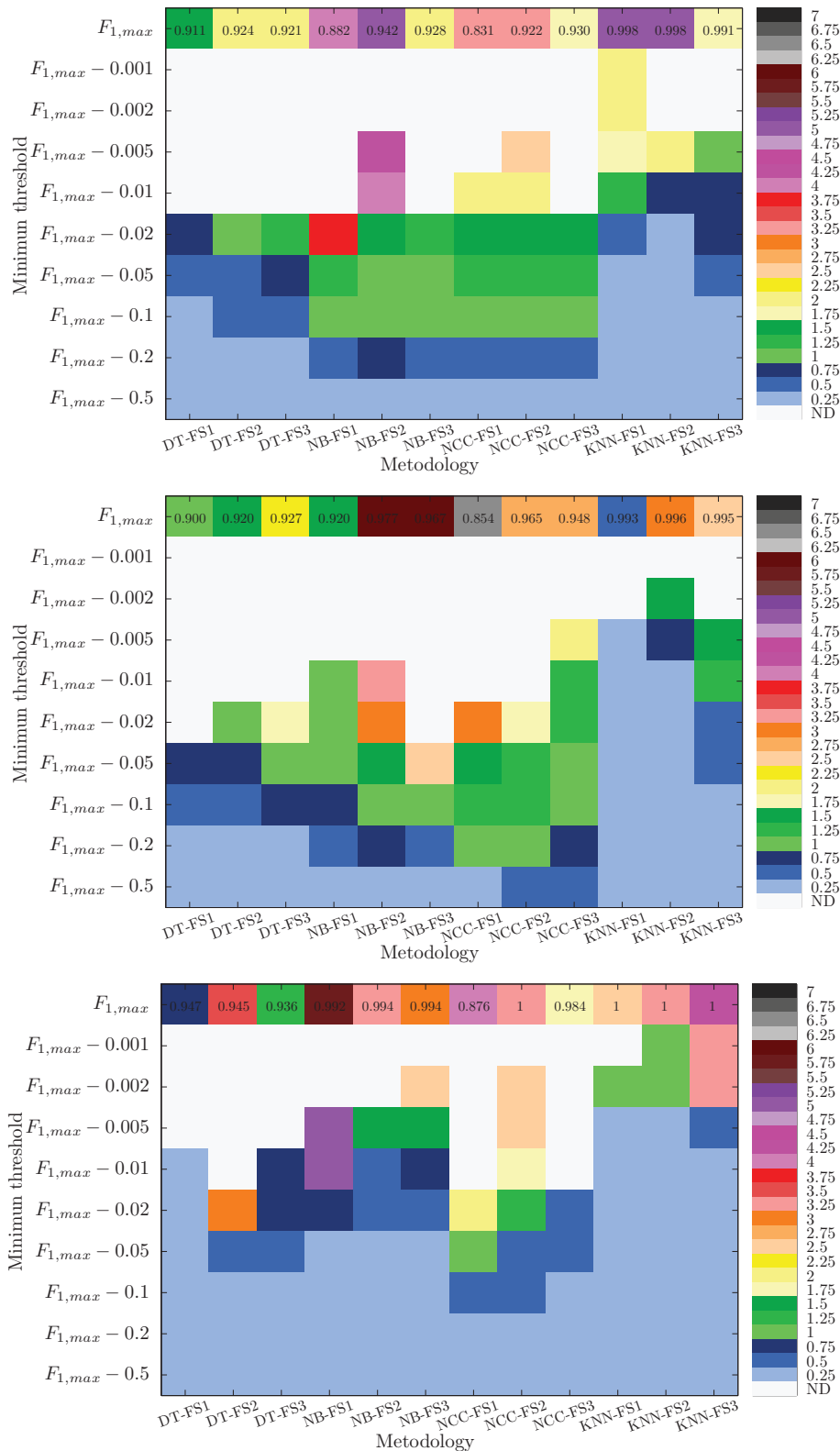


Figura 3.17: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *doblar las rodillas agachándose*, *doblar las rodillas hacia delante*(alternativamente) y *rotación de rodillas* en orden descendente

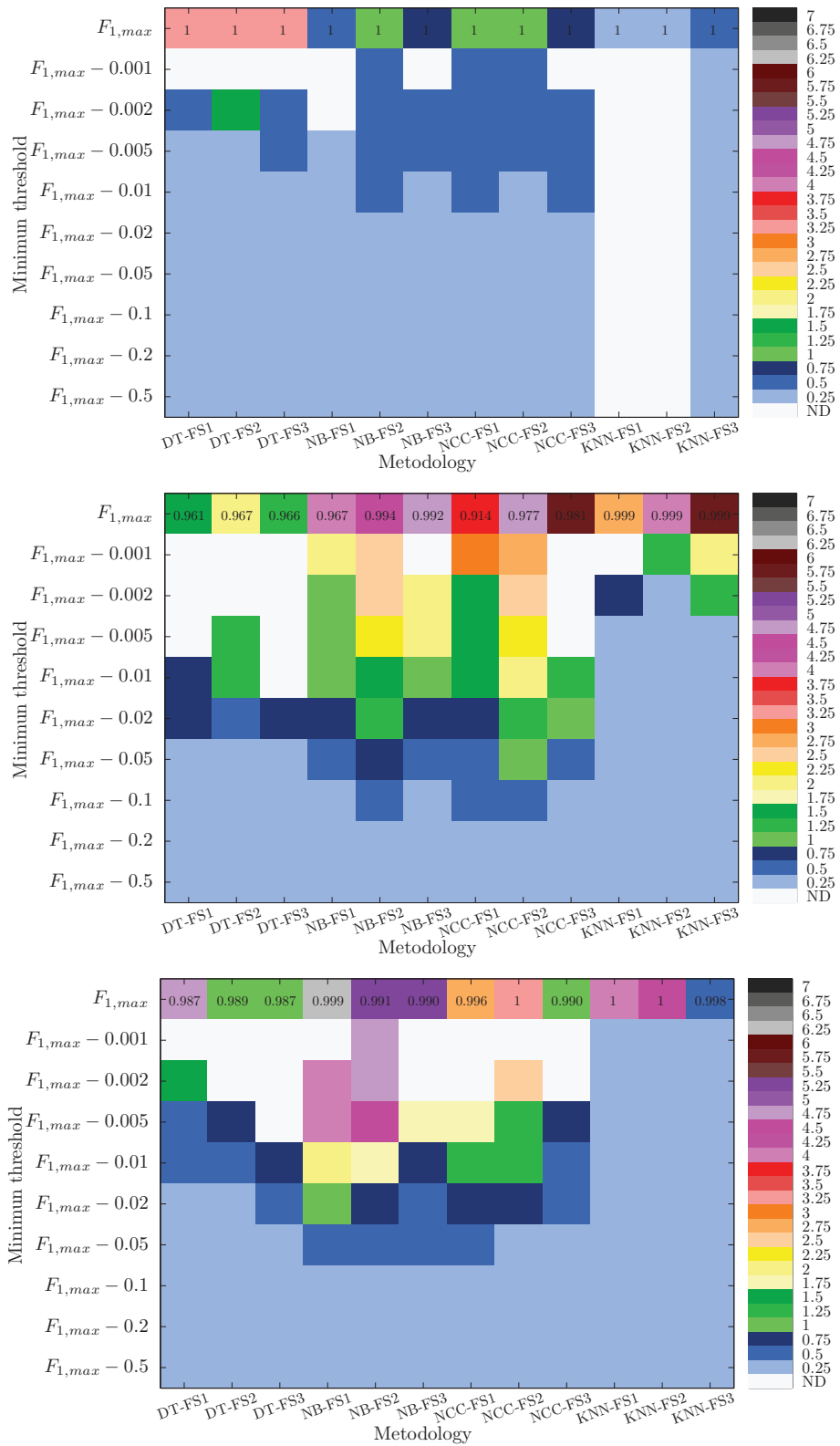


Figura 3.18: Tamaño mínimo de ventana para diversas tolerancias de F_1 - score de una actividad sobre las 12 metodologías planteadas. En filas, umbral mínimo de F_1 - score exigido. En columnas, metodologías aplicadas en este estudio (algoritmo de clasificación - conjunto de características). Las actividades presentadas son, de arriba a abajo: *remar*, *hacer bicicleta elíptica* y *hacer bicicleta estática* en orden descendente

La figura 3.8 recoge las actividades que implican movimientos de traslación: *andar*, *trotar* y *correr*. La actividad *andar* presenta buenos resultados en lo que respecta al reconocimiento preciso de su puesta en práctica. El conjunto de $F_{1,máx} - score \geq 0.96$ (excepto para *NCC* empleando la media de la señal exclusivamente) evidencia qué posibilidades actuales existen sobre el reconocimiento de esta actividad básica de la rutina diaria. Sin embargo, los tamaños de ventana requeridos para alcanzar dichas tasas son considerables empleando cualquier metodología de *NB* (6 ó 7 segundos), a partir de la media de la señal para *DT* (6 segundos) o utilizando hasta 2 características para *NCC* (5-5.25 segundos). Para otras metodologías se obtienen tamaños de ventana medios (3-4 segundos) o inferiores (0.75-2 segundos). Asumiendo una penalización moderada, se consigue reducir conjuntamente el tamaño de ventana para el abanico de metodologías planteadas. En concreto, una penalización de 0.02 implica un tamaño de ventana máximo de 2 segundos (en el caso de *NCC* que usa sólo la media de la señal). Si se asume una penalización de 0.05, dicha metodología requiere un tamaño de ventana de 1.25 segundos, pero el resto apenas necesitan 0.75 segundos. Por tanto, la actividad *andar* se puede reconocer perfectamente con un $F_1 \geq 0.91$ aplicando un tamaño de ventana máximo de 0.75 segundos. En 2005, Huynh y Schiele (64) indicaban que la actividad *andar* se reconoce con precisión para un tamaño de ventana de 1 segundo a partir del clasificador *NCC* utilizando los coeficientes 2 y 3 de la *Transformada de Fourier* de la señal. Aplicando características diferentes se ha obtenido un resultado similar. Las otras actividades que requieren movimientos de traslación presentan características similares. Destaca el reconocimiento de la actividad *correr* puesto que cualquier metodología presenta al menos $F_{1,máx} - score \geq 0.92$. Sin embargo, este ejercicio implica un tamaño de ventana relativamente alto para la optimización de $F_{1,máx}$ (7 segundos aplicando 2 características con *NB* o *KNN*, o 6 segundos con *NCC*). Puede asumirse una penalización moderada de 0.02 que permite reducir el ventanado colectivo de las metodologías hasta 1.5 segundos, conservando una magnitud del parámetro F_1 alta. Para la consideración de los diferentes conjuntos de características sobre los clasificadores *DT* y *KNN* se pueden asumir penalizaciones de 0.05 y 0.01 permitiendo disponer de un tamaño de ventana de 0.75 segundos como máximo (con $F_1 \geq 0.92$) y de 0.25 segundos (con $F_1 \geq 0.98$), respectivamente.

De la figura 3.9 hay que destacar la actividad *saltar hacia arriba*. Únicamente se pueden considerar resultados aceptables para el reconocimiento de esta actividad los presentados por el clasificador *KNN* cuando se aplican 2 y 5 características. El resto de algoritmos, además de requerir tamaños de ventana de larga duración para optimizar el reconocimiento (6.25-7 para *DT* y *NB*, y 4.75, 5 y 6.25 para *NCC*), presentan un $F_{1,máx} - score \in (0.51, 0.78)$. Reducir el tamaño de ventana es algo impensable. *KNN* con 2 características permite alcanzar $F_{1,máx} = 0.941$ con una ventana de 3 segundos. Quizás se podría tolerar una penalización de 0.05 para reducir

el tamaño de ventana hasta 0.5 segundos si existiesen exigencias en la monitorización para el reconocimiento óptimo en conjunto con otras actividades. La penalización más asumible es 0.01 que reduce el tamaño de ventana hasta 1 segundo y se consigue un reconocimiento aceptable con $F_1 \geq 0.93$. Sobre la actividad *saltar lateralmente* destacan 2 resultados contrarios: tanto *DT* como *KNN* optimizan su reconocimiento empleando el mínimo tamaño de ventana considerado para este estudio (0.25 segundos) solamente con la media de la señal. Sin embargo, obtienen $F_{1,máx} - score$ muy diferentes: 0.629 y 0.942, respectivamente.

En cuanto a la figura 3.10, la actividad *saltar la cuerda* presenta la peculiaridad de que a pesar de alcanzar un $F_{1,máx} - score = 0.741$ para *DT* con la media de la señal en 6.75 segundos, aplicando la mínima penalización se reduce el tamaño de ventana hasta 2 segundos. Al menos se puede aumentar el número de monitorizaciones en un factor superior a 3 en el caso de detectar dicha actividad. La actividad *girar el tronco con los brazos extendidos* presenta buenos resultados en general. *DT* optimiza su $F_{1,máx} - score \geq 0.93$ con 1.5 segundos y tolerando una penalización de 0.02 se podría reducir el tamaño de ventana a la mitad o incluso hasta el mínimo tamaño de ventana considerado. *KNN* con 2 características, que prácticamente maximiza su reconocimiento en 6.5 segundos, puede asumir una penalización mínima de 0.005 que permite reducir el tamaño de ventana hasta 0.25 segundos, un resultado magnífico.

El comportamiento presentado por la actividad *rotación de cintura* frente a las penalizaciones resulta destacable (figura 3.11). Salvo excepciones (como *DT*, *NCC* o *KNN* empleando 5 características), los tamaños de ventana requeridos por el resto de metodologías superan los 2.5 segundos y la tendencia son los 4-4.5 segundos. Dado que se consigue un $F_{1,máx} - score \geq 0.96$ en la mayoría de los casos, asumiendo una penalización de 0.05 se puede reducir el tamaño de ventana hasta 0.5 segundos para todos los clasificadores, excepto *KNN*, que con un umbral de 0.005 ya alcanza el tamaño mínimo de ventana.

En la figura 3.12 se puede destacar el criterio de asumir la mínima penalización en el reconocimiento de las actividades. Para las 3 actividades expuestas (*alcanzar un pie con la mano contraria y la cintura doblada*, *alcanzar los talones flexionando las rodillas* y *doblarse lateralmente*), se consigue apreciar su efecto inmediato en la reducción del tamaño de ventana sin influir apenas en el parámetro $F_{1,máx} - score$. Para la primera actividad, el mejor caso ocurre para *KNN* usando la media de la señal, que reduce el tamaño de ventana desde 1.75 hasta 0.25 segundos. Para la segunda actividad, el resultado más destacable se produce con *NB* empleando las 5 características, que reduce la ventana hasta 1.5 segundos partiendo de 5 segundos. La última actividad reduce el tamaño de ventana de 4 a 1 segundo a partir del clasificador *KNN* y la media de la señal. El parámetro $F_1 - score$ para cada una de ellas, casi intacto, queda en 0.993, 0.931 y 0.997, respectivamente.

La siguiente figura de este análisis (3.13) destaca por presentar una actividad que para su optimización requiere tamaños de ventana de larga duración: *doblar lateralmente con un brazo arriba*. Esta actividad presenta hasta 4 metodologías con un tamaño de ventana de 7 segundos y $F_{1,máx} - score \geq 0.98$. Una penalización de 0.05 los sitúa en un tamaño de ventana máximo de 1.25 en el peor de los casos (*NB* con la media de la señal) y de 0.25 segundos en el mejor (*KNN* con 2 y 5 características).

Sobre la figura 3.15 destaca la distribución de las zonas de penalización con tonos azules y verdes principalmente. Este hecho implica que las actividades alcanzan tamaños de ventana inferiores a 1.5 segundos asumiendo una serie de penalizaciones asumibles. Se trata de las actividades *cruzar brazos frontalmente*, *rotación intensa de hombros* y *rotación moderada de hombros*. A pesar de haber metodologías que no las consiguen reconocer bien (como *NCC* en alguna de las variantes consideradas), lo cierto es que otros clasificadores las consiguen monitorizar con altas tasas de rendimiento. Por ejemplo, *cruzar brazos frontalmente* se puede reconocer con *KNN* junto a la media de la señal en 0.25 segundos y un $F_{1,máx} - score = 0.989$. Otras metodologías pueden asumir penalizaciones de 0.02-0.05 para reducir el tamaño de ventana hasta 1.25 segundos como máximo.

La actividad *talones a la espalda (alternativamente)* destaca en la figura 3.16. Clasificadores como *DT* o *NB*, que necesitan tamaños de ventana superiores a 5.25 segundos, o *NCC*, que requiere al menos 4 segundos, no alcanzan o igualan siquiera en $F_{1,máx} - score$ a *KNN* con la media de la señal, que presenta un $F_{1,máx} - score = 0.961$ considerando el mínimo tamaño de ventana considerado en este proyecto.

De la figura 3.17 puede destacarse el buen reconocimiento de la actividad *rotación de rodillas*, que maximiza el parámetro F_1 para 4 metodologías. Los requerimientos en tamaño de ventana, de 2.5 segundos al menos, pueden decrementarse asumiendo una penalización de 0.005 para 2 casos de *KNN* (1 y 2 características), 0.01 para el último caso de *KNN* y 0.05 para *NCC* con 2 características, que permite reducir la ventana hasta 0.25, 0.25 y 0.5 segundos, respectivamente.

Por último, la figura 3.18 recoge las actividades específicas de *fitness*. Los resultados sobre la actividad *remar* son los mejores de todas las actividades del dataset. Se maximiza su reconocimiento para las 12 metodologías con tamaños de ventana bajos (0.25-1 segundo), salvo para *DT*, que en sus 3 variantes indica que necesita 3.25 segundos. Si se impusiera un tamaño de ventana de 0.25 segundos para todas las metodologías, supondría una penalización máxima de 0.02, lo cual no impide que la actividad se reconozca con todas las garantías. Se consigue optimizar tanto el $F_1 - score$ (~ 0.98 tras la penalización) como el tamaño de ventana (reducción hasta 0.25 segundos), por lo que su aplicación funcional es inmejorable.

Después de un análisis extenso y particularizado que permite comprobar el efecto del tamaño de ventana sobre las diferentes actividades del dataset, alguien se puede preguntar: ¿qué tamaño de ventana se podría emplear en la segmentación para penalizar lo mínimo posible el reconocimiento global de todas las actividades? Si hubiera que seleccionar un tamaño de ventana específico para la etapa de segmentación, la mejor solución es comprobar qué $F_1 - score$ se obtiene como mínimo para la actividad que peor se reconozca. La tabla 3.6 recoge dicho estudio para las 12 metodologías planteadas en este proyecto y los 28 tamaños de ventana seleccionados en el rango 0-7 segundos.

Como se puede comprobar, los clasificadores *DT*, *NB* y *NCC* presentan peores resultados que *KNN*. El mejor resultado obtenido entre los 3 primeros clasificadores se consigue para el clasificador *NB* considerando 5 características con un $F_1 - score$ garantizado para cualquier actividad del dataset de 0.748 aplicando un tamaño de ventana de 4.25 segundos. En cuanto al clasificador *KNN*, el mejor resultado se consigue aplicando un tamaño de ventana de 2 segundos a partir de la media y la desviación estándar de la señal, garantizando un $F_1 - score \geq 0.941$ para cualquier actividad del dataset. También presenta un buen resultado la variante de 5 características, que reduce el tamaño de ventana hasta 1.5 segundos garantizando un $F_1 - score \geq 0.900$. Menos destacable parece el resultado de la metodología que emplea la media de la señal, ya que el $F_1 - score \geq 0.799$, aunque se obtiene un tamaño de ventana de 0.5 segundos.

Como decisión personal, si tuviera que elegir un algoritmo de clasificación y un conjunto de características, la elección sería *KNN* utilizando la media y la desviación estándar de la señal. Puesto que además podría seleccionar el tamaño de ventana para segmentar los datos monitorizados, emplearía un tamaño de ventana de 1 segundo puesto que me garantiza un $F_1 - score \geq 0.933$ para cualquier actividad y dicho valor está muy próximo al valor óptimo alcanzado para 2 segundos (0.941).

	DT			NB			NCC			KNN		
0.25	0.557	0.587	0.584	0.283	0.382	0.434	0.052	0.237	0.286	<i>0.756</i>	0.814	0.778
0.5	0.458	<i>0.631</i>	<i>0.652</i>	0.396	0.645	0.627	0.191	0.492	0.546	0.799	<i>0.898</i>	<i>0.870</i>
0.75	0.463	0.592	0.614	0.361	0.551	0.626	0.125	0.522	0.558	0.692	0.854	0.832
1	0.488	0.646	0.676	0.224	0.615	0.667	0.183	<i>0.568</i>	0.580	0.795	0.933	0.881
1.25	0.457	0.595	0.616	0.244	0.582	0.642	0.153	0.543	0.577	0.694	0.882	0.853
1.5	0.500	0.626	0.677	0.309	0.633	<i>0.697</i>	0.148	0.564	0.599	0.801	0.920	0.900
1.75	0.489	0.613	0.650	0.354	0.604	0.673	0.140	0.575	0.587	0.721	0.881	0.872
2	0.380	0.630	0.692	0.173	0.619	0.692	0.143	0.578	0.586	0.805	0.941	0.899
2.25	0.434	0.554	0.595	0.299	0.606	0.649	0.149	0.571	0.579	0.735	0.911	0.845
2.5	0.491	0.639	0.653	0.279	0.625	0.701	0.138	0.585	0.598	0.720	0.905	0.857
2.75	0.414	0.623	0.620	0.280	0.595	0.635	0.162	0.547	0.574	0.696	0.896	0.809
3	0.485	0.655	0.601	<i>0.418</i>	0.651	0.680	0.148	0.563	0.576	0.657	0.941	0.862
3.25	0.389	0.638	0.607	0.322	0.654	0.679	0.149	0.588	<i>0.619</i>	0.700	0.903	0.837
3.5	0.372	0.579	0.643	0.237	0.632	0.704	0.130	0.602	0.550	0.591	0.892	0.802
3.75	0.383	0.623	0.633	0.300	0.561	0.527	0.160	0.535	0.517	0.505	0.835	0.699
4	0.389	0.605	0.611	0.141	0.467	0.659	0.103	0.577	0.621	0.333	0.832	0.755
4.25	0.464	0.646	0.617	0.414	0.712	0.748	0.144	0.549	0.598	0.256	0.853	0.594
4.5	0.393	0.662	0.589	0.402	0.698	0.717	0.123	0.558	0.580	0.135	0.817	0.610
4.75	0.398	0.677	0.555	0.182	0.493	0.707	0.143	0.591	0.659	0.264	0.744	0.531
5	0.338	0.649	0.644	0.122	0.422	0.594	0.120	0.576	0.559	0.105	0.665	0.524
5.25	0.398	0.669	0.648	0.460	0.691	0.703	0.160	0.585	0.525	0.171	0.638	0.523
5.5	0.446	0.606	0.611	0.177	0.450	0.440	0.115	0.516	0.481	0.110	0.532	0.437
5.75	0.381	0.651	0.694	0.064	0.401	0.551	0.139	0.593	0.585	0.146	0.564	0.518
6	0.396	0.672	0.640	0.405	0.625	0.665	0.088	0.552	0.527	0.076	0.492	0.607
6.25	0.420	0.605	0.606	0.301	0.650	0.671	0.144	0.555	0.536	0.114	0.458	0.624
6.5	0.227	0.681	0.657	0.295	0.559	0.634	0.111	0.535	0.515	0.204	0.514	0.603
6.75	0.381	0.640	0.647	0.417	0.699	0.709	0.173	0.532	0.523	0.141	0.486	0.602
7	0.473	0.566	0.688	0.399	0.631	0.724	0.117	0.560	0.565	0.153	0.524	0.507

Tabla 3.6: F_1 – score mínimo obtenido para cualquier actividad del dataset para las 12 metodologías planteadas aplicando un tamaño de ventana específico. En cada fila se incluyen los resultados para un tamaño de ventana del rango de estudio (0-7 segundos). En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el F_1 – score mínimo que permite seleccionar el menor tamaño de ventana posible optimizando el F_1 – score. En cursiva, el F_1 – score mínimo que permite seleccionar el menor tamaño de ventana posible considerando una penalización máxima ~ 0.05 al F_1 – score optimizado

3.5. Conclusiones

A lo largo de este capítulo se ha realizado un estudio sobre el efecto del tamaño de ventana en el reconocimiento de la actividad. Su influencia, presentada en el *Estado del arte* de este proyecto 1.4, ha sido confirmada con la aplicación, análisis y estudio sobre un amplio dataset de actividades, obteniendo diferentes resultados bajo diversos criterios.

La eficiencia de clasificación global de cada metodología ofrece una idea sobre el tamaño de ventana óptimo aplicable. Las diferentes tendencias y comportamientos de las metodologías frente al tamaño de ventana considerado invitan a un estudio personalizado del efecto de dicho parámetro sobre cada actividad particular. La búsqueda de una métrica apropiada para abordar el análisis exhaustivo y pormenorizado de las actividades implicaba la consideración de parámetros que recojan el impacto de los reconocimientos fallidos y acertados de los clasificadores evaluados.

La utilización de la métrica F_1 – score es estricta en las condiciones buscadas y penaliza proporcionalmente el correcto reconocimiento y la confusión frente a otras actividades. Un análisis centrado en el efecto del tamaño de ventana sobre cada actividad ha llevado a la aplicación de estudios diferentes según el criterio establecido. Intentar observar qué tamaño de ventana mínimo se podría aplicar estableciendo una exigencia en la magnitud F_1 ha supuesto la obtención de las figuras 3.2 y 3.3.

Los análisis por metodología o por actividad se convierten en un punto de referencia para comparar y evaluar las prestaciones que ofrece un clasificador específico aplicando diferentes conjuntos de características seleccionados sobre actividades que, por lo general, forman parte de la rutina diaria. Su aplicación exclusiva en disciplinas como rehabilitación, teleasistencia o control deportivo que deseen monitorizar actividades tan comunes como *andar*, *correr*, *hacer bicicleta elíptica*, *hacer bicicleta estática*, *remar*, etc., trae consigo grandes ventajas en cuanto a la reducción del tiempo en la toma de decisión asumiendo una pérdida de reconocimiento controlada.

Un reconocimiento prácticamente óptimo de cualquier actividad considerada en este dataset se puede conseguir con la aplicación del clasificador *KNN* que permite la reducción del tamaño de ventana hasta 0.25 segundos (salvo algunas actividades que impliquen saltos como *saltar hacia arriba* o *saltar adelante y atrás*). Adicionalmente, se ha optimizado el reconocimiento conjunto de las actividades garantizando un F_1 – score ≥ 0.941 para cada actividad utilizando la media y la desviación estándar de la señal con el clasificador de los K-Vecinos más cercanos.

Estos resultados se han convertido en todo un éxito a tener en cuenta en futuros trabajos y que anima a realizar estudios adicionales en la continuación de este proyecto.

4

Capítulo IV: Fusión multiventana

4.1. Introducción

Tras la finalización del capítulo 3, en el que se ha abordado, analizado y comprobado la influencia del tamaño de ventana en el reconocimiento de actividades, la principal motivación de este proyecto ha quedado cubierta.

Sin embargo, teniendo en cuenta los resultados obtenidos y el margen de mejora posible en relación a la temática tratada en este proyecto, se ha considerado oportuno implementar alguna estrategia que emplee todo el conocimiento adquirido en torno al tratamiento con diferentes tamaños de ventana durante el reconocimiento de actividades.

La aplicación de distintos tamaños de ventana en tiempo real a los datos adquiridos permite disponer de diferentes decisiones durante el reconocimiento de actividades. Ha quedado de manifiesto en el capítulo 2 que cada actividad alcanza un reconocimiento óptimo (y en algunos casos máximo) para un tamaño de ventana diferente, dependiente además del clasificador empleado y las características seleccionadas para su extracción de las señales inerciales.

En este punto, se plantea un algoritmo de fusión multiventana que permita ofrecer una decisión más robusta sobre la actividad reconocida durante la monitorización. La estrategia de fusión ya ha sido considerada en trabajos anteriores sobre el *ARC*. Aplicando dicha estrategia a nivel de clases (en nuestro caso, actividades), a nivel de fuentes (tal y como se presenta *ARC*, sensores) y a nivel de clasificadores, este año Baños et al. (71) ha definido una etapa de clasificación que se compone de 3 niveles que definen el clasificador *Hierarchical-Weighted Classification* (HWC) a partir de los clasificadores de actividades entrenados (c_{mn}), los pesos a nivel de actividad (β_{mn}) y los pesos a nivel de sensores (α_{mn}), donde m hace referencia al sensor considerado y n a la actividad clasificada (ver figura 4.1). Por otro lado, el motivo más importante para utilizar una estrategia de fusión reside en incrementar la eficiencia del sistema

de reconocimiento. En 2007, Zappi et al. (85) demuestran que la fusión a nivel de clasificación reduce la complejidad en las primeras etapas de *ARC*, lo que supone un incremento en términos de robustez y adaptabilidad para el sistema de reconocimiento.

Aplicado a nuestro estudio, se busca la implementación de un algoritmo de fusión de ventanas que permita a nivel de clasificación aportar mayor robustez al sistema *ARC* en la toma de decisión final sobre la actividad reconocida en un tiempo de referencia establecido. Además, se adopta la idea de pesado de la clasificación obtenida para cada tamaño de ventana considerado.

La descripción del algoritmo implementado se presenta en la sección 4.2. Diferentes detalles sobre el algoritmo aplicados al estudio particularizado en este proyecto son descritos en la sección 4.3. Posteriormente, se presentan los resultados obtenidos, así como las interpretaciones que se derivan de la aplicación del algoritmo en la sección 4.4. Finalmente, el capítulo termina con la inclusión de las conclusiones a las que se ha llegado durante la realización de este capítulo.

4.2. Descripción del algoritmo

Tradicionalmente, en lo que se refiere al reconocimiento de actividades, el cauce de los datos monitorizados ha seguido un curso lineal a través de las diferentes etapas del *ARC*. En cada etapa, tal y como se expuso en 2.5.2, se seleccionan una serie de parámetros que definen unívocamente el diseño del sistema *ARC*. La utilización de diferentes tamaños de ventana en un mismo sistema de monitorización en tiempo real se contempla como una *adaptación* del cauce original, permitiendo la aplicación de distintas segmentaciones a los datos procesados. Las ventajas adquiridas son evidentes vistos los resultados alcanzados con el estudio realizado en el capítulo 3, debido a que las actividades optimizan su reconocimiento para distintos tamaños de ventana. La fusión multiventana consiste en la consideración conjunta de decisiones tomadas en la etapa de clasificación debido a la aplicación durante la etapa de segmentación de diferentes tamaños de ventana. De este modo, en lugar de aceptar como válida la decisión única tomada por

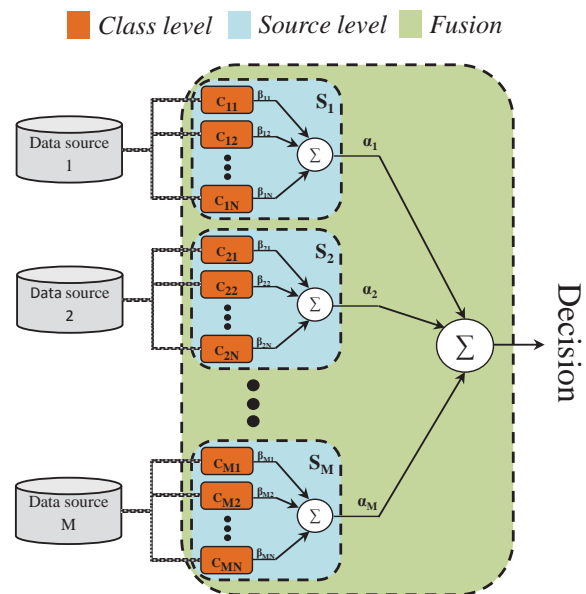


Figura 4.1: Estructura del clasificador HWC. Problema con N actividades y M sensores

el clasificador al reconocer una actividad cada tamaño de ventana, se trata de que el clasificador disponga de más decisiones en un tiempo de referencia para ofrecer una decisión final que tenga en cuenta las decisiones parciales tomadas para cada tamaño de ventana considerado.

El algoritmo de fusión multiventana depende de diversos factores:

- ★ **Tamaños de ventanas:** Se seleccionan un conjunto de K ventanas de duración inferior al *instante de decisión*. La adopción y aplicación de diferentes tamaños de ventana permite la obtención de diferentes decisiones previas al instante de decisión. Por simplicidad, se han considerado tamaños de ventana múltiplos del instante de decisión.
- ★ **Instante de decisión:** Se trata del tiempo empleado tradicionalmente por el tamaño de ventana aplicado en *ARC*. Su determinación para la aplicación de la fusión es fundamental, ya que se trata del instante temporal en que se aporta la decisión final sobre la actividad reconocida por las diferentes ventanas consideradas. Dicho instante temporal suele coincidir con la longitud de la ventana de mayor tamaño N .
- ★ **Etiquetas de clasificación:** Cada tamaño de ventana requiere una clasificación individual, por lo que el algoritmo considera las diferentes actividades reconocidas para los tamaños de ventana seleccionados.
- ★ **Métrica de pesado:** Se considera alguna métrica que pondere las clasificaciones parciales realizadas a partir de la consideración de los diferentes tamaños de ventana.

En la figura 4.2 se presenta un esquema *ARC* adaptado a las contribuciones del proyecto:

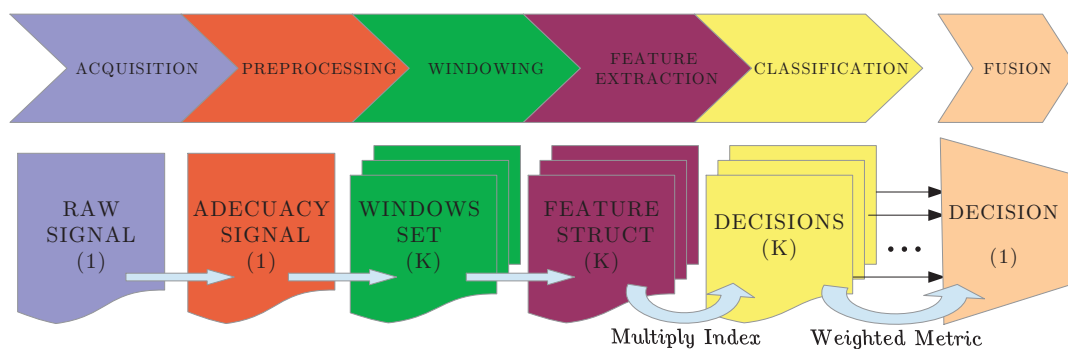


Figura 4.2: Esquema *ARC* adaptado al multiventanado. Se realiza una adecuación de la señal adquirida para conseguir una partición correcta en la etapa de segmentación de las K ventanas seleccionadas (múltiplos de la ventana de mayor tamaño, de tamaño N) y se multiplican índices de control para cada tamaño de ventana seleccionado. Se ponderan las K decisiones con una métrica seleccionada y se realiza la fusión de decisiones tras la etapa de clasificación, obteniendo una decisión única que es aceptada como actividad reconocida

Una vez presentados los factores determinantes del algoritmo, se realiza una descripción paso por paso sobre su funcionamiento:

Paso 1 Importación de parámetros: la métrica, los tamaños de ventana y las etiquetas de actividades clasificadas

Paso 2 Creación de un vector de decisiones finales: su tamaño coincide con la ventana de mayor tamaño (tamaño N)

Paso 3 Creación de un vector de índices de control para los tamaños de ventana

Paso 4 Cálculo de la relación existente entre el tamaño de la ventana de mayor tamaño y el resto

Paso 5 Aplicación del proceso de fusión

(a) Creación de un vector de decisiones parciales: su tamaño coincide con el total de tamaños de ventana considerados

(b) Selección de un tamaño de ventana: este bucle se repite para todos los tamaños de ventana considerados

★ Evaluación de datos del tamaño de ventana seleccionado

★ Elaboración de una tabla de contabilidad de las actividades predichas

★ Contabilidad de repeticiones de cada actividad

★ Ponderación de cada actividad: pesado a partir de la utilización de la métrica

★ Adopción de la decisión parcial de la actividad reconocida para el tamaño de ventana seleccionado (*nivel 1 de decisión*)

★ Actualización del índice de control asociado al tamaño de ventana seleccionado

(c) Creación de una tabla de ponderación de las clases predichas para la fusión: su tamaño coincide con el número de actividades consideradas en el estudio y el número de ventanas seleccionadas

(d) Ponderación de cada clase predicha: pesado a partir de la utilización de la métrica

(e) Contabilidad de repeticiones de cada clase predicha

(f) Adopción de la decisión final de la actividad reconocida (*nivel 2 de decisión*)

Paso 6 Repetición del **Paso 5** hasta un total de N veces

Paso 7 Obtención completa del vector de decisiones finales: fin de la fusión

Una vez presentado un esquema *ARC* adaptado al estudio particularizado de este proyecto, los parámetros del algoritmo y los pasos seguidos para su implementación, se concreta su aplicación y estudio en la sección 4.3.

4.3. Descripción del estudio. Análisis

Tal y como se ha puede ver en la descripción del algoritmo, detallada paso por paso, se adoptan decisiones de fusión a 2 niveles. En primer lugar, la decisión a nivel 1 implica que cada tipo de ventana seleccionada acuerda por ponderación qué actividad se ha reconocido tras un tiempo denominado **instante de decisión**. Este instante temporal coincide, por lo general, con la longitud de la ventana de mayor tamaño, pero se ha desarrollado un estudio en este apartado que supone una excepción de esta regla. En lo que respecta a la métrica empleada para la ponderación, se ha optado por la utilización de la métrica F_1 - *score*. Ya se han comentado las ventajas de su aplicación en el capítulo 3 sección 3.3, y se ha considerado una métrica adecuada para este estudio.

Se han considerado 2 estudios para testear el algoritmo de fusión planteado:

- ★ **Fusión a 2 niveles:** Se aplica el algoritmo de fusión considerando 5 tamaños de ventana (0.25, 0.75, 1.5, 3 y 6 segundos) y se comprueba la eficiencia de clasificación al aplicar diferentes combinaciones. En total, 19 combinaciones han sido probadas en este estudio de las 31 combinaciones posibles: 10 combinaciones incluyen el tamaño de ventana de 6 segundos, otras 5 incluyen hasta 3 segundos como la longitud de la mayor ventana, 2 combinaciones con 1.5 segundos como la ventana de mayor tamaño y 2 combinaciones más para la aplicación individualizada del ventanado a 0.25 y 0.75 segundos.
- ★ **Fusión a 1 nivel:** Se aplica una variante del algoritmo de fusión, adoptando 6 segundos como instante de decisión final y aplicando únicamente un tamaño de ventana para la obtención de la decisión en fusión. Para este estudio se han añadido el resto de tamaños de ventana considerados a lo largo de este proyecto que son múltiplos del tamaño de ventana de referencia (6 segundos). En concreto, se agregan los tamaños de ventana 0.5, 1 y 2 segundos. En total, se obtienen 8 resultados en este estudio.

Cada estudio requiere de unas condiciones particulares. Ambos análisis han precisado la adecuación de los datos inerciales monitorizados para un correcto reconocimiento en la etapa de clasificación. La selección de tamaños de ventana múltiplos de 6 segundos ha sido una decisión que, más allá de la simplicidad y comodidad que permite para la puesta en práctica del algoritmo de fusión, se encuentra respaldada por los resultados obtenidos en trabajos previos

(Bao e Intille (1) (2004) y Ravi et al. (63) (2005)). Sin embargo, la principal razón radica en la eliminación de fallos de reconocimiento por procesamiento inadecuado del particionamiento de datos. La adecuación de los datos permite un reconocimiento exacto de la información disponible y evita problemas en la etapa de extracción de características por la disposición de ventanas incompletas.

En cuanto al abanico de ventanas seleccionadas para los estudios, diferentes criterios han marcado la elección. Por un lado, se ha decidido tener en cuenta el tamaño de ventana mínimo aplicado en este proyecto (0.25 segundos) por la obtención de buenos resultados en el capítulo 3 sobre el reconocimiento de gran parte de las actividades monitorizadas. La ventana de 0.75 segundos se convierte en una ventana intermedia entre los resultados optimizados para actividades en 0.5 y 1 segundo, cuyos tamaños han sido testeados y considerados válidos para el reconocimiento de actividades cotidianas como *andar* o *trotar* en otros trabajos (Huynh y Schiele (64) (2005) y Maurer et al. (3) (2006)) y que forman parte del dataset estudiado en este proyecto. La selección de los tamaños de ventana de 1.5 y 3 segundos intentan recoger el efecto del tamaño de ventana de diferentes rangos de monitorización como soluciones intermedias hasta la aplicación del tamaño de referencia: 6 segundos. Por otra parte, la aplicación de combinaciones que prescinden del uso de la ventana de 6 segundos permite observar la eficiencia de clasificación en un menor tiempo y el efecto de la incorporación de la fusión de decisiones.

El segundo estudio considera todos los múltiplos del tamaño de ventana de 6 segundos aplicados experimentalmente, con el fin de ampliar y facilitar la interpretación posterior de los resultados. La principal motivación para su puesta en práctica radica en los resultados sobre la eficiencia de clasificación obtenidos en la variación del tamaño de ventana. La fusión de diferentes decisiones previas a 6 segundos muestra, dependiendo de la eficiencia de clasificación obtenida para cada tamaño de ventana aplicado, qué potencial ofrece su aplicación para la mejora del reconocimiento final. Con la aplicación de una ventana de 0.25 segundos se dispone de 24 decisiones individuales para la decisión final en 6 segundos. La fusión de 24 decisiones en las que la eficiencia de clasificación presenta un porcentaje considerablemente alto (por ejemplo, superior al 85% con el clasificador *DT* o del 96% con *KNN*) invita al optimismo sobre la obtención de un incremento en la eficiencia del reconocimiento. Por el contrario, la aplicación de 24 decisiones con eficiencias de clasificación discretas o bajas como se obtiene para el clasificador *NCC* (55% utilizando sólo la media de la señal) suponen una incógnita, así como un reto y una motivación que invita a su aplicación. De la misma forma ocurre con otros tamaños de ventana como 0.75 segundos (del que se disponen 8 decisiones), 1.5 segundos (4 decisiones) y 3 segundos (2 decisiones). La aplicación de la ventana de 6 segundos implica la obtención de 1

decisión en el instante de referencia, de manera que sus resultados ya se encuentran disponibles previo al inicio del estudio.

La figura 4.3 muestra un ejemplo del funcionamiento de los algoritmos de fusión a 1 nivel y 2 niveles. A su vez, se contrasta la decisión final con la que se obtendría para el mayor tamaño de ventana (en el ejemplo, W5) y que podría corresponder, tal y como se viene desarrollando durante en este proyecto, al tamaño de ventana de 6 segundos.

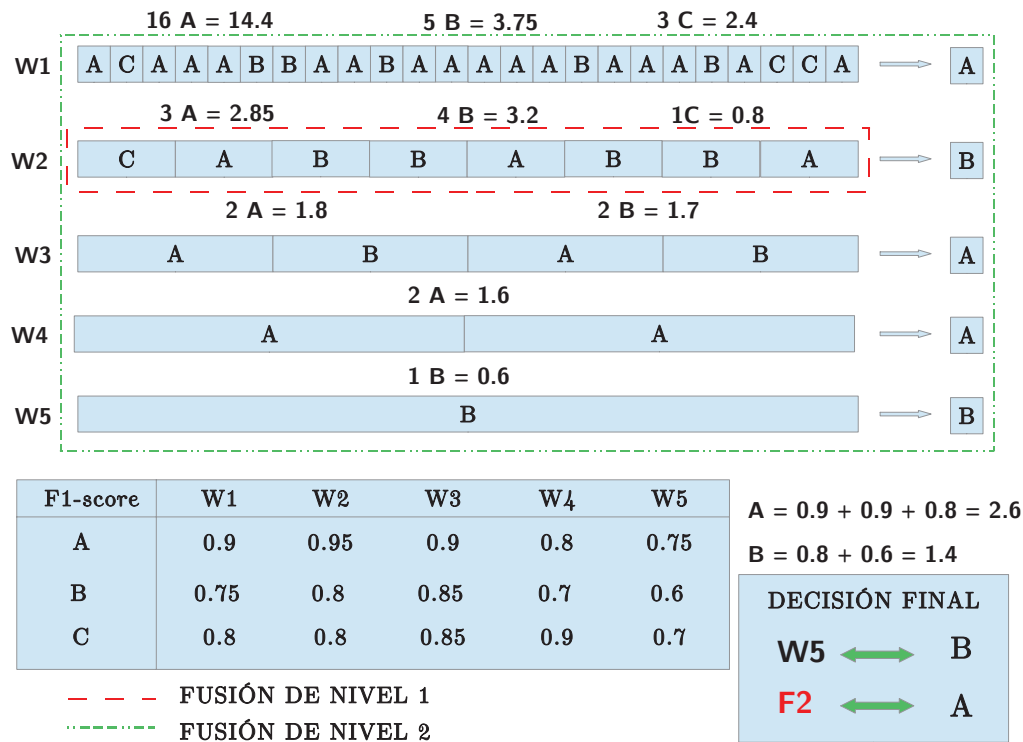


Figura 4.3: Ejemplo de funcionamiento del algoritmo de fusión a 1 nivel y 2 niveles. Se muestran las decisiones parciales adoptadas por $K = 5$ ventanas en la parte superior para las actividades reconocidas A, B y C. En la parte inferior se muestra una tabla de $F_1 - score$ correspondiente al reconocimiento de cada actividad para un tamaño de ventana específico W1-W5. La decisión final de la fusión a 2 niveles muestra la clasificación de la actividad como A, mientras que el ventanado tradicional reconocería la actividad como B. Este ejemplo está adaptado para la utilización de fusión completa expuesta en este proyecto: fusión simultánea de las ventanas de 0.25, 0.75, 1.5, 3 y 6 segundos

En la sección 4.4 se presentan tanto tablas como figuras que recogen el efecto de la fusión de ventanas en la eficiencia de clasificación. Todos los resultados y los detalles observados de estos análisis son discutidos con un enfoque crítico sobre la potencial aplicación de la fusión multiventana.

4.4. Resultados y discusión

Una vez testado el algoritmo de fusión bajo los 2 estudios considerados (19 combinaciones de los 5 tamaños de ventana seleccionados para el primer estudio y 8 resultados para los 8 tamaños de ventana escogidos para el segundo estudio), se presenta la tabla 4.1 que muestra las combinaciones de ventanas seleccionadas para el testeo del algoritmo de fusión a 2 niveles. Las tablas 4.2 y 4.3 recogen los resultados de la aplicación del algoritmo de fusión a 2 niveles tomando como mayor tamaño de ventana 6 y 3 segundos, respectivamente. Adicionalmente, se muestra la tabla 4.4 que recoge los resultados de la aplicación del algoritmo de fusión a 2 niveles para otras combinaciones con tamaños de ventana inferiores a 3 segundos. Por otra parte, la tabla 4.6 resume los resultados obtenidos tras la consideración de algoritmo de fusión en su versión de 1 nivel, aplicando únicamente un tamaño de ventana durante un tiempo total de 6 segundos para dar la decisión final. Para su comparación, se muestra la tabla 4.5 que recoge los resultados obtenidos en el capítulo 3 únicamente para los tamaños de ventana implicados en el estudio de fusión a 1 nivel.

	Con 6 segundos										Con 3 segundos					Otras			
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
0.25									X	X					X				X
0.75					X	X	X	X		X			X	X			X	X	
1.5			X	X			X	X		X		X		X		X	X		
3		X		X		X		X		X	X	X	X	X					
6	X	X	X	X	X	X	X	X	X	X									

Tabla 4.1: Tabla de combinaciones de ventanas aplicadas para el testeo del algoritmo de fusión multiventana a 2 niveles

Además de las combinaciones correspondientes a la obtención de una decisión final en 3 y 6 segundos, la tabla 4.1 incluye otras correspondientes a la toma de decisión final en 1.5 segundos (2 combinaciones), y 0.75 y 0.25 segundos (1 combinación para cada una). Especial interés se ha prestado a la obtención de resultados referentes a la toma de decisión final en 6 segundos con el fin de comparar los resultados de fusión (combinaciones C2-C10) con los obtenidos sin aplicar fusión (combinación C1), puesto que es un tiempo considerado como suficiente para reconocer cualquier actividad de un conjunto amplio de actividades (Bao e Intille (1) (2004) y Ravi et al. (63) (2005)). La combinación C10 aplica los 5 tamaños de ventana seleccionados y se le ha denominado en este proyecto como *fusión completa*. De manera complementaria, se han considerado combinaciones que presenten la decisión final en un tiempo menor para comprobar el efecto de fusión (combinaciones C11-C19). Sobre la distribución y selección de combinaciones,

se ha decidido mostrar la evolución de la eficiencia de clasificación colocando las combinaciones en orden binario creciente, así como los tamaños de ventana implicados en orden descendente (por ejemplo, C3 es equivalente a 00101, o lo que es lo mismo, combinar las decisiones adoptadas para los tamaños de ventana de 1.5 y 6 segundos con la decisión final en 6 segundos).

	DT			NB			NCC			KNN		
C1	85.09	89.54	89.25	87.82	93.38	92.50	79.57	89.75	90.09	93.59	<i>96.41</i>	94.71
C2	87.47	91.11	89.80	89.81	95.05	95.82	82.18	91.40	94.66	95.64	98.09	<i>98.24</i>
C3	90.83	94.14	93.73	89.97	95.16	95.91	82.44	91.51	94.44	95.81	98.02	<i>98.25</i>
C4	92.63	95.83	96.00	90.06	95.38	95.98	82.42	91.43	94.53	96.35	98.39	<i>98.59</i>
C5	92.31	95.31	94.22	89.71	95.03	95.60	81.01	90.99	94.30	96.03	98.22	<i>98.46</i>
C6	93.16	96.03	96.62	90.38	95.31	96.16	82.24	91.37	94.43	96.42	98.57	<i>98.60</i>
C7	94.74	97.36	97.58	89.73	95.16	95.96	82.11	91.29	94.48	96.68	98.57	<i>98.73</i>
C8	94.80	97.32	97.66	90.22	95.22	95.94	82.27	91.52	94.52	96.58	98.47	<i>98.68</i>
C9	93.77	94.61	92.86	88.84	95.00	95.61	80.43	90.80	93.40	97.35	<i>98.59</i>	98.26
C10	96.74	98.25	98.63	90.28	95.22	95.98	82.01	91.43	94.50	97.21	98.70	98.77

Tabla 4.2: Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento cada 6 segundos. Las ventanas seleccionadas son múltiplos de la ventana de mayor tamaño (6 segundos). En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada combinación aplicada

Observando la tabla 4.2 y la figura 4.4 se pueden extraer diversos resultados relevantes. Por un lado, el clasificador *DT* se ve claramente beneficiado de la aplicación del algoritmo de fusión implementado en este proyecto. Las 3 metodologías ven incrementada su eficiencia de clasificación en el reconocimiento de la actividad sustancialmente conforme se agregan ventanas diferentes para tomar la decisión final en 6 segundos. Hay que tener en cuenta que las combinaciones no se presentan en orden creciente de número de ventanas empleadas para la fusión. La eficiencia incrementa desde una tasa del 85.09% (que ya se obtuvo en el estudio realizado en el capítulo 3 aplicando una segmentación con ventanas de 6 segundos) hasta un 96.74% empleando *fusión completa* en el mismo tramo temporal y utilizando únicamente la media de la señal. Lo mismo ocurre para las variantes que emplean 2 y 5 características, consiguiendo la optimización de la eficiencia a partir de la aplicación de *fusión completa* que aumenta el porcentaje de reconocimiento global del clasificador hasta 98.25% y 98.63%, respectivamente.

En cuanto a los clasificadores *NB* y *NCC*, se observa una tendencia diferente a *DT*. En general, se observa una mejora con respecto a la segmentación de los datos en ventanas de 6 segundos para todos los casos, aumentando la eficiencia entre un 2-4.5%. Sin embargo, lo más

destacable es que todas las metodologías de ambos algoritmos, a pesar de experimentar una mejora global, no presentan una combinación de ventanas que destaque frente al resto. Las diferencias existentes en el abanico de combinaciones C2-C10 son mínimas, ya que apenas se alcanza una variación entre ellas de 0.5% aunque los resultados señalados en negrita sean considerados como óptimos entre las combinaciones seleccionadas para este proyecto. Por tanto, se puede concluir que la adición de ventanas no garantiza un incremento progresivo de la eficiencia cuando se emplean estos clasificadores, por lo que en el caso de aplicar fusión multiventana, se podría optar por la aplicación del menor número de ventanas posible que aumente la eficiencia de clasificación del ventanado tradicional y evitar una sobrecarga innecesaria en tiempo real en el procesamiento de los datos monitorizados.

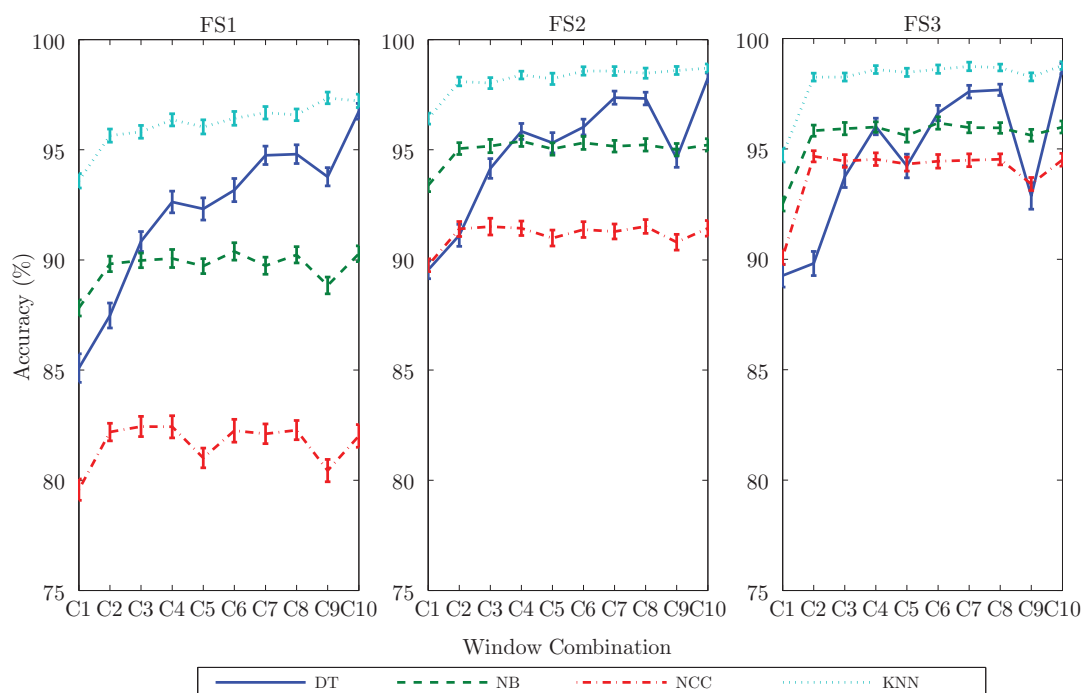


Figura 4.4: Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de las combinaciones C1-C10 escogidas para el testeo del algoritmo de fusión a 2 niveles. Todas las combinaciones utilizan el tamaño de ventana de 6 segundos (ver tabla 4.1)

Por último, el algoritmo *KNN* también ve incrementadas sus prestaciones a pesar del poco margen de mejora existente por su buen funcionamiento como clasificador de actividades. Aplicando indistintamente los 3 conjuntos de características se obtienen eficiencias de clasificación superiores al 97% (considerando la media de la señal) y al 98.7% (para 2 y 5 características), lo que supone mejoras del 2-4% aproximadamente. Quizás lo más reseñable sea el posicionamiento de la metodología con 5 características como óptima a partir de la aplicación del algoritmo de

fusión multiventana. En el capítulo 3 se comprueba la superioridad de *KNN* con 2 características frente al resto de metodologías planteadas. Sin embargo, la fusión de ventanas permite que la variante de 5 características la supere ligeramente. No obstante, las diferencias porcentuales son mínimas (inferiores al 0.3% en cualquier caso), de modo que la implementación del clasificador *KNN* utilizando la media y la desviación estándar de la señal inercial como características para el reconocimiento de la actividad se sigue postulando como la mejor opción posible entre los casos considerados ya que presenta resultados similares y aplica un repertorio de características más reducido, lo que supone un menor coste computacional en la etapa de extracción de características del sistema *ARC*.

Si se comparan estos resultados con el tamaño de ventana óptimo obtenido en el capítulo 3, se podría decir que la fusión no introduce ventajas significativas. Los porcentajes alcanzados con *KNN* ya se obtuvieron con la aplicación de tamaños de ventana reducidos (0.5-1 segundo). No obstante, se tratan de análisis y condiciones diferentes dada la respuesta del clasificador *KNN* frente a tamaños de ventana cortos. Este hecho se reafirma con el análisis de la tabla 4.3 y la figura 4.5.

	DT			NB			NCC			KNN		
C11	87.88	91.56	90.88	87.93	93.45	93.46	80.21	90.18	91.83	97.33	<i>98.77</i>	97.65
C12	89.38	92.41	92.60	89.87	95.14	96.29	82.38	91.46	94.79	97.20	98.94	<i>99.30</i>
C13	91.92	94.33	94.50	89.22	95.28	95.96	80.84	91.49	94.46	97.78	99.21	<i>99.30</i>
C14	94.18	96.89	97.21	89.88	95.29	96.30	82.05	91.41	94.74	98.36	99.48	99.53
C15	93.87	94.75	94.62	88.31	95.06	96.00	81.12	90.98	94.03	98.50	99.00	<i>99.30</i>

Tabla 4.3: Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento cada 3 segundos. Las ventanas seleccionadas son múltiplos de la ventana de mayor tamaño (3 segundos). En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada combinación aplicada

El algoritmo *KNN* continúa aumentando la eficiencia de clasificación considerando un instante temporal de decisión inferior a 6 segundos. En este caso, la aplicación de la fusión junto a la toma de decisión final en 3 segundos permite incrementar en más de un 1% los resultados comentados anteriormente para cada metodología, lo que implica para las variantes de 2 y 5 características alcanzar prácticamente la maximización de la eficiencia de clasificación en el reconocimiento de todas las actividades del dataset.

Por el contrario, el algoritmo *DT* únicamente no experimenta mejora con respecto a la fusión con decisión en 6 segundos, si no que empeora el reconocimiento global. Los resultados óptimos,

obtenidos para la combinación C14 (es decir, seleccionando las ventanas de 0.75, 1.5 y 3 segundos para la fusión) distan de los obtenidos para *fusión completa* en 6 segundos en torno al 1.5-2% (con eficiencias en el rango del 94-97% aproximadamente). Sin embargo, los clasificadores *NB* y *NCC* se muestran independientes de este factor, ya que los resultados para el conjunto global de combinaciones seleccionadas para este estudio (C11-C15) presentan similitudes porcentuales a las obtenidas para la fusión con decisión en 6 segundos.

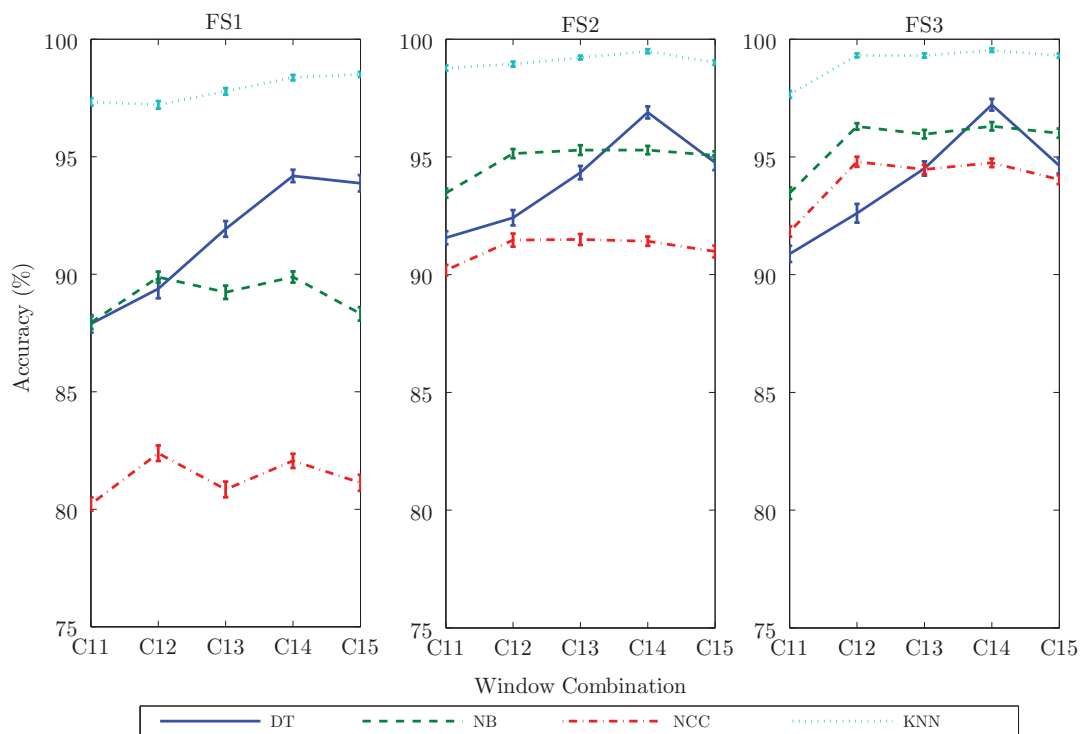


Figura 4.5: Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de las combinaciones C11-C15 escogidas para el testeo del algoritmo de fusión a 2 niveles. Todas las combinaciones utilizan el tamaño de ventana de 3 segundos (ver tabla 4.1)

Una última prueba se puede ver en la tabla 4.4 en la que se consideran otras combinaciones adicionales con decisiones finales en tiempos inferiores a 3 segundos. No obstante, 3 combinaciones proceden del estudio realizado en el capítulo 3 (C16, C18 y C19), dado que solamente se aplica un tamaño de ventana específico (1.5, 0.75 y 0.25 segundos, respectivamente) (ver tabla 4.1)

A la vista de los resultados, queda de manifiesto que continuar reduciendo el instante de decisión para los algoritmos *DT*, *NB* y *NCC* no aporta mejoras en la eficiencia de clasificación, si no que además supone una reducción de dicho parámetro. Por el contrario, el clasificador *KNN* no se observa especialmente afectado por este factor y conserva los porcentajes obtenidos

	DT			NB			NCC			KNN		
C16	89.08	92.42	92.46	86.79	92.81	93.77	79.72	89.48	91.80	97.55	<i>98.58</i>	98.08
C17	88.40	92.13	91.89	87.24	93.44	94.84	78.43	89.70	92.50	97.93	99.48	99.37
C18	88.13	90.90	90.99	83.34	89.62	91.31	72.93	84.12	87.47	97.20	<i>98.48</i>	98.04
C19	86.53	86.77	86.67	74.26	74.78	78.52	55.34	66.25	70.73	97.73	<i>98.36</i>	96.09

Tabla 4.4: Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 2 niveles con reconocimiento en tiempos reducidos. Se aplican ventanas de 1.5, 0.75 y 0.25 segundos para estas combinaciones. En cada fila se incluyen los resultados de la aplicación de una combinación específica de la tabla 4.1. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada combinación aplicada

tras la reducción del instante de decisión a 3 segundos. La combinación C17 alcanza la misma eficiencia de clasificación que la combinación C14 considerando 2 características. Es decir, que eliminando el tamaño de ventana de 3 segundos, se consigue igualar la decisión en la mitad de tiempo (1.5 segundos) manteniendo un porcentaje de reconocimiento de 99.48 %.

Un estudio adicional ha sido considerado en este proyecto. La fusión de diferentes tamaños de ventana que aporten sus reconocimientos a la decisión final dotan de mayor robustez a la decisión del clasificador. No obstante, es posible plantear un análisis de la influencia de la aplicación de una fusión con un tamaño de ventana único en un instante de decisión impuesto (por ejemplo, y por las referencias y experimentaciones anteriores, 6 segundos). El buen reconocimiento individual de un tamaño de ventana específico parece propicio para la obtención de mejoras a través de su fusión, mientras que la situación en la que los reconocimientos con tamaños de ventana concretos ofrecen porcentajes bajos se afronta como una incógnita. En este estudio se aplican los diferentes múltiplos de 6 segundos considerados en este proyecto (0.25, 0.5, 0.75, 1, 1.5, 2 y 3 segundos). La tabla 4.6 recoge todos los resultados que conciernen al testeo del algoritmo de fusión en su versión de 1 nivel, y que pueden ser comparados con los obtenidos en el capítulo 3 a partir de la tabla 4.5 seleccionando únicamente los tamaños de ventana que se han aplicado para la fusión. Adicionalmente, se presenta la figura 4.6 que permite observar las tendencias seguidas por las diferentes metodologías considerando el algoritmo de fusión a 1 nivel.

El planteamiento previo, con el que se indica el resultado esperado de la aplicación del algoritmo de fusión a 1 nivel, se reflejan en los resultados obtenidos y se puede interpretar comparando las tablas 4.5 y 4.6. Los clasificadores *DT* y *KNN* consiguen aumentar la eficiencia de clasificación progresivamente conforme se reduce el tamaño de ventana específico para la fusión y obtención de decisión en 6 segundos. Sin embargo, las tendencias presentadas por los

Sin fusión

	DT			NB			NCC			KNN		
0.25	86.54	86.77	86.68	74.26	74.79	78.53	55.35	66.25	70.73	97.74	<i>98.37</i>	96.09
0.5	86.81	89.51	89.45	79.56	85.83	87.75	68.10	79.65	83.48	97.92	<i>98.70</i>	97.95
0.75	88.14	90.91	90.99	83.35	89.63	91.31	72.94	84.12	87.48	97.21	<i>98.48</i>	98.04
1	88.88	92.03	92.27	85.14	91.89	93.32	77.30	87.38	90.44	97.71	98.84	98.35
1.5	89.08	92.43	92.47	86.80	92.81	93.77	79.72	89.48	91.81	97.55	<i>98.59</i>	98.08
2	88.90	92.29	92.42	86.60	92.97	93.85	80.26	90.18	92.21	97.65	<i>98.68</i>	98.16
3	87.89	91.56	90.88	87.94	93.46	93.46	80.22	90.19	91.83	97.33	<i>98.77</i>	97.65
6	85.09	89.55	89.26	87.82	93.39	92.50	79.58	89.76	90.10	93.59	<i>96.41</i>	94.72

Tabla 4.5: Eficiencia de clasificación (en %) de las 12 metodologías planteadas sin aplicar fusión. Las ventanas seleccionadas son múltiplos del tamaño de ventana de referencia (6 segundos). En cada fila se incluyen los resultados de la aplicación de un tamaño de ventana específico en la etapa de segmentación. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada tamaño de ventana aplicado

Con fusión

	DT			NB			NCC			KNN		
0.25	99.94	99.91	99.93	91.95	93.61	96.72	73.42	83.43	88.44	100.0	100.0	100.0
0.5	98.98	99.86	99.85	93.43	97.23	98.51	83.06	91.66	95.62	100.0	100.0	100.0
0.75	98.23	99.73	99.80	94.44	98.12	98.72	86.31	93.75	97.08	99.94	<i>99.99</i>	<i>99.99</i>
1	98.14	99.62	99.70	94.79	98.24	98.63	88.47	95.62	97.64	99.93	<i>99.99</i>	<i>99.99</i>
1.5	97.00	98.91	99.07	94.20	97.91	98.20	89.05	95.51	97.51	99.70	99.96	<i>99.98</i>
2	95.59	97.93	97.84	93.06	98.01	98.29	87.55	95.09	97.05	99.52	99.88	<i>99.92</i>
3	91.24	94.14	93.27	92.10	96.75	96.63	85.68	93.74	96.06	97.91	98.86	<i>99.24</i>
6	85.09	89.55	89.26	87.82	93.39	92.50	79.58	89.76	90.10	93.59	<i>96.41</i>	94.72

Tabla 4.6: Eficiencia de clasificación (en %) de las 12 metodologías planteadas aplicando el algoritmo de fusión multiventana a 1 nivel. Las ventanas seleccionadas son múltiplos del tamaño de ventana de referencia (6 segundos). En cada fila se incluyen los resultados de la aplicación de un tamaño de ventana específico en la fusión. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada tamaño de ventana aplicado

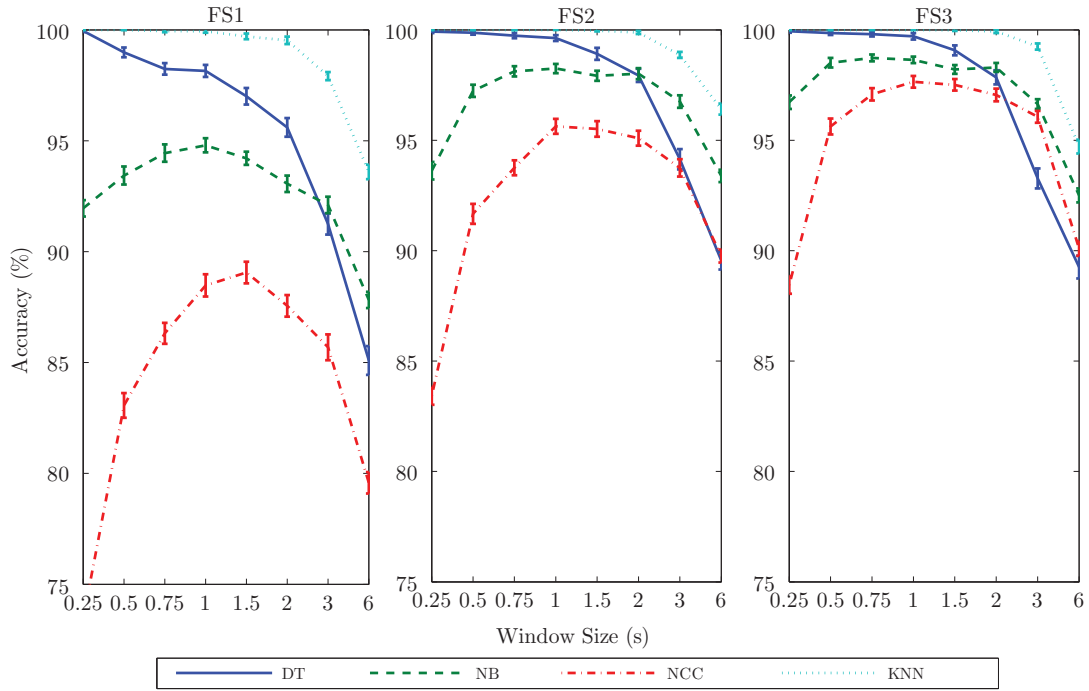


Figura 4.6: Eficiencia de clasificación (en %) de los 4 algoritmos de clasificación frente a los 3 conjuntos de características seleccionados (FS1, FS2, FS3) en función de 8 ventanas seleccionadas para el testeo del algoritmo de fusión a 1 nivel.

clasificadores *NB* y *NCC* son similares entre sí, pero diferentes a los otros 2 clasificadores. La lectura que se puede hacer es la siguiente. Observando la tabla 4.5 se comprueba que el clasificador *DT* presenta eficiencias de clasificación superiores al 86.5 % para cualquier tamaño de ventana seleccionado inferior a 6 segundos. Cuanto menor es el tamaño de ventana de la etapa de segmentación, mayor es el número de ventanas implicadas en la decisión adoptada finalmente en 6 segundos. Por ejemplo, la fusión de decisiones para ventanas de 0.25 segundos con una eficiencia de 86.54 % por ventana utilizando únicamente la media de la señal implica la fusión de 24 decisiones con un porcentaje de *acierto* del 86.54 %. La intuición marca que, probablemente, cuando haya que ofrecer un resultado tras 6 segundos de la ejecución de una actividad, el reconocimiento sea correcto y el porcentaje asociado muy cercano a la maximización, como así ocurre: se obtiene una eficiencia de clasificación de 99.94 %. De igual forma, si se opta por utilizar una segmentación en ventanas de 0.5 segundos, se obtienen 12 decisiones en 6 segundos con un porcentaje del 86.81 %, lo que lleva a que su aplicación en fusión permita obtener una eficiencia de clasificación del 98.98 %. Conforme disminuyen las decisiones parciales (menos ventanas de decisión en el tiempo de referencia de 6 segundos), *menos acertada* es la decisión de la fusión a pesar de reconocer las actividades con porcentajes altos. Si se observa la tabla 4.5,

para 1.5 segundos se obtuvo la optimización para la metodología que considera la media de la señal. No obstante, dado que únicamente se aplican 4 ventanas con un porcentaje del 89.08 % en 6 segundos, la decisión de fusión es menos *robusta* que la aplicación de 24 ventanas de 0.25 segundos con un porcentaje del 86.54 % en el tiempo de referencia. Este análisis es igualmente válido tanto para las variantes restantes de *DT* como las metodologías del clasificador *KNN*, que consigue maximizar el reconocimiento para los tamaños de ventana mínimos seleccionados para este proyecto. Por ejemplo, la consideración de 24 ventanas de 0.25 segundos con un reconocimiento del 97.74 % se traduce en una decisión de fusión del 100 %. Y es que las eficiencias de clasificación para los tamaños de ventana seleccionados alcanzan porcentajes superiores al 97.3 % sin aplicar fusión, lo que ya permite ser optimistas en cuanto a la mejora de la eficiencia de clasificación con la aplicación del algoritmo de fusión.

En lo que respecta a la implementación e integración del algoritmo en las decisiones adoptadas en un sistema *ARC* empleando *DT* o *KNN*, probablemente habría que hacer un balance de prioridades para la selección de un tamaño de ventana específico u otro. La versión del algoritmo de fusión con decisión en 1 nivel permite mantener fijado el tamaño de ventana durante el reconocimiento de actividades. Vistos los porcentajes obtenidos para la eficiencia de clasificación y con el fin de evitar sobrecargas de procesamiento y ventanado en la etapa de segmentación, mayor número de extracciones de características en la etapa de extracción y mayor toma de decisiones en la etapa de clasificación, lo adecuado es penalizar mínimamente la eficiencia de clasificación a costa de reducir el número de ventanas implicadas. El algoritmo *DT* supera el 99 % de eficiencia aplicando el tamaño de ventana mínimo seleccionado para el estudio (0.25 segundos) considerando sólo la media de la señal, pero con 2 características se supera con ventanas de 1 segundo y con 5 características a partir de la aplicación de segmentaciones de 1.5 segundos. Se trata de hacer un balance entre lo que se puede asumir como parámetro penalizable y lo que se debe exigir en cuanto a prestaciones para obtener un resultado a la altura de las expectativas esperadas en lo que respecta al mejor reconocimiento posible de cualquier actividad.

Los clasificadores *NB* y *NCC* presentan tendencias distintas a las comentadas anteriormente. El algoritmo bayesiano reduce sus prestaciones considerablemente para tamaños de ventana bajos, con una eficiencia en torno al 74 % para segmentación en ventanas de 0.25 segundos aplicando únicamente la media de la señal. Aún así, la fusión de 24 decisiones en 6 segundos consigue aumentar la eficiencia hasta el 91.95 %, lo que supone una mejora superior al 17 %. Conforme se aumenta el tamaño de ventana, consigue aumentar la eficiencia de clasificación (aunque no linealmente como ya se extrajo de los análisis realizados en el capítulo 3 sección 3.4). Por tanto, aunque por un lado se aumente la eficiencia de clasificación individual de las ventanas aplicadas en la fusión, menor es el número de ventanas disponibles para la fusión en

6 segundos, por lo que los resultados de fusión obtenidos para el clasificador *NB* consiguen una optimización de la eficiencia en la aplicación de un tamaño de ventana intermedio. En el caso de utilizar la media de la señal exclusivamente, la fusión de 6 ventanas de 1 segundo con eficiencias del 85.14 % permiten un resultado de fusión del 94.79 %, lo que supone una mejora de casi el 10 %. El resto de variantes consideradas para el clasificador *NB* presentan tendencias similares, consiguiendo la optimización para 2 y 5 características con la aplicación de ventanas de 1 y 0.5 segundos, respectivamente.

El clasificador *NCC*, que alcanza un 55.35 % de eficiencia aplicando una ventana de 0.25 segundos considerando la media de la señal, aumenta sus prestaciones hasta 73.42 % aplicando la fusión de 24 decisiones en 6 segundos, lo que supone una mejora superior al 18 %. La eficiencia óptima de esta metodología se alcanza para la fusión con ventanas de 1.5 segundos, alcanzando una eficiencia de 89.05 %. Las variantes de 2 y 5 características optimizan la eficiencia de clasificación con la fusión de ventanas de 1 segundo, que permite obtener un porcentaje de reconocimiento del 95.62 % y 97.64 %. Con el análisis de la influencia del tamaño de ventana y la introducción del algoritmo de fusión, las prestaciones del clasificador del centroide más cercano han aumentado considerablemente.

Con respecto a la principal referencia para la segmentación de un conjunto amplio de actividades (6 segundos), las mejoras alcanzadas con la aplicación de la fusión son considerables. La tabla 4.7 compara las eficiencias de reconocimiento global del ventanado tradicional con la aplicación del algoritmo de fusión multiventana en sus 2 versiones.

	DT			NB			NCC			KNN		
T	85.09	89.55	89.26	87.82	93.39	92.50	79.58	89.76	90.10	93.59	<i>96.41</i>	94.72
F1	99.94	99.91	99.93	94.79	98.24	98.72	89.05	95.62	97.64	100.0	100.0	100.0
F2	96.74	98.25	98.63	90.38	95.38	95.98	82.44	91.52	94.66	97.35	98.70	98.77

Tabla 4.7: Comparación en términos de eficiencia de clasificación (en %) de los mejores resultados alcanzados para las 12 metodologías planteadas entre la aplicación de ventanado tradicional (T), la aplicación del algoritmo de fusión a 1 nivel (F1) y la aplicación del algoritmo de fusión a 2 niveles (F2) con decisión final cada 6 segundos. En cada fila se incluyen los mejores resultados de cada opción. En cada columna se incluyen los resultados de los 4 algoritmos de clasificación y los 3 conjuntos de características seleccionados (FS1, FS2, FS3). En negrita, el mayor rendimiento posible para cada metodología. En cursiva, la metodología con el mayor rendimiento para cada opción considerada

El clasificador que emplea el algoritmo de árbol de decisión consigue aumentar su eficiencia de clasificación hasta prácticamente la maximización de dicho parámetro seleccionando cualquiera de los conjuntos de características aplicados en este proyecto, algo que sí ocurre con el clasificador de los K-Vecinos, que obtiene porcentajes del 100 % aplicando el algoritmo de fusión a 1 nivel. Las

mejoras experimentadas por los clasificadores bayesiano y centroide más cercano son también notables, con mejoras acotadas en un rango aproximado entre el 5 % y 9.5 % según la metodología seleccionada. El peor resultado se obtiene para el clasificador *NCC* empleando únicamente la media de la señal y aún así se alcanza una eficiencia de reconocimiento global superior al 89 %.

A continuación, tras los análisis realizados en esta sección, se presentan las principales conclusiones extraídas del estudio y testeo del algoritmo de fusión multiventana.

4.5. Conclusiones

La implementación y aplicación del algoritmo de fusión multiventana ha introducido mejoras significativas en lo que respecta al aumento de la eficiencia de clasificación para las 12 metodologías planteadas en este proyecto. Dependiendo de las características intrínsecas que definen el funcionamiento de cada clasificador, el aumento de prestaciones de los clasificadores ha seguido diferentes evoluciones con la introducción de la fusión en el sistema *ARC*.

Teniendo siempre presente el punto de partida de este proyecto, en el que existían estudios específicos que demuestran el reconocimiento de ciertas actividades para tamaños de ventana diferentes (Mantjarvi et al. (60) (2001), Kern et al. (61) (2003), Huynh y Schiele (64) (2005), Maurer et al. (3) (2006) y Ward et al. (65) (2006)) y otros trabajos con un tratamiento amplio de actividades que indican que un tamaño de ventana de 6 segundos es suficiente para un buen reconocimiento del conjunto de actividades seleccionadas (Bao e Intille (1) (2004) y Ravi et al. (63) (2005)), o que incluso indican un tamaño de ventana mayor (Krause et al. (62) (2003)), la mejora obtenida con la introducción del algoritmo de fusión en este capítulo 4 ha quedado de manifiesto.

Las diferentes posibilidades existentes permiten un diseño y una configuración a medida para el reconocimiento de actividades, lo que supone alcanzar un compromiso entre la selección de los diferentes parámetros de influencia en un sistema *ARC*.

La introducción del concepto de fusión permite obtener una mayor robustez final en el reconocimiento global del clasificador a costa de requerir más recursos de procesamiento. La gran ventaja del algoritmo de fusión multiventana reside en el aprovechamiento de los datos procesados para ser segmentados en tiempo real sin necesidad de otros parámetros externos para la toma de decisiones.

Su implementación y consideración en futuros trabajos del sistema *ARC* se postula como fundamental para mejorar el reconocimiento de actividades tradicional.

5

Capítulo V: Conclusiones generales y trabajo futuro

5.1. Introducción

En este último capítulo, se resumen las principales conclusiones presentadas a lo largo del proyecto. Además se incluyen algunas ideas sobre la aplicabilidad del efecto del tamaño de ventana y la viabilidad de su potencial implementación. Por último, se presentan las principales ventajas e inconvenientes del algoritmo de fusión presentado, así como posibles mejoras que motiven su desarrollo en futuros trabajos sobre el reconocimiento de la actividad.

5.2. Conclusiones

A lo largo de este proyecto se ha realizado un tratamiento amplio de la temática que concierne al reconocimiento de la actividad. Un primer análisis sobre las tendencias actuales en el reconocimiento de actividades (*concept drift, self-calibration, adaptation, windowing* etc.) ha permitido tomar conciencia sobre los principales factores de influencia que se encuentran en investigación y por resolver adecuadamente para un mejor reconocimiento de la actividad humana. En concreto, se ha encontrado una motivación especial en abordar el efecto del *windowing* en un sistema *ARC* y se ha convertido en el eje principal sobre el que ha girado este proyecto.

En primer lugar se ha realizado un estudio previo sobre los trabajos anteriores que han tratado este tema. Varios autores han demostrado la influencia del tamaño de ventana en el reconocimiento particular de una o varias actividades, mostrando resultados dispares sobre la optimización del tamaño de ventana para alcanzar el mejor reconocimiento posible de las actividades tratadas. Otros han indicado que, en búsqueda de reconocer un amplio abanico de actividades, lo idóneo es considerar una segmentación de los datos monitorizados en ventanas de

larga duración, por ejemplo, 6 segundos. Sin embargo, las condiciones experimentales de cada trabajo no permiten ofrecer una idea generalizada sobre el reconocimiento de actividades por diferentes decisiones: estudio de actividades muy específicas (en muchos casos, 4-5 actividades), monitorización de pocos sujetos (en algunos casos, únicamente el investigador), utilización de un número reducido de sensores (en ciertas ocasiones, 1-2 sensores), etc. Por tanto, ya se observa que existe cierto vacío en lo que respecta a ofrecer un estudio amplio en el que se consideran muchas actividades, un número de sujetos considerable y la monitorización a partir de varios sensores localizados. Disponer de un dataset con el que se han recogido datos biomecánicos a partir de 33 actividades, 17 sujetos y un total de 9 sensores ha ofrecido la oportunidad de realizar un análisis experimental que permite rellenar en parte ese vacío, así como ofrecer una visión más general sobre el efecto del tamaño de ventana al tratar de reconocer un abanico amplio de actividades. Por supuesto, a la vez que se indica que se trata de un estudio que permite dar una idea global sobre la consideración de este factor, también se sabe que no deja de ser un estudio específico en el que se tratan 33 actividades muy concretas. Sin embargo, cualquier estudio está constituido por una serie de datos y no deja de ser específico, por lo que simplemente se intenta ofrecer un estudio lo más robusto y concluyente posible que permita disponer de unas conclusiones más generales.

Con estas premisas, la prioridad del estudio llevado a cabo en este proyecto se centra en averiguar qué efecto tiene la variación del tamaño de ventana sobre la eficiencia global del clasificador aplicado. A partir de 4 clasificadores y 3 conjuntos de características muy contrastados en el panorama científico en el tratamiento de esta temática y otras diferentes, se ha podido comprobar que no solamente influye, sino que por lo general se consigue optimizar la eficiencia del clasificador para tamaños de ventana inferiores a 6 segundos. Dependiendo del funcionamiento específico de cada algoritmo de clasificación, el tamaño de ventana óptimo alcanza diferentes cotas. Desde un tamaño de ventana óptimo de 0.5 segundos, obtenido a partir del testeo del clasificador *KNN* empleando únicamente la media de la señal y que ofrece una eficiencia de clasificación de casi el 98 %, hasta un tamaño de ventana óptimo de 4.25 segundos para diferentes metodologías de los clasificadores bayesiano y centroide más cercano, con eficiencias de clasificación superiores al 88.8 %. En este sentido, la importancia del estudio cobra mayor relevancia al comprobar que reduciendo el tamaño de ventana se conseguía adicionalmente optimizar la eficiencia de clasificación.

Dado que, por estudios previos se conoce que la influencia sobre cada actividad en particular influye de manera diferente, el siguiente objetivo supone comprobar qué tamaño de ventana permite obtener una eficiencia óptima de cada actividad considerada en este proyecto. De las simulaciones experimentales se han extraído diferentes parámetros sobre el reconocimiento de

actividades. Entre ellos, el parámetro que ofrece una idea general sobre el reconocimiento de actividades es la matriz de confusión que permite observar el reconocimiento particular de cada actividad frente al resto. Sin embargo, no se dispone de una métrica directa que permita evaluar tanto los fallos en el reconocimiento (al ser confundida con otras) como los aciertos de reconocimiento. Tras una búsqueda sobre métricas que tengan en cuenta estas consideraciones se ha encontrado la métrica $F_1 - score$, que utiliza la precisión y la sensibilidad para elaborar una media ponderada que ofrece una magnitud del reconocimiento particular de cada actividad. La métrica $F_1 - score$ ha permitido realizar diferentes estudios y análisis sobre el efecto del tamaño de ventana en el reconocimiento particular de cada actividad.

En general, todas las actividades se pueden reconocer con altas garantías (minimizando el error que se comete al intentar reconocerlas) ateniéndose a la medida que ofrece $F_1 - score$ y seleccionando adecuadamente la metodología que ofrezca mejores resultados. El peor reconocimiento se ha obtenido para la actividad *saltar hacia arriba* con un valor de $F_1 - score$ máximo de 0.941 aplicando un tamaño de ventana óptimo de 3 segundos y el clasificador *KNN* junto a la media y la desviación estándar de la señal. Esto quiere decir que esta actividad no se va a reconocer mejor permitiendo un mayor tiempo de monitorización antes de segmentar (un tamaño de ventana mayor), así que habría que pensar si otros conjuntos de características y el uso de otro clasificador pudiese mejorar su reconocimiento en estas condiciones. El resto de actividades presentan un $F_1 - score \geq 0.975$ como reconocimiento óptimo para alguna de las metodologías de *KNN*, lo que indica que cualquier actividad puede ser reconocida con alta probabilidad de éxito en tiempo real pudiendo tomar cualquier decisión al respecto. No obstante, algunas actividades requieren tamaños de ventana de larga duración para alcanzar estos valores de optimización (6-7 segundos). Debido a este hecho, se han obtenido representaciones en las que se puede observar qué tamaño de ventana es necesario para reconocer una actividad según las exigencias marcadas por el propio investigador o por las restricciones que vengán impuestas en el sistema *ARC*. Sin duda, se trata de alcanzar un compromiso entre diferentes parámetros (tamaño de ventana, eficiencia de clasificación, recursos computacionales, tiempo de respuesta, etc.) hasta alcanzar un diseño que responda a las expectativas creadas para un reconocimiento lo más óptimo posible de cada actividad. En este sentido, se trata de alcanzar un grado de exigencia sobre la monitorización explícita de cualquier actividad.

En el caso de que se desee aplicar alguna implementación concreta que permita minimizar la penalización en el reconocimiento global de todas las actividades consideradas en el dataset, se ha presentado la tabla 3.6 en el capítulo 3 que permite seleccionar a título personal un tamaño de ventana específico para una metodología concreta. El clasificador *KNN*, utilizando un tamaño de ventana de 2 segundos y las características de la media y la desviación estándar

de la señal, se ha postulado como la mejor opción entre las metodologías estudiadas ya que es el clasificador que menos penaliza el reconocimiento global de las actividades con un $F_1 - score \geq 0.941$ garantizado para reconocer cualquier actividad. Si por el contrario se desea reconocer una actividad específica, se ha presentado en la sección 3.4 un estudio amplio para cada actividad incluida en el dataset donde se indican las mejores implementaciones posibles entre las metodologías y tamaños de ventana seleccionados para este proyecto.

Posteriormente, vistos los resultados alcanzados en el capítulo 3, surge un interés sobre si es posible conseguir aumentar la eficiencia del reconocimiento de actividades aplicando algún algoritmo que utilice los conocimientos adquiridos sobre el efecto del tamaño de ventana en el sistema *ARC* y que introduzca mayor robustez en la decisión final. Un algoritmo de fusión multiventana ha sido implementado y testeado aplicando 2 versiones diferentes: tomando decisiones a 1 nivel de fusión y a 2 niveles de fusión. De esta forma, el peor resultado se ha obtenido para el clasificador *NB* que utiliza únicamente la media de la señal, con una eficiencia de 89.05% y fusión multiventana de 1 nivel aplicando 4 ventanas de 1.5 segundos para ofrecer la decisión en 6 segundos. Clasificadores como *DT* prácticamente alcanzan la maximización del *performance* o *KNN*, que maximiza el reconocimiento considerando 24 decisiones de 0.25 segundos y ofrecer la decisión en 6 segundos. En lo que respecta a la versión de 2 niveles, la aplicación del algoritmo de fusión también incrementa la eficiencia de clasificación considerablemente aplicando *fusión completa* para 6 de las 12 metodologías estudiadas con tasas que superan en cualquier caso el 96%.

La mejora final obtenida tanto en términos de reducción del tamaño de ventana como de optimización de la eficiencia de clasificación ha quedado patente y los resultados obtenidos se espera que sean de gran ayuda para el planteamiento de futuros trabajos.

5.3. Aplicabilidad y viabilidad

Ha quedado de manifiesto que cualquier actividad se puede reconocer con altas garantías de éxito considerando tamaños de ventana inferiores a 6 segundos. Un sondeo continuo en situaciones en las que se desean monitorizar actividades muy concretas permite reducir el tiempo de decisión considerablemente.

Aplicado a la teleasistencia, se podría controlar si una persona mayor está andando o no y si drásticamente se interrumpe dicho ejercicio (por ejemplo, debido a una caída), ya que se puede controlar eficientemente si el sistema *ARC* se ha implementado con el clasificador *KNN* utilizando la media y la desviación estándar de la señal inercial con un tamaño de ventana mínimo de 0.25 segundos y un $F_1 - score = 0.984$. Probablemente, imponer 0.25 segundos

suponga una atención desmesurada para cubrir la monitorización de esta actividad y se pueda asumir un ventanado de mayor duración, que requiera un sondeo menos agresivo. Entonces, se podría optar por un clasificador *DT* utilizando únicamente la media de la señal y una ventana de 1.75 segundos, que permite reconocerla con un $F_1 - score = 0.945$. De esta forma, se reduce el repertorio de características, lo que supone un menor coste computacional del sistema *ARC* y se toman decisiones con un margen de tiempo mayor.

Aplicado al control deportivo, puede ser interesante monitorizar actividades como *correr*, *trotar* o *remar*. A lo largo de este proyecto se ha comprobado que la actividad *remar* se consigue reconocer aplicando tamaños de ventana muy reducidos para cualquier metodología estudiada (la figura superior en 3.18). De esta forma, un equipo de remo podría querer monitorizar la realización de dicho ejercicio para optimizar sus movimientos, así como su entrenador podría conocer el estado de ejecución de la actividad para corrección de errores. Cualquier actividad deportiva puede encontrar similitudes en el patrón de ejecución de sus acciones en gran parte de las actividades tratadas en este proyecto. Ciclismo (a partir de la actividad *hacer bicicleta estática*), la ejecución del estilo *mariposa* en natación (a partir de las actividades *rotación intensa de hombros* y *rotación moderada de hombros* junto a alguna otra actividad que monitorice las piernas), halterofilia (a partir de la actividad *elevación frontal de brazos*), esquí de fondo (a partir de la mezcla de la actividad *hacer bicicleta elíptica* y *andar*), etc.

Aplicado a la rehabilitación física, monitorizar el movimiento específico de cualquier extremidad que está dañada o ha sido operada. El dataset que se ha tratado en este proyecto presenta multitud de actividades que están dedicadas exclusivamente a la ejercitación de una parte específica del cuerpo (rodillas, brazos, codos, piernas, cintura, etc.)

Sobre la viabilidad de implementación del algoritmo de fusión en el sistema *ARC*, no sería un problema ya que supondría la adición de una etapa más. Incluso sin la utilización del algoritmo de fusión, se obtienen muy buenos resultados reconociendo cualquier actividad considerada, por lo que se trataría de ajustar específicamente el tamaño de ventana óptimo para la monitorización de alguna actividad en concreto. Sobre la viabilidad de aplicación sobre personas, hay que pensar en la idea presentada por Amft et al. (68) en 2009: se espera la llegada de una nueva generación de sensores diminutos que podrán ser integrados en la ropa o cualquier objeto que nos acompaña en nuestra vida diaria (cartera, teléfono móvil, reloj de pulsera, gafas, etc.).

5.4. Ventajas e inconvenientes

La principal ventaja adquirida con el efecto del tamaño de ventana es conseguir optimizar el reconocimiento de cada actividad en tiempos de segmentación de datos muy pequeños. En muchos casos, se pueden adoptar decisiones con altas garantías de reconocimiento en tiempos tan reducidos como 0.25 segundos, un tiempo en el que prácticamente se ha podido llevar a cabo la realización de un gesto. En cuanto a la aplicación del algoritmo de fusión, permite una mejora de la eficiencia de clasificación con respecto al ventanado tradicional, lo cual es motivo suficiente para pensar en su potencial aplicación.

El principal inconveniente reside en la introducción de mayor coste computacional con la implementación del algoritmo de fusión y la segmentación de los datos monitorizados en varios tamaños simultáneamente, más si cabe en los casos en los que se consiguen optimizaciones considerando 5 características en la etapa de extracción. Aún así, se obtienen buenos resultados aplicando 1 y 2 características exclusivamente (como se ha demostrado para *KNN* con 2 características y una ventana de 1 segundo) sin necesidad de aplicar el algoritmo de fusión.

5.5. Posibles mejoras y trabajo futuro

El estudio realizado en este proyecto presenta resultados importantes en aspectos como la optimización del tamaño de ventana para cada actividad (capítulo 3) o la aplicación de un algoritmo de fusión que mejora la eficiencia de clasificación (capítulo 4), pero también presenta ciertas limitaciones en lo que respecta a la toma de decisiones en tiempo real.

Pensando en trabajos futuros, los resultados presentados en este proyecto son un punto de partida para la obtención de un sistema adaptativo que modifique en tiempo real el tamaño de ventana en el reconocimiento de actividades. Puesto que se pueden identificar perfectamente actividades con porcentajes próximos al reconocimiento perfecto, se puede implementar una etapa de realimentación que modele un ventanado adaptativo en tiempo real y conmute el segmentado de datos a otro tamaño de ventana.

La utilización de métricas como el F_1 - *score* ofrecen una alta confiabilidad en términos de reconocimiento ponderando lo preciso que es el clasificador reconociendo cada actividad y la confusión que puede aparecer con otras actividades. Sin embargo, se podrían evaluar otro tipo de métricas que consideren otros parámetros obtenidos de la clasificación de actividades (especificidad, valor predictivo negativo, matriz de confusión, etc.) o incluso los mismos que se han considerado en este proyecto (precisión y sensibilidad). Sokolova et al. (83) utiliza la especificidad y la sensibilidad para obtener otros parámetros a partir de éstos como el *índice*

de *Youden*, *Likelihood test* (algo así como prueba de probabilidad) o *Discriminant Power* (algo así como capacidad de discriminación).

En lo que respecta a otras consideraciones en torno a la etapa de segmentación, se podría pensar en la introducción del solapamiento de ventanas. Se ha demostrado que un solapamiento del 50 % ha sido todo un éxito en trabajos previos como los presentados por De Vault y Dunn (86) en 2001 o Van Laerhoven y Cakmakci (87) en el año 2000. Este factor permite incrementar la eficiencia de clasificación a costa de introducir un procesamiento adicional de los datos monitorizados.

En cuanto a los datos de monitorización empleados, se han empleado los datos inerciales por simplificar el estudio y que pueda ser abordable el tratamiento y el análisis para el desarrollo de este proyecto. Sin embargo, Baños et al. (71) presentan un dataset con diferentes datos adquiridos por giróscopos o magnetómetros, entre otros, que pueden ser agregados y considerados en cualquier estudio de reconocimiento de actividades para comprobar su influencia. Probablemente, existan mejores combinaciones de datos monitorizados que permitan incrementar el reconocimiento de actividades en conjunción con una selección óptima de los sensores que aportan información determinante sobre la identificación de cada actividad en particular. Lo mismo ocurre si se consideran los otros 2 escenarios de monitorización recogidos en el dataset: *calibrados* y *mutuos*. De sus análisis se pueden extraer conclusiones adicionales a las presentadas en este proyecto.

De la misma forma, en el capítulo 4 se consideran 19 de las 31 combinaciones posibles para los 5 tamaños de ventana seleccionados para el testeo del algoritmo de fusión a 2 niveles. Para este proyecto se han descartado las 12 combinaciones que faltan debido a las exigencias computacionales que requiere temporalmente al incluir la ventana de 0.25 segundos. Aún así, para este proyecto se han seleccionado las 3 combinaciones más interesantes que emplean este tamaño de ventana: la *fusión completa* (C10) y otras dos combinaciones en las que se emplea conjuntamente con el tamaño de ventana que coincide con el instante de decisión (C9 para 6 segundos y C15 para 3 segundos)(ver tabla 4.1). Nuestra idea es completar las simulaciones que faltan en próximas fechas para disponer de una opinión e interpretación más objetiva de los resultados que se alcanzan en dicho capítulo, además de que dichos resultados estén disponibles para la realización de futuros análisis y trabajos que se desarrollen sobre esta temática.

Lo que queda claro es que la temática que engloba al reconocimiento de actividades tiene aún mucho recorrido por delante, un amplio campo de investigación para optimizar cualquier etapa del sistema *ARC* y lo que se espera es que este proyecto aporte su granito de arena en la consecución de futuros trabajos.

Referencias

- [1] LING BAO AND STEPHEN S INTILLE. **Activity recognition from user-annotated acceleration data.** In *Pervasive Computing*, pages 1–17. Springer, 2004. 4, 10, 20, 21, 22, 74, 76, 86
- [2] MIKKA ERMES, JUHA PARKKA, JANI MANTYJARVI, AND ILKKA KORHONEN. **Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions.** *Information Technology in Biomedicine, IEEE Transactions on*, **12**(1):20–26, 2008. 4
- [3] UWE MAURER, ASIM SMAILAGIC, DANIEL P SIEWIOREK, AND MICHAEL DEISHER. **Activity recognition and monitoring using multiple sensors on different body positions.** In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE, 2006. 4, 11, 21, 22, 74, 86
- [4] NAM PHAM AND TAREK ABDELZAHER. **Robust dynamic human activity recognition based on relative energy allocation.** In *Distributed Computing in Sensor Systems*, pages 525–530. Springer, 2008. 4
- [5] RAGHU K. GANTI, PRAVEEN JAYACHANDRAN, TAREK F. ABDELZAHER, AND JOHN A. STANKOVIC. **SATIRE: a software architecture for smart AtTIRE.** In *Proceedings of the 4th international conference on Mobile systems, applications and services, MobiSys '06*, pages 110–123, New York, NY, USA, 2006. ACM. 4
- [6] JHUN-YING YANG, JEEN-SHING WANG, AND YEN-PING CHEN. **Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers.** *Pattern recognition letters*, **29**(16):2213–2220, 2008. 5
- [7] ANTHONY F DALTON AND GEARÓID O'LAIGHIN. **Identifying activities of daily living using wireless kinematic sensors and data mining algorithms.** In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, pages 87–91. IEEE, 2009. 5
- [8] ORESTI BAÑOS, MIGUEL DAMAS, HÉCTOR POMARES, IGNACIO ROJAS, MÁTÉ ÁTTILA TÓTH, AND OLIVER AMFT. **A benchmark dataset to evaluate sensor displacement in activity recognition.** In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12*, pages 1026–1035, New York, NY, USA, 2012. ACM. 5, 17, 22
- [9] KAI KUNZE AND PAUL LUKOWICZ. **Dealing with sensor displacement in motion-based onbody activity recognition systems.** In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 20–29. ACM, 2008. 5, 8
- [10] RICARDO CHAVARRIAGA, HAMIDREZA BAYATI, AND JOSÉ DEL MILLÁN. **Unsupervised adaptation for acceleration-based activity recognition: robustness to sensor displacement and rotation.** *Personal and Ubiquitous Computing*, **17**(3):479–490, 2011. 5
- [11] HAMIDREZA BAYATI, JOSÉ DEL R MILLÁN, AND RICARDO CHAVARRIAGA. **Unsupervised adaptation to on-body sensor displacement in acceleration-based activity recognition.** In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pages 71–78. IEEE, 2011. 5
- [12] DEA ROGGEN, K FORSTER, ALBERTO CALATRONI, THOMAS HOLLECZEK, YU FANG, G TROSTER, PAUL LUKOWICZ, GERALD PIRKL, DAVID BANACH, KAI KUNZE, ET AL. **OPPORTUNITY: Towards opportunistic activity and context recognition systems.** In *World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a*, pages 1–6. IEEE, 2009. 5, 22
- [13] SHUANGQUAN WANG, JIE YANG, NINGJIANG CHEN, XIN CHEN, AND QINFENG ZHANG. **Human activity recognition with user-free accelerometers in the sensor networks.** In *Neural Networks and Brain, 2005. ICNN&B'05. International Conference on*, **2**, pages 1212–1217. IEEE, 2005. 5
- [14] I ZLIOBAITE AND M PECHENIZKIY. **Handling Concept Drift in Information Systems.** **2010.** 6
- [15] JOSEP CARMONA AND RICARD GAVALDÀ. **Online techniques for dealing with concept drift in process mining.** In *Advances in Intelligent Data Analysis XI*, pages 90–102. Springer, 2012. 6
- [16] DIPTI D PATIL, JYOTI G MUDKANA, DNYANESHWAR ROKADE, AND VIJAY M WADHAI. **Concept Adapting Real-Time Data Stream Mining for Health Care Applications.** In *Advances in Computer Science, Engineering & Applications*, pages 341–351. Springer, 2012. 6
- [17] ROBERTO CALANDRA, TAPANI RAIKO, MARC PETER DEISENROTH, AND FEDERICO MONTESINO POUZOLS. **Learning deep belief networks from non-stationary streams.** In *Artificial Neural Networks and Machine Learning-ICANN 2012*, pages 379–386. Springer, 2012. 6
- [18] BARRY SMIT AND JOHANNA WANDEL. **Adaptation, adaptive capacity and vulnerability.** *Global environmental change*, **16**(3):282–292, 2006. 6
- [19] CARLO GHEZZI. **Evolution, adaptation, and the quest for incrementality.** In *Large-Scale Complex IT Systems. Development, Operation and Management*, pages 369–379. Springer, 2012. 6
- [20] LIANG GUANG, ETHIOPIA NIGUSSIE, JUHA PLOSILA, JOUNI ISOAHO, AND HANNU TENHUNEN. **Survey of self-adaptive nocs with energy-efficiency and dependability.** *International Journal of Embedded and Real-Time Communication Systems (IJERTCS)*, **3**(2):1–22, 2012. 6
- [21] RENÉ JIMENEZ-FABIAN AND OLIVIER VERLINDEN. **Review of control algorithms for robotic ankle systems in lower-limb orthoses, prostheses, and exoskeletons.** *Medical engineering & physics*, **34**(4):397–408, 2012. 6
- [22] NIKOLA ŠERBEDŽIJA AND GILBERT BEYER. **Reflective Assistance-Pervasive Adaptation in Real Life Computing.** *Procedia Computer Science*, **7**:297–300, 2011. 6
- [23] IOANNIS MAGNISALIS, STAVROS DEMETRIADIS, AND ANASTASIOS KARAKOSTAS. **Adaptive and intelligent systems for collaborative learning support: A review of the field.** *Learning Technologies, IEEE Transactions on*, **4**(1):5–20, 2011. 6
- [24] SANG-SIK AHN AND YOON-CHANG LEE. **An improved statistical model-based VAD algorithm with an adaptive threshold.** *Journal of the Chinese Institute of Engineers*, **29**(5):783–789, 2006. 6

- [25] MARYAM RAVAN AND SOOSAN BEHESHTI. **Speech recognition from adaptive windowing PSD estimation.** In *Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on*, pages 524–527. IEEE, 2011. 6
- [26] CHIH-WEN KAN, ANDY Y LEE, LINDA T NIEMAN, KONSTANTIN SOKOLOV, AND MIA K MARKEY. **Adaptive spectral window sizes for extraction of diagnostic features from optical spectra.** *Journal of biomedical optics*, 15(4), 2010. 7
- [27] M. ROS, M.P. CUAR, M. DELGADO, AND A. VILA. **Online recognition of human activities and adaptation to habit changes by means of learning automata and fuzzy temporal windows.** *Information Sciences*, 220(0):86–101, 2013. **Online Fuzzy Machine Learning and Data Mining.** 7
- [28] ARCHAN MISRA AND LIPYEOW LIM. **Optimizing sensor data acquisition for energy-efficient smartphone-based continuous event processing.** In *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*, 1, pages 88–97. IEEE, 2011. 7
- [29] KIRAN K RACHURI, CECILIA MASCOLO, AND MIRCO MUSOLESI. **Energy-Accuracy Trade-offs of Sensor Sampling in Smart Phone Based Sensing Systems.** In *Mobile Context Awareness*, pages 65–76. Springer, 2012. 7
- [30] PEKKA SIIRTOLA AND JUHA RÖNING. **User-Independent Human Activity Recognition Using a Mobile Phone: Offline Recognition vs. Real-Time on Device Recognition.** In *Distributed Computing and Artificial Intelligence*, pages 617–627. Springer, 2012. 7
- [31] HIROTAKA HACHIYA, MASASHI SUGIYAMA, AND NAONORI UEDA. **Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition.** *Neurocomputing*, 80:93–101, 2012. 7
- [32] DAVIDE ANGUITA, ALESSANDRO GHIO, LUCA ONETO, XAVIER PARRA, AND JORGE L REYES-ORTIZ. **Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine.** In *Ambient Assisted Living and Home Care*, pages 216–223. Springer, 2012. 7
- [33] P AUGUSTYNIAK. **Distance Measures in Behavioral Pattern Analysis.** In *5th European Conference of the International Federation for Medical and Biological Engineering*, pages 894–897. Springer, 2011. 7
- [34] JC LÖTTTERS, J SCHIPPER, PH VELTINK, W OLTUIS, AND P BERGVELD. **Procedure for in-use calibration of triaxial accelerometers in medical applications.** *Sensors and Actuators A: Physical*, 68(1):221–228, 1998. 7
- [35] JAMES R LACKNER AND PAUL A DIZIO. **Aspects of body self-calibration.** *Trends in cognitive sciences*, 4(7):279–288, 2000. 7
- [36] PETER STURM. **Self-calibration of a moving zoom-lens camera by pre-calibration.** *Image and Vision Computing*, 15(8):583–589, 1997. 8
- [37] FERRAN ESPUNY. **On the Global Self-calibration of Central Cameras using Two Infinitesimal Rotations.** In *Proceedings of the 11th Asian Conference on Computer Vision-ACCV*, 2012. 8
- [38] HANNO ACKERMANN, BODO ROSENHAHN, FR SCHMIDT, HANNO ACKERMANN, BODO ROSENHAHN, NILS HASLER, THORSTEN THORMÄHLEN, HANS-PETER SEIDEL, KENICHI KANATANI, AND YASUYUKI SUGAYA. **Non-Rigid Self-Calibration Of A Projective Camera.** In *Proceedings of the 11th Asian Conference on Computer Vision (ACCV)*, 2012. 8
- [39] JING JIN AND XIAOFENG LI. **Efficient camera self-calibration method based on the absolute dual quadric.** *JOSA A*, 30(3):287–292, 2013. 8
- [40] FAHAD JAVED, MALIK TAHIR HASSAN, KHURUM NAZIR JUNEJO, NAVEED ARSHAD, AND ASIM KARIM. **Self-Calibration: Enabling Self-Management in Autonomous Systems by Preserving Model Fidelity.** In *Engineering of Complex Computer Systems (ICECCS), 2012 17th International Conference on*, pages 308–317. IEEE, 2012. 8
- [41] JI WEI ZHANG, XIAO DONG XU, AND BO WANG. **Multi-Position Self-Calibration Method of Inertial Navigation System.** *Advanced Materials Research*, 580:146–150, 2012. 8
- [42] OSCAR S VAN DE VEN, RUIMIN YANG, SHA XIA, JEROEN P VAN SCHIEVEEN, JO W SPRONCK, ROBERT H MUNNIG SCHMIDT, AND STOYAN NIHTANOV. **Autonomous self-aligning and self-calibrating capacitive sensor system.** In *Autonomous and Intelligent Systems*, pages 10–17. Springer, 2012. 8
- [43] ANDREA BARTOLINI, MATTEO CACCIARI, ANDREA TILLI, AND LUCA BENINI. **Thermal and energy management of high-performance multicores: Distributed and self-calibrating model-predictive controller.** 2013. 8
- [44] MATTHIAS GIETZELT, K-H WOLF, MICHAEL MARSSCHOLLEK, AND REINHOLD HAUX. **Automatic self-calibration of body worn triaxial-accelerometers for application in healthcare.** In *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on*, pages 177–180. IEEE, 2008. 8
- [45] KILIAN FORSTER, DANIEL ROGGEN, AND GERHARD TROSTER. **Un-supervised classifier self-calibration through repeated context occurrences: is there robustness against sensor displacement to gain?** In *Wearable Computers, 2009. ISWC'09. International Symposium on*, pages 77–84. IEEE, 2009. 8, 22
- [46] LEI GAO, ALAN K BOURKE, AND JOHN NELSON. **A system for activity recognition using multi-sensor fusion.** In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 7869–7872. IEEE, 2011. 8
- [47] DANIEL ROGGEN. **Designing and sharing activity recognition systems across platforms: methods from wearable computing.** 2011. 8
- [48] ORESTI BANOS, MIGUEL DAMAS, HECTOR POMARES, AND IGNACIO ROJAS. **On the Use of Sensor Fusion to Reduce the Impact of Rotational and Additive Noise in Human Activity Recognition.** *Sensors*, 12(6):8039–8054, 2012. 8
- [49] RALF KLINKENBERG AND THORSTEN JOACHIMS. **Detecting concept drift with support vector machines.** In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML)*, 11. Citeseer, 2000. 8
- [50] MOAMAR SAYED-MOUCHAWEH, JANAN ZAYTOON, AND PATRICE BILLAUDEL. **Adaptive Time Window Size to Track Concept Drift.** In *Machine Learning and Applications and Workshops (ICMLA), 2011 10th International Conference on*, 2, pages 41–46. IEEE, 2011. 9
- [51] ARTUR ŠILIĆ AND BOJANA DALBELO BAŠIĆ. **Exploring classification concept drift on a large news text corpus.** In *Computational Linguistics and Intelligent Text Processing*, pages 428–437. Springer, 2012. 9
- [52] QUN ZHU, XUEGANG HU, YUHONG ZHANG, PEIPEI LI, AND XINDONG WU. **A double-window-based classification algorithm for concept drifting data streams.** In *Granular Computing (GrC), 2010 IEEE International Conference on*, pages 639–644. IEEE, 2010. 9

- [53] ALBERT BIFET AND RICARD GAVALDA. **Learning from time-changing data with adaptive windowing.** In *SIAM International Conference on Data Mining*, pages 443–448, 2007. 9
- [54] TAO GU, ZHANQING WU, XIANPING TAO, HUNG KENG PUNG, AND JIAN LU. **epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition.** In *Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on*, pages 1–9. IEEE, 2009. 9
- [55] MINA MEMAR, MAHMOOD DEYPIR, MOHAMMAD HADI SADREDDINI, AND SEYYED MOSTAFA FAKHRAHMAD. **An Efficient Frequent Itemset Mining Method over High-speed Data Streams.** *The Computer Journal*, 55(11):1357–1366, 2012. 9
- [56] JAN PETER PATIST. **Optimal window change detection.** In *Data Mining Workshops, 2007. ICDM Workshops 2007. Seventh IEEE International Conference on*, pages 557–562. IEEE, 2007. 9
- [57] OLIVER AMFT AND GERHARD TRÖSTER. **Recognition of dietary activity events using on-body sensors.** *Artificial Intelligence in Medicine*, 42(2):121–136, 2008. 9
- [58] NARAYANAN C KRISHNAN AND DIANE J COOK. **Activity recognition on streaming sensor data.** *Pervasive and Mobile Computing*, 2012. 9
- [59] GEORGE OKEYO, LIMING CHEN, HUI WANG, AND ROY STERRITT. **Dynamic sensor data segmentation for real-time knowledge-driven activity recognition.** *Pervasive and Mobile Computing*, (0), 2012. 9
- [60] JANI MANTYJARVI, JOHAN HIMBERG, AND TAPIO SEPPANEN. **Recognizing human motion with multiple acceleration sensors.** In *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, 2, pages 747–752. IEEE, 2001. 10, 17, 86
- [61] NICKY KERN, BERNT SCHIELE, AND ALBRECHT SCHMIDT. **Multi-sensor activity context detection for wearable computing.** In *Ambient Intelligence*, pages 220–232. Springer, 2003. 10, 86
- [62] ANDREAS KRAUSE, DANIEL P SIEWIOREK, ASIM SMALAGIC, AND JONNY FARRINGTON. **Unsupervised, Dynamic Identification of Physiological and Activity Context in Wearable Computing.** In *ISWC*, 3, page 88, 2003. 10, 86
- [63] NISHKAM RAVI, NIKHIL DANDEKAR, PREETHAM MYSORE, AND MICHAEL L LITTMAN. **Activity recognition from accelerometer data.** In *AAAI*, pages 1541–1546, 2005. 10, 21, 22, 74, 76, 86
- [64] TÂM HUYNH AND BERNT SCHIELE. **Analyzing features for activity recognition.** In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, pages 159–163. ACM, 2005. 10, 21, 62, 74, 86
- [65] JAMIE A WARD, PAUL LUKOWICZ, GERHARD TROSTER, AND THAD E STARNER. **Activity recognition of assembly tasks using body-worn microphones and accelerometers.** *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1553–1567, 2006. 11, 86
- [66] SUSANNA PIRTIKANGAS, KAORI FUJINAMI, AND TATSUO NAKAJIMA. **Feature selection and activity recognition from wearable sensors.** In *Ubiquitous Computing Systems*, pages 516–527. Springer, 2006. 11
- [67] LIN SUN, DAQING ZHANG, BIN LI, BIN GUO, AND SHIJIAN LI. **Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations.** In *Ubiquitous intelligence and computing*, pages 548–562. Springer, 2010. 11, 21, 26
- [68] OLIVER AMFT AND PAUL LUKOWICZ. **From backpacks to smartphones: past, present, and future of wearable computers.** *Pervasive Computing, IEEE*, 8(3):8–13, 2009. 14, 91
- [69] STEPHEN J PREECE, JOHN Y GOULERMAS, LAURENCE PJ KENNEY, DAVE HOWARD, KENNETH MELJER, AND ROBIN CROMPTON. **Activity identification using body-mounted sensors: a review of classification techniques.** *Physiological measurement*, 30(4):R1, 2009. 16
- [70] RAJEEV SHARMA, VLADIMIR I PAVLOVIC, AND THOMAS S HUANG. **Toward multimodal human-computer interface.** *Proceedings of the IEEE*, 86(5):853–869, 1998. 17
- [71] ORESTI BANOS, MIGUEL DAMAS, HECTOR POMARES, FERNANDO ROJAS, BLANCA DELGADO-MARQUEZ, AND OLGA VALENZUELA. **Human activity recognition based on a sensor weighting hierarchical classifier.** *Soft Computing*, 17(2):333–343, 2013. 17, 23, 69, 93
- [72] JONATHAN LESTER, TANZEEM CHOUDHURY, NICKY KERN, GAETANO BORRIELLO, AND BLAKE HANNAFORD. **A Hybrid Discriminative/Generative Approach for Modeling Human Activities.** In *IJCAI*, 5, pages 766–772, 2005. 21
- [73] JONATHAN LESTER, TANZEEM CHOUDHURY, AND GAETANO BORRIELLO. **A practical approach to recognizing physical activities.** In *Pervasive Computing*, pages 1–16. Springer, 2006. 21
- [74] JIAHUI WU, GANG PAN, DAQING ZHANG, GUANDE QI, AND SHIJIAN LI. **Gesture recognition with a 3-d accelerometer.** In *Ubiquitous intelligence and computing*, pages 25–38. Springer, 2009. 21
- [75] JUN YANG. **Toward physical activity diary: motion recognition using simple acceleration features with mobile phones.** In *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics*, pages 1–10. ACM, 2009. 21, 22
- [76] RÓ DUDA. **PE hart and DG Stork.** *Pattern Classification*, 2000. 21
- [77] S THEODORIDIS AND K KOUTROUMBAS. **Pattern Recognition 4th edn (London: Academic).** 2008. 22
- [78] THOMAS COVER AND PETER HART. **Nearest neighbor pattern classification.** *Information Theory, IEEE Transactions on*, 13(1):21–27, 1967. 22
- [79] DAVID MINNEN, TRACY WESTEYN, THAD STARNER, J WARD, AND PAUL LUKOWICZ. **Performance metrics and evaluation issues for continuous activity recognition.** *Performance Metrics for Intelligent Systems*, page 4, 2006. 26
- [80] EMMANUEL MUNGUIA TAPIA, STEPHEN S INTILLE, WILLIAM HASKELL, KENT LARSON, JULIE WRIGHT, ABBY KING, AND ROBERT FRIEDMAN. **Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor.** In *Wearable Computers, 2007 11th IEEE International Symposium on*, pages 37–40. IEEE, 2007. 26
- [81] QUANFU FAN, RUSSELL BOBBITT, YUN ZHAI, AKIRA YANAGAWA, SHARATH PANKANTI, AND ARUN HAMPAPUR. **Recognition of repetitive sequential human activity.** In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 943–950. IEEE, 2009. 26

- [82] HENRI BOUMA, PATRICK HANCKMANN, JAN-WILLEM MARCK, LEO PENNING, RICHARD DEN HOLLANDER, JOHAN-MARTIJN TEN HOVE, SEBASTIAAN VAN DEN BROEK, KLAMER SCHUTTE, AND GERTJAN BURGHOUTS. **Automatic human action recognition in a scene from visual inputs.** In *SPIE Defense, Security, and Sensing*, pages 83880L–83880L. International Society for Optics and Photonics, 2012. 26
- [83] MARINA SOKOLOVA, NATHALIE JAPKOWICZ, AND STAN SZPAKOWICZ. **Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation.** In *AI 2006: Advances in Artificial Intelligence*, pages 1015–1021. Springer, 2006. 27, 92
- [84] MARK DREDZE, TESSA LAU, AND NICHOLAS KUSHMERICK. **Automatically classifying emails into activities.** In *Proceedings of the 11th international conference on Intelligent user interfaces*, pages 70–77. ACM, 2006. 27
- [85] PIERO ZAPPI, THOMAS STIEFMEIER, ELISABETTA FARELLA, DANIEL ROGEN, LUCA BENINI, AND G TROSTER. **Activity recognition from on-body sensors by classifier fusion: sensor scalability and robustness.** In *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, pages 281–286. IEEE, 2007. 70
- [86] RICHARD W DEVAUL AND STEVE DUNN. **Real-time motion classification for wearable computing applications.** 2001, project paper, <http://www.media.mit.edu/wearables/mithril/realtime.pdf>, 2001. 93
- [87] KRISTOF VAN LAERHOVEN AND OZAN CAKMAKCI. **What shall we teach our pants?** In *Wearable Computers, The Fourth International Symposium on*, pages 77–83. IEEE, 2000. 93