

**CENTRO POLITÉCNICO SUPERIOR
UNIVERSIDAD DE ZARAGOZA**

**CLASIFICADOR AUTOMÁTICO
DE LATIDOS BASADO EN LAS
CARACTERÍSTICAS DE LA
TRANSFORMADA WAVELET
DEL ECG**

PAULA ROMEO SÁNCHEZ

2 de Septiembre del 2011

**DIRECTOR: JUAN PABLO MARTINEZ CORTES
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA Y COMUNICACIONES
INGENIERÍA DE TELECOMUNICACIONES
ESPECIALIDAD COMUNICACIONES
CURSO 2010-2011**

*A Manuel y Gloria,
Por haber sido el mejor ejemplo de fortaleza, superación y generosidad.
Gracias por haberme dado todo lo que tengo y enseñado a ser lo que ahora soy.
Nunca podré demostraros todo mi agradecimiento.*

CLASIFICADOR AUTOMÁTICO DE LATIDOS BASADO EN CARACTERÍSTICAS DE LA TRANSFORMADA WAVELET DEL ECG

RESUMEN

En este proyecto se ha implementado un clasificador automático de latidos en la señal electrocardiográfica (ECG), cuyas características destacables son su sencillez y el empleo de la transformada wavelet, que nos permite hacer uso de un delineador ya creado, basado en dicha transformada. De este modo, el mismo procesado ya usado en la delineación permite obtener las características que usamos para la clasificación. Además su utilización viene justificada porque proporciona una descripción de la señal en el dominio del tiempo, permitiendo la representación de las características temporales de ésta a distintas escalas, lo cual resulta muy adecuado para analizar señales ECG, que se caracterizan por una ocurrencia cíclica de ondas con diferentes contenidos frecuenciales.

Para entrenar y evaluar el clasificador contamos con la base de datos MIT-BIH Arrhythmia database [6], que contiene 44 registros, cada uno de los cuales supone alrededor de una media hora de captura de latidos. Ésta base la dividimos en dos partes, de las cuales una es utilizada para entrenar al clasificador, y la otra para la evaluación de los resultados.

La clasificación se hace en base a una serie de características obtenidas de la señal ECG y su transformada wavelet a diferentes escalas (características temporales y de amplitud), así como información del ritmo cardíaco instantáneo. Se consideran así mismo características de cada una de las dos derivaciones registradas y características conjuntas.

Para la clasificación se emplea un clasificador sencillo de tipo Bayesiano, que se aplica sobre distintos conjuntos de características (configuraciones) para determinar aquellas que mejor discriminan entre las clases de latidos. Se considera la clasificación de latidos recomendada por la AAMI (Administration for the Advance of Medical Instrumentation): normales, ectópicos, supraventriculares, ectópicos ventriculares, fusión de normales y ventriculares y latidos marcapaseados.

ÍNDICE GENERAL

<u>DESCRIPCIÓN</u>	<u>PÁG.</u>
1. INTRODUCCIÓN	7
2. OBJETIVO	9
3. MATERIALES	10
4. METODOLOGÍA	11
5. PREPROCESADO Y OBTENCIÓN DE LAS CARACTERÍSTICAS DE LOS LATIDOS	12
5.1 DELINEACIÓN	12
5.2 TRANSFORMADA WAVELET DE LA SEÑAL	12
5.3 FUNCIONES DE COVARIANZA DE LA SEÑAL	15
5.4 CARACTERÍSTICAS DE LA SEÑAL	19
5.4.1 Características temporales y de amplitud.....	19
5.4.2 Características del ritmo de la señal	20
5.4.3 Tipos de latido.....	21
5.4.4 Matriz de características	21
6. DEFINICIÓN DEL CLASIFICADOR	24
6.1 CONJUNTOS DE CARACTERÍSTICAS	24
6.2 ESTRUCTURA DEL CLASIFICADOR	25
6.2.1 Clasificador cuadrático	26
6.2.2 Clasificador lineal	27
6.3 ENTRENAMIENTO	27
6.4 EVALUACIÓN DE PRESTACIONES.....	28
7. RESULTADOS	29
7.1 RESULTADOS DEL CLASIFICADOR CUADRÁTICO	31
7.2 RESULTADOS DEL CLASIFICADOR LINEAL	33
7.3 ELECCIÓN DE CONFIGURACIÓN O CLASIFICADOR	35
8. REDUCCIÓN DE CARACTERÍSTICAS	36
8.1 REDUCCIÓN DE LA CONFIGURACIÓN FS4_D3.....	37
8.2 REDUCCIÓN DE LA CONFIGURACIÓN FS4_ECG.....	43
9. VALIDACIÓN-CRUZADA	47
10. COMPARATIVA CON TRABAJOS ANTERIORES	52
11. CONCLUSIONES	53
12. BIBLIOGRAFÍA	55
13. ANEXOS	56
13.1 ANEXO1: RESULTADOS DEL CLASIFICADOR FS1_A1.....	56
13.2 ANEXO2: RESULTADOS DEL CLASIFICADOR FS1_A2.....	56

13.3 ANEXO3: RESULTADOS DEL CLASIFICADOR FS1_A3.....	56
13.4 ANEXO4: RESULTADOS DEL CLASIFICADOR FS1_A4.....	57
13.5 ANEXO5: RESULTADOS DEL CLASIFICADOR FS1_A5.....	57
13.6 ANEXO6: RESULTADOS DEL CLASIFICADOR FS1_D1.....	57
13.7 ANEXO7: RESULTADOS DEL CLASIFICADOR FS1_D2.....	58
13.8 ANEXO8: RESULTADOS DEL CLASIFICADOR FS1_D3.....	58
13.9 ANEXO9: RESULTADOS DEL CLASIFICADOR FS1_D4.....	58
13.10 ANEXO10: RESULTADOS DEL CLASIFICADOR FS1_D5.....	59
13.11 ANEXO11: RESULTADOS DEL CLASIFICADOR FS1_ECG.....	59
13.12 ANEXO 12: RESULTADOS DEL CLASIFICADOR FS2_A1	59
13.13 ANEXO 13: RESULTADOS DEL CLASIFICADOR FS2_A2.....	60
13.14 ANEXO 14: RESULTADOS DEL CLASIFICADOR FS2_A3.....	60
13.15 ANEXO 15: RESULTADOS DEL CLASIFICADOR FS2_A4	60
13.16 ANEXO 16: RESULTADOS DEL CLASIFICADOR FS2_A5.....	61
13.17 ANEXO 17: RESULTADOS DEL CLASIFICADOR FS2_D1.....	61
13.18 ANEXO 18: RESULTADOS DEL CLASIFICADOR FS2_D2.....	61
13.19 ANEXO 19: RESULTADOS DEL CLASIFICADOR FS2_D3.....	62
13.20 ANEXO 20: RESULTADOS DEL CLASIFICADOR FS2_D4.....	62
13.21 ANEXO 21: RESULTADOS DEL CLASIFICADOR FS2_D5.....	62
13.22 ANEXO 22: RESULTADOS DEL CLASIFICADOR FS2_ECG.....	63
13.23 ANEXO 23: RESULTADOS DEL CLASIFICADOR FS3_A1	63
13.24 ANEXO 24: RESULTADOS DEL CLASIFICADOR FS3_A2.....	63
13.25 ANEXO 25: RESULTADOS DEL CLASIFICADOR FS3_A3.....	64
13.26 ANEXO 26: RESULTADOS DEL CLASIFICADOR FS3_A4.....	64
13.27 ANEXO 27: RESULTADOS DEL CLASIFICADOR FS3_A5.....	64
13.28 ANEXO 28: RESULTADOS DEL CLASIFICADOR FS3_D1.....	65
13.29 ANEXO 29: RESULTADOS DEL CLASIFICADOR FS3_D2.....	65
13.30 ANEXO 30: RESULTADOS DEL CLASIFICADOR FS3_D3.....	65
13.31 ANEXO 31: RESULTADOS DEL CLASIFICADOR FS3_D4.....	66
13.32 ANEXO 32: RESULTADOS DEL CLASIFICADOR FS3_D5.....	66
13.33 ANEXO 33: RESULTADOS DEL CLASIFICADOR FS3_ECG.....	66
13.34 ANEXO 34: RESULTADOS DEL CLASIFICADOR FS4_A1	67
13.35 ANEXO 35: RESULTADOS DEL CLASIFICADOR FS4_A2.....	67
13.36 ANEXO 36: RESULTADOS DEL CLASIFICADOR FS4_A3.....	67
13.37 ANEXO 37: RESULTADOS DEL CLASIFICADOR FS4_A4.....	68
13.38 ANEXO 38: RESULTADOS DEL CLASIFICADOR FS4_A5.....	68
13.39 ANEXO 39: RESULTADOS DEL CLASIFICADOR FS4_D1.....	68
13.40 ANEXO 40: RESULTADOS DEL CLASIFICADOR FS4_D2.....	69
13.41 ANEXO 41: RESULTADOS DEL CLASIFICADOR FS4_D3.....	69
13.42 ANEXO 42: RESULTADOS DEL CLASIFICADOR FS4_D4.....	69

13.43 ANEXO 43: RESULTADOS DEL CLASIFICADOR FS4_D5.....	70
13.44 ANEXO 44: RESULTADOS DEL CLASIFICADOR FS4_ECG.....	70

ÍNDICE DE FIGURAS Y TABLAS

<u>DESCRIPCIÓN</u>	<u>PÁG.</u>
FIGURAS	
Figura 1-1 Esquema de un sistema de análisis clínico basado en ECG [2]	8
Figura 5-1 Banco de Filtros	13
Figura 5-2 Filtros 'P(w)' a diferentes escalas con frecuencia de muestreo 360 Hz	14
Figura 5-3 Señal ECG y sus aproximaciones de distintos órdenes	15
Figura 5-4 Superposición de la función de autocovarianza de los latidos de la señal ECG separada por tipos (derivación 1)	16
Figura 5-5 Superposición de la función de autocovarianza de los latidos de la aproximación de orden 2 de la señal ECG separada por tipos (derivación 1)	17
Figura 5-6 Superposición de la función de autocovarianza de los latidos del Detalle de orden 1 de la señal ECG separada por tipos (derivación 1)	17
Figura 5-7 Superposición de la función de autocovarianza de los latidos del Detalle de orden 5 de la señal ECG separada por tipos (derivación 1)	18
Figura 5-8 Forma de la onda QRS	20
TABLAS	
Tabla 1 Matriz de confusión	29
Tabla 2 Resultados de los clasificadores cuadráticos	33
Tabla 3 Resultados de los clasificadores lineales	34
Tabla 4 Tabla de resultados de la validación cruzada por registros y agregados de la configuración FS4_D3_Reducida	48
Tabla 5 Matriz de confusión de la validación cruzada de la configuración FS4_D3_Reducida	48
Tabla 6 Resultados finales de la configuración FS4_D3_Reducida	49
Tabla 7 Tabla de resultados de la validación cruzada por registros y agregados de la configuración FS4_ECG_Reducida	50
Tabla 8 Matriz de confusión de la configuración FS4_ECG_Reducida	50
Tabla 9 Resultados finales de la configuración FS4_ECG_Reducida	51
Tabla 10 Tabla comparativa de resultados con trabajos anteriores	52

1. INTRODUCCIÓN

Las enfermedades cardiovasculares constituyen una de las primeras causas de mortalidad en los países occidentales. Las arritmias son el resultado de alteraciones en el ritmo, la regularidad y en el lugar de origen o conducción de los impulsos eléctricos del corazón. A rasgos generales, pueden dividirse en dos grupos: el primero incluye las taquicardias y las fibrilaciones ventriculares, que suponen una grave amenaza para la vida, por lo que requieren una terapia inmediata con desfibrilador. Para estos casos ya han sido diseñados detectores con una alta sensibilidad y especificación. El segundo grupo lo componen las arritmias que no suponen una amenaza inminente, pero que igualmente requieren una detección y, en función del diagnóstico, pueden requerir terapia para evitar problemas posteriores [1].

Este estudio se va a centrar en la detección del segundo tipo y para ello nos basaremos en el análisis del electrocardiograma (ECG), que resulta un método muy apropiado puesto que nos permite hacer un análisis del ritmo cardiaco y de sus anomalías de una forma barata, sencilla y no invasiva.

Un paso importante para la identificación de arritmias a través de ECG es hacer una buena clasificación de los diferentes tipos de latidos ya que conociendo la clasificación de latidos consecutivos del corazón, es posible determinar el ritmo de la señal ECG. Sin embargo, la clasificación de latidos puede ser un proceso largo y tedioso, por lo que cualquier automatización de este proceso resulta de gran ayuda [1].

El propósito de este proyecto es realizar una clasificación automática de latidos del corazón para el diagnóstico de enfermedades cardiacas a través de ECG. Los métodos y objetivos a perseguir se explicarán con más detalle posteriormente.

Un sistema típico de análisis clínico basado en ECG comprende un dispositivo de adquisición de datos, seguido de una etapa de eliminación de ruido. La siguiente fase es la de delineación, que nos permite detectar las diferentes partes de la señal ECG como el complejo QRS, las ondas P y T así como sus puntos significativos (inicios, picos y finales). Y por último, la etapa de clasificación de latidos, en la que nos centraremos en este proyecto. [2]

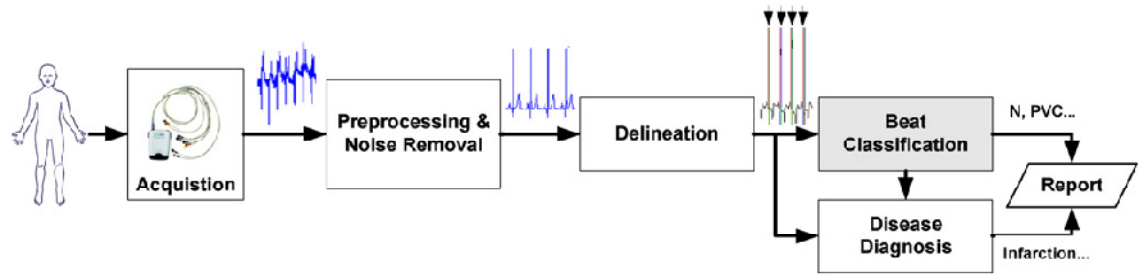


Figura 1-1 Esquema de un sistema de análisis clínico basado en ECG [2]

2. **OBJETIVO**

Actualmente existen ya numerosos estudios de clasificación automática de latidos basados en diferentes métodos y analizando características diferentes de la señal ECG. El objetivo de este proyecto es diseñar y evaluar un clasificador sencillo y robusto que permita emplear características que se puedan obtener de forma simple a partir de la delineación.

La peculiaridad que presenta el clasificador automático que se va a implementar es que está basado en la transformada wavelet, lo que nos permitirá hacer uso de un delineador ya creado basado en dicha transformada. De este modo, el mismo procesado ya usado en la delineación, servirá para obtener las características que usaremos para la clasificación, aportando así sencillez al clasificador.

Además, el uso de la transformada wavelet resulta muy adecuado para analizar señales ECG puesto que dichas señales se caracterizan por una ocurrencia cíclica de ondas con diferentes contenidos frecuenciales y la transformada wavelet nos proporciona una descripción de la señal en el dominio del tiempo, permitiendo la representación de las características temporales de ésta a diferentes resoluciones.

3. MATERIALES

Para llevar a cabo este proyecto, contaremos con la base de datos MIT-BIH Arrhythmia database [6], que contiene 44 registros, cada uno de los cuales supone alrededor de una media de hora de captura de latidos y contiene información de 2 derivaciones de la señal: derivación 1 y derivación 2.

La base de datos incluye también una anotación manual de los latidos realizada por expertos cardiólogos, que incluye la posición de cada latido y el tipo de latido. En este trabajo utilizamos las clases recomendadas por la AAMI (Administration for the Advance of Medical Instrumentation), que diferencia entre 5 tipos: N o latido normal, S o latido ectópico supraventricular, V o latido ectópico ventricular, F o fusión de latido normal y ventricular y Q o latidos desconocidos o marcapaseados.

Para realizar la clasificación de latidos en base a las características obtenidas, se hará uso de un clasificador sencillo de tipo Bayesiano, para determinar aquellas que mejor discriminan entre las 5 clases de latidos.

Todos estos procesos de extracción de características y clasificación se harán mediante programación de funciones y procedimientos en Matlab.

4. METODOLOGÍA

Dividiremos los registros de la base de datos en dos partes de 22 registros cada una: la primera será la de entrenamiento (registros: 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 y 230) y otra la de test (registros: 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232 y 234). Esta división ya ha sido empleada en trabajos previos [1], [5], y se adopta en este trabajo para facilitar la comparación.

Con la primera, que llamaremos de entrenamiento, enseñaremos al clasificador. Para ello obtendremos una serie de características extraídas a partir de la señal ECG y de su transformada wavelet a diferentes escalas (características temporales y de amplitud), así como información del ritmo cardíaco instantáneo (distancia entre puntos R de las ondas del ECG). Se considerarán así mismo características de cada una de las dos derivaciones registradas (derivación 1 y derivación 2) y características conjuntas.

Determinaremos 4 configuraciones o agrupaciones de características diferentes que nos permitan discriminar entre los distintos tipos de latido. Estas características se aplicarán tanto a la señal ECG registrada como las 10 señales de aproximación y detalle resultantes de pasar la señal por el banco de filtros equivalente de la transformada wavelet discreta (en las 5 primeras escalas). Por tanto contaremos con 4 configuraciones de características en cada uno de los 11 tipos de señal, lo que hace 44 configuraciones distintas para analizar (y por tanto 44 posibles clasificadores). Tras la fase de test y análisis de resultados numéricos, nos quedaremos de entre todas ellas, con la que mejor cumpla nuestro objetivo.

Una vez que tenemos las 44 configuraciones distintas de características que nos han permitido construir el sistema de clasificación, aplicaremos el clasificador obtenido a la parte de test de la base de datos, consiguiendo así haber mantenido estos datos fuera del proceso de creación del clasificador, para evitar resultados sesgados. Del resultado de aplicar los 44 tipos de clasificador diferentes a los nuevos registros, obtendremos resultados numéricos que analizaremos y plasmaremos en tablas, para finalmente seleccionar nuestro clasificador definitivo.

5. PREPROCESADO Y OBTENCIÓN DE LAS CARACTERÍSTICAS DE LOS LATIDOS

En este punto se va a explicar los pasos que se han seguido para la construcción de los 44 posibles clasificadores que vamos a considerar en este proyecto, de entre los que elegiremos posteriormente el que mejores resultados obtenga, para considerarlo nuestro clasificador definitivo.

5.1 DELINEACIÓN

Se recorren todos los registros (en sus dos derivaciones) de la base de latidos que tenemos, aplicando el delineador ya existente [3] que, como comentábamos en la introducción, está basado en la transformada wavelet.

Esto nos permite de forma sencilla y rápida, obtener la posición y las anotaciones de cada latido del registro, para poder centrarnos en su estudio y clasificación.

5.2 TRANSFORMADA WAVELET DE LA SEÑAL

Como ya se ha mencionado, el uso de la transformada wavelet resulta muy adecuado para analizar señales ECG puesto que dichas señales se caracterizan por una ocurrencia cíclica de ondas con diferentes contenidos frecuenciales y la transformada wavelet nos proporciona una descripción de la señal en el dominio del tiempo, permitiendo la representación de las características temporales de ésta a diferentes resoluciones [2]. Por este motivo, vamos a trabajar, además de con la señal ECG directamente, con su transformada.

Para la obtención de la transformada wavelet, pasamos nuestra señal ECG a través de un banco de filtros diádicos como el que se muestra a continuación (figura adaptación de [3]):

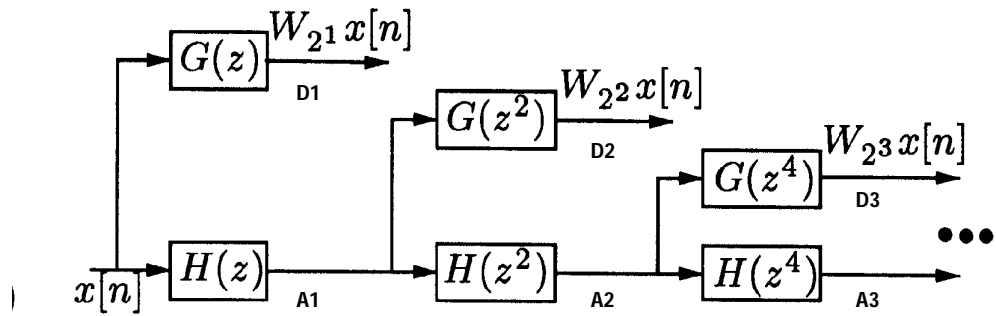


Figura 5-1 Banco de Filtros

En este trabajo, como ya se hacía en el delineador [3], se emplea como wavelet prototipo la derivada de una spline cuadrática, de forma que los filtros $H(z)$ son filtros FIR paso bajo cuya respuesta frecuencial viene dada por:

$$H(z) = H(e^{j\omega}) = e^{j\omega} \cdot \left(\cos\left(\frac{\omega}{2}\right)\right)^3 \quad (1)$$

Los filtros $G(z)$ son filtros FIR paso alto (en realidad, dada la wavelet utilizada, son aproximaciones a filtros derivadores) cuya respuesta frecuencial viene dada por:

$$G(z) = G(e^{j\omega}) = 4 \cdot j \cdot e^{j\omega} \cdot \left(\sin\left(\frac{\omega}{2}\right)\right) \quad (2)$$

Cuando la señal pasa únicamente por los filtros paso-bajo (filtros H), obtenemos lo que llamamos **Aproximaciones**, que son versiones paso bajo de la señal a distintas escalas. El orden o la escala de estas aproximaciones vendrá dado por el número de filtros por el que pase:

$$P1(\omega) = H(\omega) \rightarrow \text{Aproximación de orden 1 (A1)}$$

$$P2(\omega) = H(\omega)H(2\omega) \rightarrow \text{Aproximación de orden 2 (A2)}$$

$$P3(\omega) = H(\omega)H(2\omega)H(4\omega) \rightarrow \text{Aproximación de orden 3 (A3)}$$

$$P4(\omega) = H(\omega)H(2\omega)H(4\omega)H(8\omega) \rightarrow \text{Aproximación de orden 4 (A4)}$$

...

Cuando la señal pasa por filtros paso bajo (filtros H) y paso alto (filtros G), obtenemos los **Detalles**, que constituyen propiamente la transformada wavelet de la señal y para la wavelet utilizada, son versiones derivadas paso-bajo de la señal original. De forma análoga tendremos:

$$Q1(w) = G(w) \rightarrow \text{Detalle de orden 1 (D1)}$$

$$Q2(w) = H(w)G(2w) \rightarrow \text{Detalle de orden 2 (D2)}$$

$$Q3(w) = H(w)H(2w)G(4w) \rightarrow \text{Detalle de orden 3 (D3)}$$

$$Q4(w) = H(w)H(2w)H(4w)G(8w) \rightarrow \text{Detalle de orden 4 (D4)}$$

En nuestro caso trabajaremos con aproximaciones y detalles de la señal hasta el orden 5.

En las siguientes figuras se representan como ejemplo los filtros 'p' y 'q' en sus 5 primeras escalas, normalizados y alineados. La frecuencia de estos filtros es 360 Hz, que es la frecuencia de muestreo que tiene la base de datos del MIT-BIH.

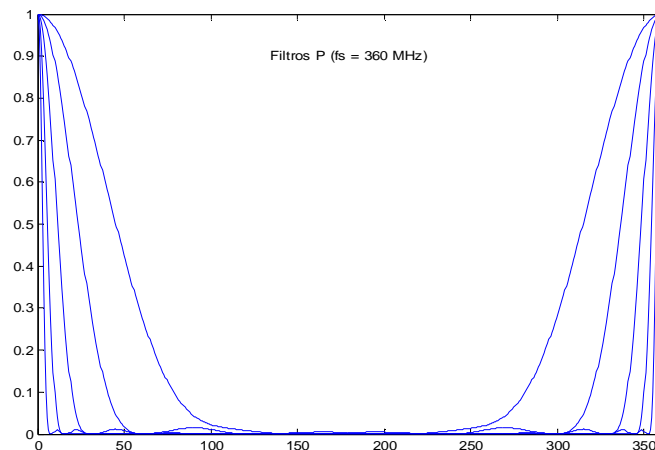


Figura 5-2 Filtros 'P(w)' a diferentes escalas con frecuencia de muestreo 360 Hz

La siguiente imagen muestra un ejemplo de cómo sería la apariencia de las aproximaciones de la señal ECG

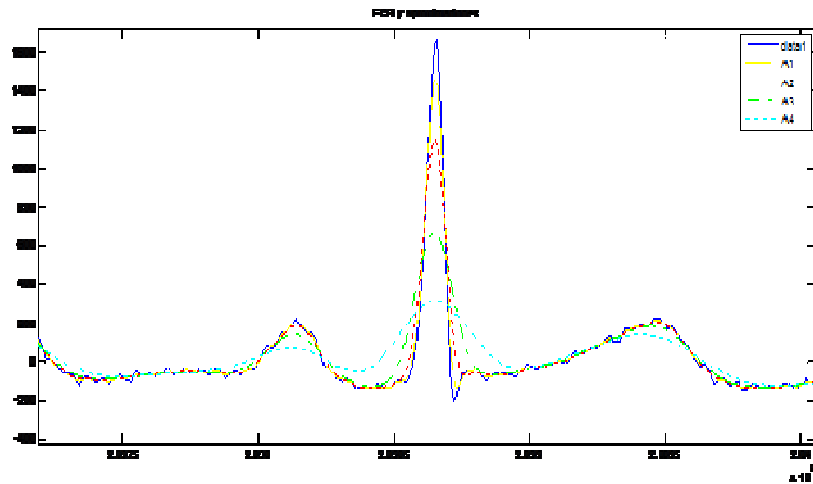


Figura 5-3 Señal ECG y sus aproximaciones de distintos órdenes

5.3 FUNCIONES DE COVARIANZA DE LA SEÑAL

En lugar de obtener las características directamente de la señal ECG y sus aproximaciones / detalles de distintas escalas, vamos a trabajar a partir de ahora con su función de autocovarianza, puesto que si trabajásemos con muestras de la señal las características serían muy dependientes de la referencia temporal y por tanto, cualquier error en la detección del QRS, afectaría a dicha característica. Por ejemplo, si pusiéramos un cruce por cero tendríamos que darlo con respecto al punto fiducial por lo que, como decíamos, un error en la detección de QRS afectará a la característica asociada.

La idea de hacer funciones de autocovarianza y covarianzas cruzadas (que son autocorrelaciones y correlaciones cruzadas de las señales una vez les quitamos la media), es que hacemos los características menos dependientes de las fases de detección de QRS y delineación y, dado que la autocorrelación por definición ya está sincronizada, podemos incluso considerar muestras de la autocorrelación como características.

Para hacernos una idea de la forma que adopta la covarianza de la señal en sus distintas expresiones, y de las diferencias de estas formas entre los distintos tipos de latidos, se muestra a continuación una representación gráfica de algunas de las expresiones de la señal, agrupando por separado los latidos de los

distintos tipos, que se muestran superpuestos. Para la agrupación de latidos por tipos, usamos la información de referencia de la base de datos, que contiene las etiquetas de clasificación de todos los latidos, que han sido anotados y etiquetados por expertos cardiólogos. Esto nos permitirá comprobar si el tipo que prediga finalmente nuestro clasificador, es correcto o no.

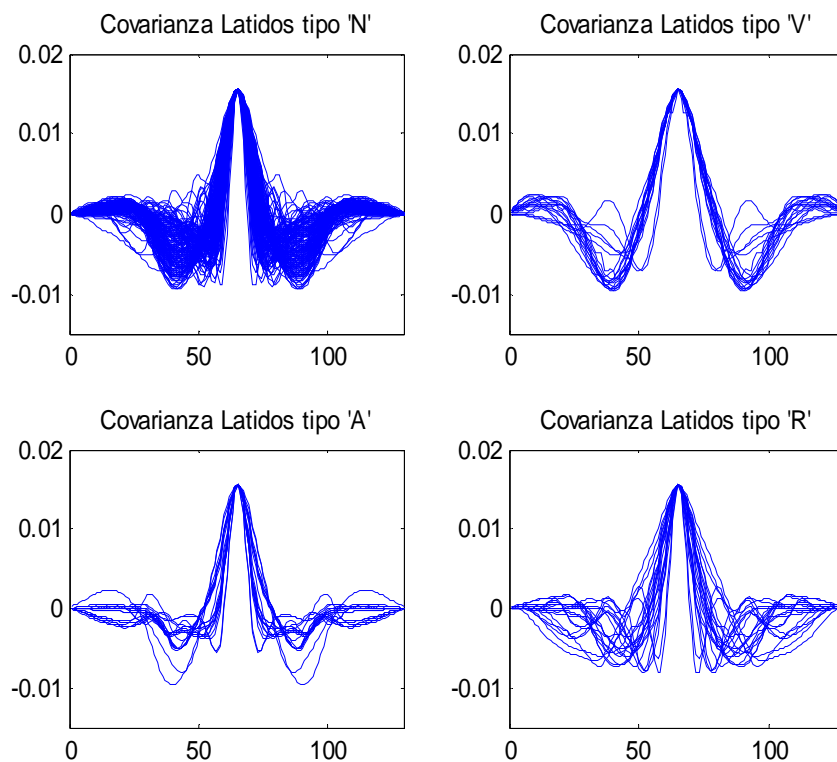


Figura 5-4 Superposición de la función de autocovarianza de los latidos de la señal ECG separada por tipos (derivación 1)

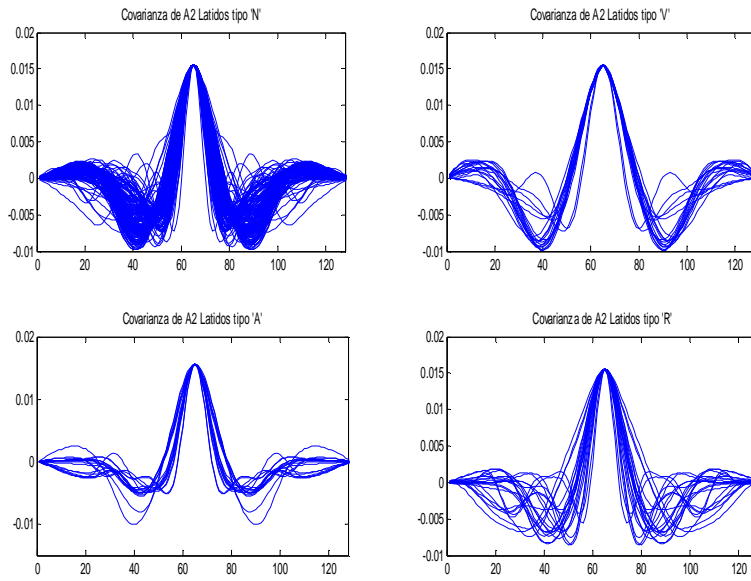


Figura 5-5 Superposición de la función de autocovarianza de los latidos de la aproximación de orden 2 de la señal ECG separada por tipos (derivación 1)

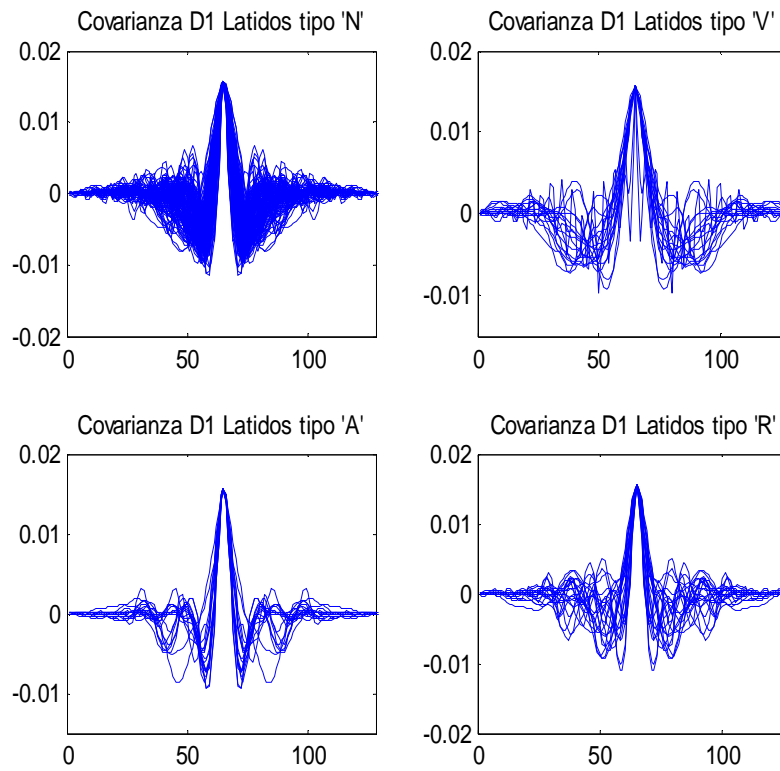


Figura 5-6 Superposición de la función de autocovarianza de los latidos del Detalle de orden 1 de la señal ECG separada por tipos (derivación 1)

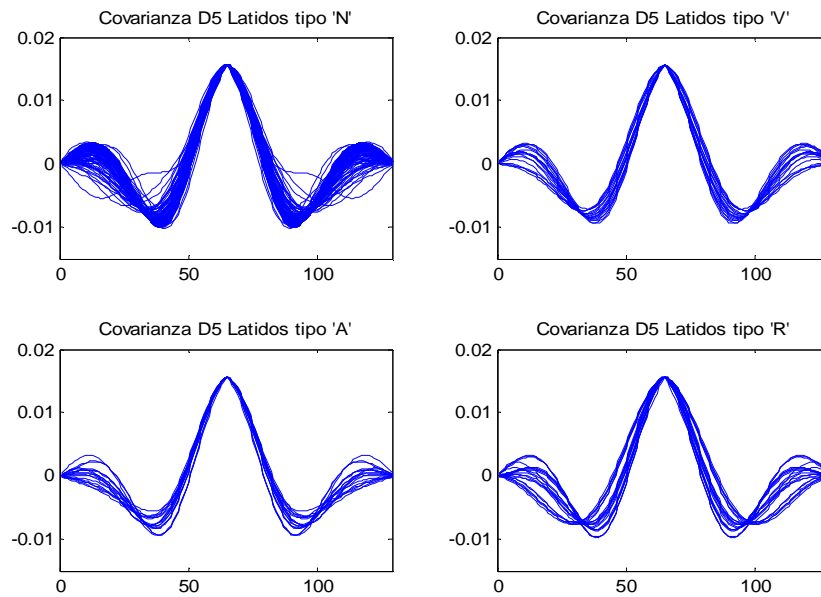


Figura 5-7 Superposición de la función de autocovarianza de los latidos del Detalle de orden 5 de la señal ECG separada por tipos (derivación 1)

Como ya hemos dicho, a partir de ahora trabajaremos con las funciones de autocovarianza y covarianza cruzada de la señal por lo que, creamos unas matrices que almacenarán esta información. Tendremos una matriz por cada registro y derivación del registro (derivación 1 y derivación 2). Estas matrices contendrán en cada columna la covarianza de uno de los latidos del registro, por tanto tendrán tantas columnas como latidos haya en dicho registro, y 65 filas que son el resultado de realizar la covarianza de 64 puntos.

También calcularemos para cada uno de los 44 registros la covarianza cruzada de la derivación 1 con la derivación 2 en cada una de las expresiones de la señal (A1, A2,...).

Por tanto, al final tendremos 3 matrices de covarianzas: covarianza de la derivación 1, covarianza de la derivación 2 y covarianza cruzada de la derivación 1 con la derivación 2, por cada uno de los 44 registros. En total $44 \times 3 = 132$ matrices con las covarianzas de todos los latidos.

5.4 CARACTERÍSTICAS DE LA SEÑAL

Llegado este punto ya tenemos almacenada la información de la covarianza y covarianza cruzada de todos los registros de la base (de sus dos derivaciones) en sus 11 expresiones fruto de aplicar la transformada wavelet a la señal ECG (ECG, A1, A2, A3, A4, A5, D1, D2, D3, D4, D5).

El siguiente paso va a ser escoger las características de la señal que vamos a emplear para la posterior clasificación.

5.4.1 Características temporales y de amplitud

Estas características para covarianzas sencillas de la derivación 1 o de la derivación 2 van a ser las siguientes:

1. muestra 11
2. muestra 21 => 4 características de diezmado
3. muestra 31
4. muestra 41
5. primer cruce por cero (descendente)
6. Segundo cruce por cero (ascendente)
7. Instante en milisegundos del primer mínimo.
8. Amplitud del primer mínimo.
9. Instante en milisegundos del segundo máximo.
10. Amplitud del segundo máximo.

Las características escogidas para las covarianzas cruzadas de la derivación 1 con la derivación 2 son:

1. Muestra -40
2. Muestra -30
3. Muestra -20
4. Muestra -10
5. Muestra 0
6. Muestra 10
7. Muestra 20
8. Muestra 30
9. Muestra 40

10. Instante en milisegundos del máximo/mínimo (el que sea mayor)
11. Amplitud de ese máximo/mínimo
12. Distancia en milisegundos entre los dos cruces por cero a ambos lados del máximo/mínimo.

5.4.2 Características del ritmo de la señal

Además de estas características relacionadas con la amplitud y el tiempo de la señal, vamos a considerar también otro tipo de características como son las relacionadas con el ritmo de la señal: las distancias entre puntos R de la onda QRS del ECG (distancias RR). Las características de este tipo son siete:

1. RRant(distancia entre el punto R $i-1$ y el anterior),
2. RR (distancia entre el punto R i y el anterior),
3. RRpos (distancia entre el punto R $i+1$ y el anterior),
4. $dRR = RR_i - RR_{i-1}$,
5. $dRR_{pos} = RR_{i+1} - RR_i$,
6. RRmediana = mediana de los 10 últimos intervalos RR,
7. RR_{iqr} = rango intercuartílico de los 10 últimos intervalos RR

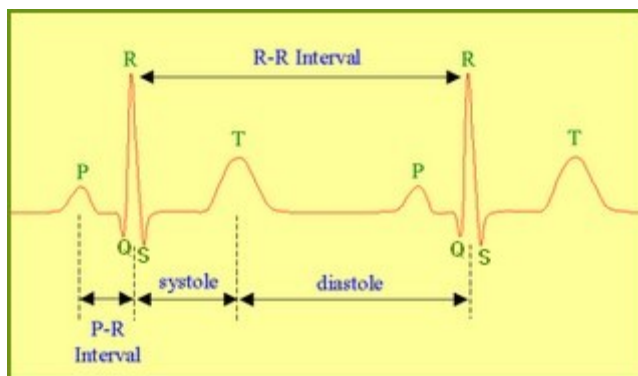


Figura 5-8 Forma de la onda QRS

Con estas características últimas, habrá que tener algunas consideraciones a la hora de almacenarlas en una matriz ya que, por ejemplo, los diez primeros valores de RR_{iqr} y $RR_{mediana}$, tendrán un valor 'NaN' (not a number), puesto que se calculan con los diez RR anteriores, por lo que no

tendremos que tenerlos en cuenta cuando posteriormente vayamos a operar de manera conjunta con estas matrices por filas o por columnas, de lo contrario nos estropearían el cálculo.

5.4.3 Tipos de latido

Otra información que nos resultará de utilidad almacenar junto con estas características es el tipo de latido correspondiente a cada uno de ellos. Para ello, mediante una función consultaremos esta información de las anotaciones que tienen los registros para cada latido y les asignaremos un tipo numérico del 1 al 5, que se corresponderá con los tipos de latidos que vamos a considerar:

Tipo 1 => Tipo N (pertenece a este grupo los latidos 'N', 'R', 'L', 'j', 'e')

Tipo 2 => Tipo S (pertenece a este grupo los latidos 'A', 'J', 'a', 'S')

Tipo 3 => Tipo V (pertenece a este grupo los latidos 'V', 'E')

Tipo 4 => Tipo F (pertenece a este grupo los latidos 'F')

Tipo 5 => Tipo Q (pertenece a este grupo los latidos 'Q', 'f', '/')

Por último, almacenaremos también el número de registro al que pertenece ese latido. Aunque no es una información indispensable, nos vendrá bien para tener más control de los latidos si tratamos todos los registros de forma conjunta.

5.4.4 Matriz de características

Ahora que ya tenemos seleccionadas todas las características que vamos a emplear en la clasificación, vamos a obtener dichas características de todos los registros de la base y a almacenarlas de manera que obtengamos una matriz de características por cada registro. Estas matrices tendrán una columna por latido del registro, y en cada una de esas columnas, estarán todas las características de ese latido.

Vamos a recordar que estas funciones de extracción de características van a ser aplicadas directamente a las matrices de covarianzas que hemos calculado anteriormente, es decir, que las características serán de la covarianza de la señal. Por tanto, hay que tener en cuenta que por cada latido, tendremos 10 características por cada expresión de la covarianza de la señal (ECG, A1, A2, A3, A4, A5, D1, D2, D3, D4, D5) y por cada una de las dos derivaciones (d1 y d2) y

12 características por cada expresión de la señal de la covarianza cruzada de las 2 derivaciones (dx). De esta manera tendremos por cada latido:

$(10 \text{ características} \times 11 \text{ expresiones} \times 2 \text{ derivaciones}) + (12 \text{ características} \times 11 \text{ expresiones}) = 352 \text{ características}$

Además de éstas, tenemos las 7 características de distancias RR, el tipo de latido y el número de registro.

Así tenemos 361 características por latido, que vendrán dispuestos de la siguiente forma:

latido1 latido2 latido3..... latido (numero de latidos)

Número de registro

tipo de latido (N=1,S=2,V=3,F=4,O=5)

característica1 (ecgd1)

característica2

....

característica10

característica1 (ecgd2)

....

característica10

.....

característica1 (ecgdx)

....

característica12

característica1 (A1d1)

....

característica10

.....

característica1 (A1d2)

....

característica10

característica1 (A1dx)

....

característica12

.....
característica1 (D5dx)
....
característica12
característica1 (RRs)
....
característica7 (RRs)

Tendremos por tanto 44 matrices de características correspondientes a los 44 registros de la base de datos, cada una de las cuales con una dimensión de 361 x numero de latidos del registro.

6. DEFINICIÓN DEL CLASIFICADOR

En este punto se explica el entrenamiento del clasificador. Para ello usaremos sólo los 22 primeros registros de la base. Los otros 22 los reservaremos para estudiar los resultados, y así evitar sesgar la evaluación.

6.1 CONJUNTOS DE CARACTERÍSTICAS

No podemos utilizar todas las características de las que disponemos para hacer un clasificador puesto que, no por tener muchas, la clasificación va a ser mejor. Llega un punto en el que, cuantas más características se tienen, la discriminación empeora al no poderse entrenar convenientemente el modelo.

Por esta razón, vamos a crear 4 agrupaciones de características diferentes para estudiar y comparar cuál de ellas consigue discriminar mejor entre los 5 tipos de latidos.

Estos conjuntos de características van a ser:

- FS1=> 7 características de ritmo + 10 características obtenidas de las funciones de autocovarianza de la primera derivación. Así, FS1 tendrá 17 características y se evaluarán 11 configuraciones de FS1, en función de si se aplican al ECG o a las 10 señales de aproximación o detalle a distintas escalas.
- FS2 => 7 características de ritmo + 10 características obtenidas de las funciones de autocovarianza de la segunda derivación. Así, FS2 tendrá 17 características y se evaluarán 11 configuraciones de FS2, en función de si se aplican al ECG o a las 10 señales de aproximación o detalle a distintas escalas.
- FS3 => 7 características de ritmo + 10 características obtenidas de las funciones de autocovarianza de la primera derivación + 10 características obtenidas de las funciones de autocovarianza de la segunda derivación. Así, FS3 tendrá 27 características y se evaluarán 11 configuraciones de FS3, en función de si se aplican al ECG o a las 10 señales de aproximación o detalle a distintas escalas.

- FS4 => 7 características de ritmo + 10 características obtenidas de las funciones de autocovarianza de la primera derivación + 10 características obtenidas de las funciones de autocovarianza de la segunda derivación + 12 características obtenidas de las funciones de covarianza cruzada de la primera con la segunda derivación. Así, FS4 tendrá 39 características y se evaluarán 11 configuraciones de FS4, en función de si se aplican al ECG o a las 10 señales de aproximación o detalle a distintas escalas.

Para seleccionar de una matriz de características las correspondientes a una de las 44 configuraciones posibles, creamos una función, a la que se le indica la configuración que se quiere obtener (conjunto de características y expresión de la señal) de una determinada matriz de características, y nos indica en que filas se encuentran las características deseadas.

Una vez que tenemos claro cómo recuperar las características de cada configuración, lo que se intenta ahora es agrupar, para cada una de las 44 configuraciones, los 5 tipos de latidos en 5 matrices distintas, para trabajar por separado con cada uno de los tipos (N, S, V, F, Q). Para ello implementamos una función que, tomando como entrada todas las matrices de características y la configuración que se desea, nos devuelve 5 matrices (una por cada uno de los tipos considerados) con tantas filas como características haya en esa configuración y tantas columnas como latidos de ese tipo haya. Este proceso se realizará para cada una de las 44 configuraciones, por lo que se genera un total de 220 matrices.

6.2 ESTRUCTURA DEL CLASIFICADOR

El clasificador que vamos a emplear en este proyecto es un clasificador sencillo de tipo Bayesiano, que discrimina utilizando un criterio Máximo a Posteriori (MAP) [1], [5]. Para ello son necesarios dos conjuntos de parámetros para cada clase: vector de medias $[\mu]$ y la matriz de covarianzas sigma $[\Sigma]$.

El siguiente paso será calcular estos parámetros que van a definir nuestro clasificador a partir de las matrices de características agrupadas por tipos y separadas por configuraciones que acabamos de obtener. Para ello creamos una función que, dada una configuración (conjunto de características y E_x) y una matriz de las anteriores matrices de tipos, obtiene su matriz Σ y su vector μ .

De esta forma, obtendremos por cada configuración 5 vectores de medias μ y 5 matrices de covarianza Σ (un vector y una matriz por cada tipo de latido i) calculados según las siguientes ecuaciones:

$$\mu_i = \frac{1}{M_i} \sum_{m=1}^{M_i} x_m \quad (3)$$

$$\Sigma_i = \frac{1}{M_i - 1} \sum_{m=1}^{M_i} (x_m - \mu_i) (x_m - \mu_i)^T \quad (4)$$

Siendo M_i el número de latidos de tipo i y x_m el vector de características del latido m .

Por último, crearemos unas estructuras para almacenar los parámetros que van a caracterizar el clasificador. Tendremos una estructura por cada configuración y cada una de ellas contendrá la siguiente información: descripción para saber de qué configuración se trata, los 5 vectores de medias μ y las 5 matrices Σ correspondientes a cada uno de los 5 tipos de latidos y, por último, la probabilidad de ocurrencia de cada uno de los tipos, que van a ser consideradas iguales para todos los tipos, excepto para el tipo Q que la consideraremos cero puesto que a partir de ahora no la vamos a considerar en este estudio, por su poca relevancia.

Estos parámetros definen el clasificador. Tenemos 44 estructuras de parámetros, es decir, 44 posibles clasificadores entre los que nos quedaremos con el que mejor clasifique.

Además de estos 44 posibles clasificadores, vamos a tener en cuenta la posibilidad de trabajar con clasificadores lineales o cuadráticos, y veremos con cuál se obtiene mejor resultado.

6.2.1 Clasificador cuadrático

Para el caso del clasificador cuadrático, se emplean los vectores de medias y matrices de covarianza definidos en las ecuaciones (3) y (4), es decir, un vector de medias y una matriz de covarianza diferente para cada tipo de latido.

Como ya hemos mencionado antes, el criterio que vamos a utilizar para nuestro clasificador es el de máximo a posteriori, que en el caso del clasificador cuadrático vendrá definido por la siguiente función de discriminación:

$$g_i(x) = -\frac{1}{2}x^T \Sigma_i^{-1} x + \mu_i^T \Sigma_i^{-1} x - \frac{1}{2} \mu_i^T \Sigma_i^{-1} \mu_i - \frac{1}{2} \log(|\Sigma_i|) + \log(P(\omega_i)) \quad (5)$$

6.2.2 Clasificador lineal

El clasificador lineal utiliza el mismo criterio de discriminación de máximo a posteriori, pero en este caso, las matrices de covarianza son consideradas iguales para todos los tipos ($\Sigma = \Sigma_i = \Sigma_j \quad \forall i \neq j$). Esta matriz de covarianza será el promedio ponderado de las covarianzas de los distintos tipos, es decir, a cada matriz de covarianza se le dará un peso determinado ω_i .

$$\Sigma = \frac{\sum_{i=1}^C \omega_i \sum_{m=1}^{M_i} (x_m - \mu_i)(x_m - \mu_i)^T}{\sum_{i=1}^C \omega_i M_i} \quad (6)$$

En nuestro caso, vamos a dar un peso de diez a los tipos N, S, V y un peso de uno al tipo F, puesto que hay menos latidos de ese tipo en nuestra base de datos.

De modo que la función discriminante en este caso será:

$$g_i(x) = \mu_i^T \Sigma^{-1} x - \frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i + \log(P(\omega_i)) \quad (7)$$

6.3 ENTRENAMIENTO

Una vez que ya tenemos todas las matrices de características de todos los registros de nuestra base de datos, separadas por configuraciones y por tipos de latido, el entrenamiento del clasificador ha consistido en la obtención de los parámetros que definen el clasificador (vectores de medias y matrices de covarianza) a partir de estas características, como hemos explicado anteriormente. La obtención de estos parámetros se ha realizado únicamente

con los primeros 22 registros de la base de datos, que hemos considerado los de entrenamiento.

6.4 EVALUACIÓN DE PRESTACIONES

Esta es la última fase del proceso. En ella vamos a recuperar la información de la segunda parte de la base de datos que habíamos reservado para hacer el test del clasificador ya entrenado con la otra parte. Recordemos que la información que teníamos de estos 22 registros de test, eran matrices de características de cada uno de dichos registros para cada una de las 44 posibles configuraciones.

Nuestro propósito ahora es introducir todos estos registros que no han participado en la fase de entrenamiento, por cada uno de los 44 posibles clasificadores, y evaluar los resultados para comprobar cuál es el que realiza una mejor predicción de tipos.

Hemos quedado en que cada uno de los 44 posibles clasificadores que vamos a estudiar, viene definido por una estructura de parámetros diferente, por tanto, lo primero que se va a hacer es adaptar las entradas (matrices de características completas de los registros de la parte de(test) al clasificador que se va a utilizar, es decir, de todas las características de la matriz, se quedará sólo con las que ese clasificador va a tener en cuenta para la clasificación.

Por último, se aplicará las funciones discriminantes del criterio MAP explicado anteriormente (tanto la lineal (7) como la cuadrática (5)) a todos los latidos de la parte de test de la base de datos, es decir, a todas las columnas de sus matrices de características, donde cada una de estas columnas es un vector de características x_m del latido m .

La función $g_i(x)$ que dé máximo valor, será la que determine que dicho latido es clasificado como tipo 'i'. Esta clasificación realizada se compara con la información almacenada en las anotaciones de la base de datos, para comprobar qué tipo de latido es en realidad y si se trata de un acierto o de un fallo.

En el apartado 7 se encuentran los resultados obtenidos y se explican las notaciones utilizadas.

7. RESULTADOS

A continuación se muestra la metodología que se ha seguido para evaluar las prestaciones de los clasificadores. Se han utilizado los mismos índices que se han usado en estudios previos [1], [5].

Una vez aplicado el clasificador a todos los latidos en las diferentes configuraciones, en cada una de ellas, podemos generar una tabla o matriz de confusión similar a la de la Tabla 1, donde se muestra cuántos del conjunto de test hay en cada una de las combinaciones posibles de etiquetas de referencia y clasificación por el algoritmo. En la matriz de confusión, los aciertos se sitúan en la diagonal (cuando la clasificación que da el algoritmo coincide con la etiqueta de referencia) y los errores fuera de la diagonal.

		ETIQUETA DEL ALGORITMO DE CLASIFICACIÓN					
		n	s	v	F	q	
ETIQUETA DE REFERENCIA	N	Nn	Ns	Nv	Nf	Nq	ΣN
	S	Sn	Ss	Sv	Sf	Sq	ΣS
	V	Vn	Vs	Vv	Vf	Vq	ΣV
	F	Fn	Fs	Fv	Ff	Fq	ΣF
	Q	Qn	Qs	Qv	Qf	Qq	ΣQ
							Σ

Tabla 1 Matriz de confusión

El siguiente paso es obtener, para cada clase de latidos "no normales", el número de verdaderos positivos (TP_x, latidos de la clase x que son clasificados correctamente como x), verdaderos negativos (TN_x, latidos de otro tipo distinto de x que no son clasificados como x), falsos positivos (FP_x, latidos de clases diferentes a x, que son clasificados incorrectamente como x) y falsos negativos (FN_x, latidos de la clase x, que son clasificados en otra clase).

Así, siguiendo los criterios utilizados en [1], definimos:

- Verdaderos Positivos:

$$TN = TPn = Nn;$$

$$TPv = Vv;$$

$$TPs = Ss;$$

$$TPf = Ff;$$

$$TPq = Qq;$$

- Falsos Positivos:

$$FPn = Vn + Sn + Fn + Qn;$$

$$FPv = Nv + Sv + Fv + Qv;$$

$$FPs = Ns + Vs + Fs + Qs;$$

$$FPf = Nf + Vf + Sf + Qf;$$

$$FPq = Nq + Vq + Sq + Fq;$$

- Verdaderos Negativos:

$$TNn = Vv + Vs + Vf + Vq + Sv + Ss + Sf + Sq + Fv + Fs + Ff + Fq + Qv + Qs + Qf + Qq;$$

$$TNv = Nn + Ns + Nf + Nq + Sn + Ss + Sf + Sq + Fn + Fs + Ff + Fq + Qn + Qs + Qf + Qq;$$

$$TNs = Nn + Nv + Nf + Nq + Vn + Vv + Vf + Vq + Fn + Fv + Ff + Fq + Qn + Qv + Qf + Qq;$$

$$TNf = Nn + Nv + Ns + Nq + Vn + Vv + Vs + Vq + Sn + Sv + Ss + Sq + Qn + Qv + Qs + Qq;$$

$$TNq = Nn + Nv + Ns + Nf + Vn + Vv + Vs + Vf + Sn + Sv + Ss + Sf + Fn + Fv + Fs + Ff;$$

- Falsos Negativos:

$$FNn = Nv + Ns + Nf + Nq;$$

$$FNv = Vn + Vs + Vf + Vq;$$

$$FNs = Sn + Sv + Sf + Sq;$$

$$FNf = Fn + Fv + Fs + Fq;$$

$$FNq = Qn + Qv + Qs + Qf;$$

A partir de aquí se definen la especificidad y las sensibilidades para cada clase, así como la exactitud (accuracy), que son los parámetros que usaremos para comparar las prestaciones de los distintos modelos de clasificación:

- Especificidad: $Sp = TN / \sum N;$
- Sensibilidades: $VEB Se = TPv / (TPv + FNv);$
 $SVEB Se = TPs / (TPs + FNs);$
 $F Se = TPf / \sum F;$
 $Q Se = TPq / \sum Q;$
- Exactitud: $Acc = (TN + TPs + TPv + TPf + TPq) / \sum ;$

La matriz de confusión mostrada en la tabla 1 se calcula para cada una de las 44 configuraciones y se pueden consultar en los anexos. Aquí vamos a mostrar únicamente los resultados finales de cada configuración (los parámetros arriba explicados), tanto para clasificador lineal como cuadrático, y nos quedaremos con la que mejor resultados obtenga.

7.1 RESULTADOS DEL CLASIFICADOR CUADRÁTICO

A continuación se muestran los resultados obtenidos en el conjunto de test usando el clasificador cuadrático (entrenado en el conjunto de entrenamiento) para cada una de las configuraciones de características estudiadas.

CONFIGURACION	SENSITIVITIES (%)				
	ACC (%)	SPEC (%)	SVEB	VEB	F
FS1_A1	90,20	100,00	0,00	0,00	0,00
FS1_A2	57,74	56,79	37,71	87,48	93,69
FS1_A3	52,96	53,21	26,24	64,80	92,87
FS1_A4	48,24	48,35	18,01	66,16	87,67
FS1_A5	60,03	61,40	16,00	78,23	33,42
FS1_D1	73,36	74,78	35,53	86,02	41,09
FS1_D2	69,19	70,33	32,90	81,04	63,56
FS1_D3	60,95	61,29	35,53	72,49	86,57
FS1_D4	57,04	57,21	27,53	73,65	93,42
FS1_D5	49,74	50,63	24,00	47,66	94,24

FS1_ECG	64,77	64,51	42,52	85,36	88,49
---------	-------	-------	-------	-------	-------

FS2_A1	58,64	60,18	20,70	56,81	93,97
FS2_A2	54,68	55,25	34,24	54,85	94,52
FS2_A3	44,49	42,57	62,17	56,20	95,06
FS2_A4	43,22	44,44	14,88	35,59	95,89
FS2_A5	50,54	52,61	15,44	34,43	94,24
FS2_D1	67,11	69,50	21,20	61,78	70,68
FS2_D2	65,02	67,00	19,41	64,55	83,56
FS2_D3	63,04	65,08	20,48	57,76	86,84
FS2_D4	56,97	58,46	17,40	56,96	95,06
FS2_D5	35,09	35,17	14,49	40,92	96,71
FS2_ECG	60,12	61,98	19,30	60,68	93,69

FS3_A1	86,32	0	0	0	100
FS3_A2	62,69	61,74	50,02	86,72	92,32
FS3_A3	48,60	47,95	35,08	65,20	92,87
FS3_A4	44,58	44,61	21,48	56,30	90,68
FS3_A5	56,43	58,56	16,95	54,55	37,80
FS3_D1	74,47	76,09	29,43	93,56	22,19
FS3_D2	72,50	74,05	26,74	93,01	23,56
FS3_D3	67,36	68,36	27,53	83,30	70,68
FS3_D4	60,77	60,89	30,94	79,63	92,05
FS3_D5	40,59	40,31	26,18	49,22	92,87
FS3_ECG	71,71	71,86	43,59	91,10	87,12

FS4_A1	90,20	100	0	0	0
FS4_A2	70,15	72,04	17,62	77,22	90,41
FS4_A3	63,13	64,31	22,49	72,14	90,13
FS4_A4	67,40	69,72	15,89	66,66	81,91
FS4_A5	66,45	69,53	21,71	51,63	44,65
FS4_D1	74,54	75,78	31,16	97,73	30,13
FS4_D2	73,78	74,94	29,43	97,88	37,80
FS4_D3	75,18	76,33	25,62	96,38	82,73

FS4_D4	69,82	71,16	16,06	88,53	91,50
FS4_D5	66,82	68,16	22,32	77,22	87,67
FS4_ECG	75,82	77,23	31,11	87,83	81,64

Tabla 2 Resultados de los clasificadores cuadráticos

*En los cálculos no se ha tenido en cuenta la clase de latidos Q, que como ya dijimos anteriormente, no va a ser considerada en este estudio.

7.2 RESULTADOS DEL CLASIFICADOR LINEAL

A continuación, los resultados obtenidos en el conjunto de test usando el clasificador lineal (entrenado en el conjunto de entrenamiento) para cada una de las configuraciones de características estudiadas.

CONFIGURACION	SENSITIVITIES (%)				
	ACC (%)	SPEC (%)	SVEB	VEB	F
FS1_A1	90,20	1,00	0,00	0,00	0,00
FS1_A2	76,97	80,06	26,24	67,82	52,05
FS1_A3	77,60	81,02	24,17	65,31	48,49
FS1_A4	79,79	83,76	17,35	66,11	44,93
FS1_A5	70,40	73,98	19,25	56,31	24,11
FS1_D1	74,18	76,45	34,25	66,31	76,16
FS1_D2	74,60	77,66	32,23	58,62	48,49
FS1_D3	74,03	77,54	25,85	61,99	8,22
FS1_D4	76,34	79,27	26,41	68,58	56,99
FS1_D5	76,46	80,11	16,12	69,98	25,20
FS1_ECG	76,15	79,26	21,82	67,12	66,30

FS2_A1	74,36	77,77	25,24	58,17	46,57
FS2_A2	75,20	78,43	31,50	57,68	47,40
FS2_A3	74,86	78,48	27,59	52,54	49,59
FS2_A4	69,28	73,78	13,60	30,57	83,01
FS2_A5	70,84	74,97	22,55	38,51	51,78

FS2_D1	73,19	77,03	19,47	49,17	66,02
FS2_D2	69,75	73,27	17,63	50,07	63,56
FS2_D3	75,73	79,76	17,90	56,10	44,10
FS2_D4	75,90	79,99	18,80	54,35	45,75
FS2_D5	70,38	74,67	14,21	37,46	76,71
FS2_ECG	72,76	76,10	23,28	57,01	51,50

FS3_A1	90,20	1,00	0,00	0,00	0,00
FS3_A2	79,77	82,55	48,68	58,72	56,71
FS3_A3	81,42	85,31	30,94	56,31	58,36
FS3_A4	76,45	80,45	15,44	54,05	78,63
FS3_A5	68,55	71,07	42,98	45,30	56,99
FS3_D1	75,28	77,96	34,97	58,97	80,82
FS3_D2	73,00	75,82	30,22	56,36	79,18
FS3_D3	74,50	78,06	23,56	61,69	21,09
FS3_D4	79,54	82,33	45,05	60,68	59,18
FS3_D5	76,56	80,17	13,71	62,89	80,82
FS3_ECG	75,00	78,25	29,32	54,80	68,77

FS4_A1	90,21	1,00	0,00	0,00	0,00
FS4_A2	65,94	68,89	18,52	52,74	62,19
FS4_A3	70,31	73,66	17,24	55,05	63,29
FS4_A4	75,41	79,02	12,81	60,23	87,67
FS4_A5	69,25	72,93	13,49	46,20	83,29
FS4_D1	67,45	70,66	19,08	46,86	80,82
FS4_D2	67,53	70,85	18,30	47,46	71,23
FS4_D3	64,37	67,23	18,30	54,45	45,75
FS4_D4	68,67	71,96	19,08	51,33	61,92
FS4_D5	74,68	78,47	10,13	57,11	89,86
FS4_ECG	61,67	64,13	17,18	53,69	65,48

Tabla 3 Resultados de los clasificadores lineales

7.3 ELECCIÓN DE CONFIGURACIÓN O CLASIFICADOR

Tras comparar los resultados obtenidos, se considera que los mejores resultados se dan con el clasificador cuadrático y con dos configuraciones: 'FS4_D3' y 'FS4_ECG', por lo que nos quedamos con ellas para continuar el estudio y validación del clasificador.

8. REDUCCIÓN DE CARACTERÍSTICAS

Intentamos hacer una mejora de los resultados obtenidos mediante una reducción de características, puesto que es posible que algunas características no esté aportando nada en cuanto a clasificación, por lo que las prestaciones podrían mejorar si las eliminamos y dejamos sólo aquellas que sí tienen poder discriminante. El propósito es ver cuántas y cuáles de ellas sería necesario eliminar para mejorar los resultados ya obtenidos.

Para ello nos centramos en las 2 configuraciones que nos han dado mejores resultados. Ambas son del Conjunto de Características 4, que contiene 39 características de la señal. Lo que se va a hacer ahora es ir reduciendo características de uno en uno para ver si de esta forma mejoran los resultados, utilizando un algoritmo "hacia atrás" o "backwards". Primero vemos que ocurre cuando quitamos sólo una, probando con cada una de las 39. Si al quitar una de ellas el resultado mejora, quitamos esa característica (en el caso de que sean varias tomaremos la que más mejore el resultado) y probamos a quitar una segunda característica de las 38 restantes, probando para ello a quitar la primera junto con cada una de las 38 para quedarnos con la segunda característica que consiga mejorar el resultado. Continuaremos así sucesivamente hasta que el resultado deje de mejorar. El valor que vamos a tomar para comparar resultados entre unos y otros será la media del Promedio de Sensibilidad Media y el Promedio Predictivo Medio, teniendo solamente en cuenta para este cálculo los tipos N, V, S (que son los más relevantes). A este valor lo llamaremos 'Índice de mérito (IM)':

$$\text{Sensibilidad_Media_V} = V_v / (V_v + V_n + V_s)$$

$$\text{Sensibilidad_Media_S} = S_s / (S_s + S_n + S_v)$$

$$\text{Sensibilidad_Media_N} = N_n / (N_n + N_s + N_v)$$

Sensibilidad_Media_Promedio =

$$(\text{Sensibilidad_Media_V} + \text{Sensibilidad_Media_S} + \text{Sensibilidad_Media_N}) / 3$$

$$\text{Predictivo_Medio_V} = V_v / (V_v + N_v + S_v)$$

$$\text{Predictivo_Medio_S} = S_s / (S_s + N_s + V_s)$$

$$\text{Predictivo_Medio_N} = N_n / (N_n + S_n + V_n)$$

$$\text{Predictivo_Medio_Promedio} = (\text{Predictivo_Medio_S} + \text{Predictivo_Medio_V} + \text{Predictivo_Medio_N}) / 3$$

Indice de mérito (IM) =

$$(\text{Sensibilidad_Media_Promedio} + \text{Predictivo_Medio_Promedio}) / 2$$

Para que los resultados sean más fiables, realizaremos la reducción de características con la parte de la base de datos de entrenamiento, y una vez obtenidas estas configuraciones reducidas, obtendremos los resultados de los clasificadores que conformarán estas 2 configuraciones reducidas en el conjunto de la base de datos, utilizando una estrategia de validación cruzada.

8.1 REDUCCIÓN DE LA CONFIGURACIÓN FS4_D3

Las características correspondientes a esta configuración son las siguientes:

1. RRant (distancia entre el punto R(i-1) y el anterior).
2. RR (distancia entre en punto R(i) y el anterior).
3. RRpos (distancia entre el punto R(i+1) y el anterior).
4. dRR = RR(i) - RR(i-1).
5. dRRpos = RR (i+1) - RR(i).
6. RRmediana = mediana de los 10 últimos intervalos RR.
7. RRIqr = rango intercuartílico de los 10 últimos intervalos RR.
8. muestra 11 [d1]
9. muestra 21 [d1]
10. muestra 31 [d1]
11. muestra 41 [d1]
2. Primer cruce por cero (descendente) [d1]
13. Segundo cruce por cero (ascendente) [d1]
14. Instante en milisegundos del primer mínimo. [d1]
15. Amplitud del primer mínimo [d1]
16. Instante en milisegundos del segundo máximo. [d1]
17. Amplitud del segundo máximo. [d1]

- 18. muestra 11 [d2]
- 19. muestra 21 [d2]
- 20. muestra 31 [d2]
- 21. muestra 41 [d2]
- 22. Primer cruce por cero (descendente) [d2]
- 23. Segundo cruce por cero (ascendente) [d2]
- 24. Instante en milisegundos del primer mínimo. [d2]
- 25. Amplitud del primer mínimo [d2]
- 26. Instante en milisegundos del segundo máximo. [d2]
- 27. Amplitud del segundo máximo. [d2]
- 28. muestra -40 [dx]
- 29. muestra -30 [dx]
- 30. muestra -20 [dx]
- 31. muestra -10 [dx]
- 32. muestra 0 [dx]
- 33. muestra 10 [dx]
- 34. muestra 20 [dx]
- 35. muestra 30 [dx]
- 36. muestra 40 [dx]
- 37. Instante en milisegundos del máximo/mínimo (el que sea mayor) [dx]
- 38. Amplitud de ese máximo/mínimo [dx]
- 39. Distancia en milisegundos entre los dos cruces por cero [dx]

(siendo $d1$: derivación 1, $d2$: derivación 2 y dx : correlación cruzada de $d1$ y $d2$).

El Índice de mérito (que llamaremos para abreviar IM a partir de ahora) que se obtiene de esta configuración con las 39 características (según los resultados obtenidos en el apartado 8) es:

IM₃₉ características= 0.8442 (si la clasificación la hacemos con la parte de training)

Tomando como referencia ese índice de mérito, vamos quitando una característica distinta cada vez, y calculamos los resultados que se obtienen con

estos nuevos 39 clasificadores basados en las 39 configuraciones de 38 características cada una.

De este proceso observamos que la que mejor resultado obtiene es la configuración que prescinde de la característica 27 (Amplitud del segundo máximo de la derivación 2 de la señal):

1. IM_38 características= 0.8483 (sin característica 27 de las 39)

Como el valor mejora con respecto a la configuración de 39 características, continuamos quitando una segunda. Además de quitar la característica 27, quitamos otra distinta en cada nueva configuración, calculando así 38 nuevos resultados de clasificadores de 37 características cada uno. El mejor resultado se obtiene con la que, además de eliminar la característica 27, elimina la 1 de las 38 restantes (RRant):

2. IM_37 características= 0.8507 (sin características: 27 de 39 y 1 de 38)

El resultado mejora de nuevo con respecto a la anterior, por tanto continuamos procediendo de la misma manera hasta que el resultado deje de mejorar y obtenemos lo siguiente:

3. IM_36 características = 0.8524 (sin características 27/39 , 1/38 y 20/37)
4. IM_35 características = 0.8544 (sin características 27/39 , 1/38, 20/37, 33/36)
5. IM_34 características = 0.8569 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35)
6. IM_33 características =0.8578 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34)
7. IM_32 características = 0.8606 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33)

8. IM_31 características = 0.8626 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32)
9. IM_30 características = 0.8624 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31)

Al quitar esta última característica el resultado empeora un poco, pero seguimos el proceso una iteración más para asegurarnos que ya a partir de ahí sólo empeora si seguimos quitando.

10. IM_29 características = 0.8631 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30)

Al contrario de lo que podía parecer, el resultado sigue mejorando, por lo que continuamos quitando características. Al quitar la 11, tenemos 2 casos que nos dan el mejor resultado:

11. a) IM_28 características = 0.8628 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29)
- b) IM_28 características = 0.8628 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 29/29)

En ambos casos, el resultado ha empeorado un poco, pero continuamos una iteración más para asegurarnos. Seguimos tanto el camino a) como el b) para ver cuál tiene mejores resultados:

12. a) IM_27 características = 0.8662 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28)
- b) IM_27 características = 0.8662 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 29/29, 19/28)

El resultado mejora por lo que continuamos quitando características:

13. a) IM_26 características = 0.8653 (sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27)

b) IM_26características = 0.8653 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 29/29, 19/28, 20/27)

14. a) IM_25características = 0.8670(sin características 27/39 , 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26)

b) IM_25características = 0.8670 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 29/29, 19/28, 20/27, 19/26)

Comprobamos que por ambos caminos se obtienen los mismos resultados, por lo que continuamos únicamente por el camino a) para simplificar:

15. IM_24características = **0.8680** (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26, 11/25)

16. IM_23características = 0.8678 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26, 11/25, 17/24)

En el paso 16, con 23 características el resultado ha empeorado. Continuamos quitando para ver si sigue empeorando o por el contrario ocurre como en las ocasiones anteriores en las que volvía a mejorar:

17. IM_22características = 0.8673 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26, 11/25, 17/24, 17/23)

18. IM_21características = 0.8665 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26, 11/25, 17/24, 17/23, 13/22)

19. IM_20características = 0.8650 (sin características 27/39, 1/38 y 20/37, 33/36, 32/35, 13/34, 7/33, 17/32, 15/31, 20/30, 19/29, 28/28, 20/27, 19/26, 11/25, 17/24, 17/23, 13/22, 15/21)

Vemos como a partir de la configuración de 24 características, los

resultados sólo empeoran, por tanto nos quedamos con esta nueva configuración en la que hemos eliminado 15 de las 39 iniciales. Las características con las que nos quedamos por tanto son:

2. RR (distancia entre en punto R(i) y el anterior.
3. RRpos (distancia entre el punto R(i+1) y el anterior).
4. $dRR = RR(i) - RR(i-1)$.
5. $dRRpos = RR(i+1) - RR(i)$.
6. RRmediana = mediana de los 10 últimos intervalos RR.
7. RRIqr = rango intercuartílico de los 10 últimos intervalos RR.
9. muestra 21 [d1]
10. muestra 31 [d1]
11. muestra 41 [d1]
12. Primer cruce por cero (descendente) [d1]
15. Amplitud del primer mínimo [d1]
16. Instante en milisegundos del segundo máximo. [d1]
17. Amplitud del segundo máximo. [d1]
19. muestra 21 [d2]
22. Primer cruce por cero (descendente) [d2]
23. Segundo cruce por cero (ascendente) [d2]
24. Instante en milisegundos del primer mínimo. [d2]
30. muestra -20 [dx]
31. muestra -10 [dx]
32. muestra 0 [dx]
33. muestra 10 [dx]
34. muestra 20 [dx]
37. Instante en milisegundos del máximo/mínimo (el que sea mayor) [dx]
38. Amplitud de ese máximo/mínimo [dx]

Con este nuevo clasificador compuesto por 24 características y entrenado con la parte de la base de datos de entrenamiento, pasaremos más adelante a evaluar los resultados con la base de datos completa, mediante validación cruzada, que se explicara en el apartado 10.

8.2 REDUCCIÓN DE LA CONFIGURACIÓN FS4_ECG

Hacemos el mismo procedimiento con la otra configuración que vamos a considerar: FS4_ECG. Esta configuración tiene las mismas 39 características que hemos enumerado para la anterior, con la diferencia de que estas características son aplicadas directamente a la señal y en la anterior se aplicaban sobre el detalle de orden 3 de la señal. Por este motivo, el proceso de reducción va a hacerse de la misma forma, se irán reduciendo características una a una mientras el resultado vaya mejorando. Cuando deje de mejorar, en varias iteraciones sucesivas, nos quedaremos con ese número de características que da mejor resultado para la nueva configuración. En ocasiones hay más de una característica que al reducirla obtiene el mejor resultado, en esos casos se seguirán ambos caminos y se descartará el que dé peor resultado de entre ellos.

Hemos podido comprobar tras hacer todo el proceso que, al igual que pasaba para el caso de la reducción de FS4_D3, cuando había más de un camino, ambos caminos iban obteniendo los mismos resultados, por lo que vamos a mostrar en este apartado sólo uno de los caminos para abreviar.

Estos son los resultados obtenidos:

La configuración FS4_ECG con sus 39 características:

IM_39características = 0.8326

Vamos reduciendo características:

1. IM_38características = 0.8363 (sin características 16/39)
2. IM_37características = 0.8436 (sin características 16/39, 13/38)
3. IM_36características = 0.8477 (sin características 16/39, 13/38, 12/37)
4. IM_35características = 0.8510 (sin características 16/39, 13/38, 12/37, 20/36)
5. IM_34características = 0.8531 (sin características 16/39, 13/38, 12/37, 20/36, 4/35)
6. IM_33características = 0.8565 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34)

7. IM_32características = 0.8589 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33)
8. IM_31características = 0.8614 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32)
9. IM_30características = 0.8639 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31)
10. IM_29características = 0.8665 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30)
11. IM_28características = 0.8694 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29)
12. IM_27características = 0.8715 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28)
13. IM_26características = 0.8724 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27)
14. IM_25características = 0.8727 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26)
15. IM_24características = 0.8748 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25)
16. IM_23características = 0.8770 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24)
17. IM_22características = 0.8777 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23)
18. IM_21características = **0.8790** (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23, 14/22)

19. IM_20características = 0.8771 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23, 14/22, 12/21)
20. IM_19características = 0.8768 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23, 14/22, 12/21, 1/20)
21. IM_18características = 0.8764 (sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23, 14/22, 12/21, 1/20, 8/19)
22. IM_17características = 0.8757(sin características 16/39, 13/38, 12/37, 20/36, 4/35, 16/34, 18/33, 16/32, 28/31, 27/30, 14/29, 14/28, 8/27, 10/26, 20/25, 20/24, 23/23, 14/22, 12/21, 1/20, 8/19, 11/18)

Vemos como la configuración de 21 características es la que obtiene mejor resultado (y lo mejora con respecto a la inicial de 39 características), y a partir de la cual no conseguimos mejorar el valor cuando seguimos quitamos más características. Por tanto nos quedamos con esta nueva configuración que elimina 18 de la original FS4_ECG. Esta nueva configuración 'FS4_ECG_Reducida' contiene 21 características y son las siguientes:

1. RRant (distancia entre el punto R(i-1) y el anterior.
2. RR (distancia entre en punto R(i) y el anterior.
3. RRpos (distancia entre el punto R(i+1) y el anterior).
5. dRRpos = RR (i+1) - RR(i).
6. RRmediana = mediana de los 10 últimos intervalos RR.
7. RRIqr = rango intercuartílico de los 10 últimos intervalos RR.
8. muestra 11 [d1]
10. muestra 31 [d1]
11. muestra 41 [d1]
15. Amplitud del primer mínimo [d1]
17. Amplitud del segundo máximo. [d1]
22. Primer cruce por cero (descendente) [d2]
25. Amplitud del primer mínimo [d2]

- 27. Amplitud del segundo máximo. [d2]
- 28. muestra -40 [dx]
- 29. muestra -30 [dx]
- 30. muestra -20 [dx]
- 31. muestra -10 [dx]
- 34. muestra 20 [dx]
- 37. Instante en milisegundos del máximo/mínimo (el que sea mayor) [dx]
- 38. Amplitud de ese máximo/mínimo [dx]

Al igual que la configuración que hemos obtenido anteriormente mediante reducción de características, ésta también será evaluada mediante el método de validación cruzada que se explica a continuación.

9. VALIDACIÓN-CRUZADA

Por último, hacemos una técnica de validación cruzada con las dos configuraciones que hemos obtenido tras la reducción de características.

Utilizaremos la técnica de “dejar uno fuera” (leave-one-out), pero aplicada por registros. Consiste en tomar para construir el clasificador todos los registros de la base de datos menos uno que dejamos fuera. El que dejamos fuera será el que pasemos por el clasificador para analizar los resultados. Haremos esto 44 veces, una por cada registro de la base.

A continuación se muestra la tabla con los latidos clasificados por el nuevo clasificador basado en la configuración FS4_D3_Reducida:

FS4_D3_Red																
	Nn	Ns	Nv	Nf	Sn	Ss	Sv	Sf	Vn	Vs	Vv	Vf	Fn	Fs	Fv	Ff
100	2213	1	0	0	2	30	0	0	0	0	0	0	0	0	0	0
103	2051	0	18	0	0	0	2	0	0	0	0	0	0	0	0	0
105	598	0	36	1535	0	0	0	0	0	1	36	0	0	0	0	0
111	115	5	1940	24	0	0	0	0	0	0	1	0	0	0	0	0
113	1767	8	0	0	0	1	5	0	0	0	0	0	0	0	0	0
117	599	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
121	1560	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0
123	1497	5	0	0	0	0	0	0	0	0	3	0	0	0	0	0
200	549	6	91	40	7	5	4	0	3	7	193	2	1	0	0	0
202	1604	67	4	363	0	40	12	0	0	0	19	0	1	0	0	0
210	1144	41	4	826	4	2	10	0	5	3	87	1	3	1	1	0
212	917	2	1802	0	0	0	0	0	0	0	0	0	0	0	0	0
213	1434	0	0	1110	5	2	0	20	0	0	81	122	7	0	3	335
214	1798	7	35	107	0	0	0	0	0	0	246	0	0	0	1	0
219	1610	368	0	0	4	3	0	0	0	3	49	0	0	0	1	0
221	1832	164	1	1	0	0	0	0	1	0	389	0	0	0	0	0
222	1597	205	82	45	13	177	2	0	0	0	0	0	0	0	0	0
228	1491	1	56	0	1	1	1	0	0	0	320	0	0	0	0	0
231	1503	45	6	0	1	0	0	0	0	2	0	0	0	0	0	0
232	364	27	2	0	663	644	68	0	0	0	0	0	0	0	0	0
233	1823	1	0	306	4	2	0	1	2	2	797	0	7	0	4	0
234	2686	0	0	0	49	1	0	0	0	0	3	0	0	0	0	0
101	1800	2	32	0	1	2	0	0	0	0	0	0	0	0	0	0
106	1427	20	0	0	0	0	0	0	3	24	447	0	0	0	0	0

108	390	2	705	0	0	0	4	0	0	0	7	0	1	0	0	0
109	1094	0	1323	0	0	0	0	0	1	0	33	0	2	0	0	0
112	2498	24	1	0	1	1	0	0	0	0	0	0	0	0	0	0
114	918	0	14	0	2	3	1	0	2	0	13	0	1	0	0	0
115	1935	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
116	929	0	4	1313	0	1	0	0	0	0	104	0	0	0	0	0
118	1610	0	163	0	19	0	63	0	3	7	4	0	0	0	0	0
119	1337	195	0	0	0	0	0	0	0	0	442	0	0	0	0	0
122	1739	0	723	0	0	0	0	0	0	0	0	0	0	0	0	0
124	1493	0	0	0	14	1	0	0	1	0	25	0	3	0	0	0
201	1054	191	0	370	3	39	80	0	5	0	193	0	1	0	0	0
203	1194	77	595	27	0	0	0	0	12	22	233	0	0	1	0	0
205	2055	5	0	450	1	1	0	0	0	0	32	0	5	0	0	5
207	20	0	4	0	0	0	0	0	3	0	108	0	0	0	0	0
208	1243	40	34	171	2	0	0	0	0	0	946	2	159	0	110	84
209	2585	3	7	0	357	22	3	0	0	0	1	0	0	0	0	0
215	2912	245	2	0	0	2	0	0	12	2	145	2	1	0	0	0
220	1935	5	0	0	13	81	0	0	0	0	0	0	0	0	0	0
223	1941	0	16	7	39	2	28	0	66	3	364	0	11	0	3	0
230	2233	8	0	0	0	0	0	0	0	0	1	0	0	0	0	0
TOTAL	65094	1772	7703	6696	1205	1064	283	21	119	76	5323	129	203	2	123	424

Tabla 4 Tabla de resultados de la validación cruzada por registros y agregados de la configuración FS4_D3_Reducida

La matriz de confusión de esta clasificación sería:

ETIQUETA DEL ALGORITMO DE CLASIFICACIÓN							
ETIQUETA DE REFERENCIA		n	s	V	F	q	
	N	65094	1772	7703	6696	Nq	81265
	S	1205	1064	283	21	Sq	2573
	V	119	76	5323	129	Vq	5647
	F	203	2	123	424	Fq	752
	Q	Qn	Qs	Qv	Qf	Qq	ΣQ
							90237

Tabla 5 Matriz de confusión de la validación cruzada de la configuración FS4_D3_Reducida

Los resultados de este clasificador son los siguientes:

CONFIGURACION	SENSITIVITIES (%)				
	ACC (%)	SPEC (%)	SVEB	VEB	F
FS4_D3_24características	79,68	80,10	29,25	48,52	56,38

Tabla 6 Resultados finales de la configuración FS4_D3_Reducida

Hacemos lo mismo con el clasificador basado en la configuración FS4_ECG_Reducida y obtenemos:

FS4_ECG_Red																
	Nn	Ns	Nv	Nf	Sn	Ss	Sv	Sf	Vn	Vs	Vv	Vf	Fn	Fs	Fv	Ff
100	2214	0	0	0	3	29	0	0	0	0	0	0	0	0	0	0
103	2047	3	19	0	2	0	0	0	0	0	0	0	0	0	0	0
105	348	5	94	1722	0	0	0	0	0	1	36	0	0	0	0	0
111	483	99	511	991	0	0	0	0	0	0	1	0	0	0	0	0
113	1768	7	0	0	0	1	5	0	0	0	0	0	0	0	0	0
117	599	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
121	1561	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
123	1499	3	0	0	0	0	0	0	0	0	3	0	0	0	0	0
200	549	5	97	35	4	7	5	0	7	13	182	3	0	0	0	1
202	1504	79	6	449	2	37	13	0	1	0	18	0	1	0	0	0
210	552	38	6	1419	7	3	6	0	5	6	84	1	4	0	0	1
212	2153	0	568	0	0	0	0	0	0	0	0	0	0	0	0	0
213	57	0	0	2487	2	0	2	23	0	0	51	152	7	0	2	336
214	1786	2	21	138	0	0	0	0	1	0	245	0	0	0	1	0
219	1667	307	4	0	4	3	0	0	0	8	44	0	0	0	1	0
221	1871	127	0	0	0	0	0	0	1	0	389	0	0	0	0	0
222	1514	348	47	20	10	180	1	1	0	0	0	0	0	0	0	0
228	1495	33	20	0	1	2	0	0	0	0	320	0	0	0	0	0
231	1508	43	3	0	1	0	0	0	0	2	0	0	0	0	0	0
232	330	57	6	0	647	655	72	1	0	0	0	0	0	0	0	0
233	1607	8	2	513	1	2	1	3	16	2	783	0	5	0	6	0
234	2686	0	0	0	49	1	0	0	0	0	3	0	0	0	0	0
101	1822	6	6	0	1	2	0	0	0	0	0	0	0	0	0	0
106	1416	30	1	0	0	0	0	0	1	71	402	0	0	0	0	0
108	767	0	330	0	0	0	4	0	0	0	7	0	1	0	0	0

109	1575	0	842	0	0	0	0	0	3	0	31	0	2	0	0	0
112	2512	9	2	0	1	1	0	0	0	0	0	0	0	0	0	0
114	855	6	71	0	2	4	0	0	1	0	14	0	1	0	0	0
115	1931	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0
116	1048	3	5	1190	0	1	0	0	0	0	104	0	0	0	0	0
118	1452	2	319	0	9	2	71	0	1	9	4	0	0	0	0	0
119	1319	207	0	6	0	0	0	0	0	0	442	0	0	0	0	0
122	2377	0	85	0	0	0	0	0	0	0	0	0	0	0	0	0
124	1493	0	0	0	14	1	0	0	1	0	25	0	3	0	0	0
201	1060	174	0	381	4	49	69	0	4	0	194	0	1	0	0	0
203	1071	89	686	47	0	0	0	0	29	20	218	0	0	1	0	0
205	2439	10	1	60	1	1	0	0	3	0	29	0	5	0	0	5
207	22	0	2	0	0	0	0	0	3	0	108	0	0	0	0	0
208	1325	22	21	120	2	0	0	0	1	0	944	3	139	0	114	100
209	2589	0	6	0	371	10	1	0	0	0	1	0	0	0	0	0
215	3125	34	0	0	0	2	0	0	5	0	149	7	1	0	0	0
220	1933	7	0	0	17	77	0	0	0	0	0	0	0	0	0	0
223	1954	0	8	2	32	4	33	0	71	2	360	0	6	0	6	2
230	2235	3	3	0	0	0	0	0	0	0	1	0	0	0	0	0
TOTAL	66118	1772	3795	9580	1187	1075	283	28	154	134	5193	166	176	1	130	445

Tabla 7 Tabla de resultados de la validación cruzada por registros y agregados de la configuración FS4_ECG_Reducida

La matriz de confusión de esta configuración es:

ETIQUETA DEL ALGORITMO DE CLASIFICACIÓN							
ETIQUETA DE REFERENCIA		n	s	v	F	q	
	N	66118	1772	3795	9580	Nq	81265
	S	1187	1075	283	28	Sq	2573
	V	154	134	5193	166	Vq	5647
	F	176	1	130	445	Fq	752
	Q	Qn	Qs	Qv	Qf	Qq	$\sum Q$
							90237

Tabla 8 Matriz de confusión de la configuración FS4_ECG_Reducida

Y de este clasificador obtenemos los siguientes resultados:

CONFIGURACION	SENSITIVITIES (%)				
	ACC (%)	SPEC (%)	SVEB	VEB	F
FS4_ECG_21caracteristicas	80,71	80,10	29,25	48,52	59,17

Tabla 9 Resultados finales de la configuración FS4_ECG_Reducida

10. COMPARATIVA CON TRABAJOS ANTERIORES

En este apartado hacemos una comparativa de los resultados obtenidos con nuestras dos configuraciones obtenidas y los obtenidos en el trabajo de referencia anterior. No se incluye la comparación con otros trabajos, ya que o bien no han utilizado el mismo método de evaluación o bien no proporcionan datos para calcular índices comparables.

CONFIGURACION	SENSITIVITIES (%)				
	ACC (%)	SPEC (%)	SVEB	VEB	F
FS4_D3_24características	79,68	80,10	29,25	48,52	56,38
FS4_ECG_21características	80,71	80,10	29,25	48,52	59,17
Philip de Chazal [1]	84.2	87.3	45.3	58.7	64.8

Tabla 10 Tabla comparativa de resultados con trabajos anteriores

Con esta comparativa comprobamos que no se ha conseguido obtener las prestaciones que se consiguen con otros trabajos basados en otras características y/o clasificadores más complejos, pero la ventaja es que tenemos un clasificador de baja complejidad, al ser sencillo tanto el clasificador como la obtención de las características a partir de la delineación.

11. CONCLUSIONES

En el presente proyecto se ha abordado el desarrollo y evaluación de un clasificador de latidos en el ECG basado en características de ritmo y morfológicas, extraídas de la función de covarianza de la transformada wavelet. El objetivo era conseguir un clasificador sencillo, basado en un esquema simple de clasificación, con características fáciles de obtener a partir del delineador de latidos utilizado, y poco sensibles a errores en la delineación.

Se han considerado diferentes conjuntos de características, en función de la derivación o derivaciones utilizadas, y de la señal considerada para extraer las características (bien el ECG, o bien las señales de aproximación o detalle de la transformada wavelet a diversas escalas). Los mejores resultados del clasificador se han obtenido utilizando información de las dos derivaciones en la señales de detalle en la escala 3 (que equivale a un filtro derivador más un filtro paso bajo), así como en el propio ECG.

El proceso posterior de reducción de características ha permitido mejorar las prestaciones del clasificador usando modelos de menor dimensión, evitando la sobreadaptación al conjunto de entrenamiento y mejorando por tanto su capacidad de generalización. Se ha conseguido una exactitud en la clasificación del 80%, si bien la sensibilidad de latidos no normales se ha situado por debajo de los obtenidos en otros trabajos publicados.

A continuación se citan algunas de las posibilidades para continuar el trabajo y mejorar el clasificador desarrollado y su validación, que no se han podido abordar en este proyecto.

- Cambios en el conjunto de características inicial: uso de otro tipo de características basados en la transformada wavelet (escala media, energías por escala, etc.).
- Utilización de características que tengan en cuenta la información espacial presente en las dos derivaciones, de forma que el clasificador no sea tan dependiente de las derivaciones grabadas.
- Dado que la distribución de las clases en el espacio de características puede no ser gaussiana (incluso puede ser multimodal), sería interesante comprobar el funcionamiento con clasificadores más complejos (no basados en una asunción de gaussianidad): mezcla de gaussianas, máquinas de vectores soporte o redes neuronales.
- Sería importante evaluar la capacidad de generalización de los modelos obtenidos mediante su evaluación en una base de datos distinta.

Además de las conclusiones técnicas del trabajo, me gustaría destacar que me ha resultado muy interesante su desarrollo, en particular conocer la cantidad de información que se puede sacar de una señal de este tipo y lo útiles que son algunos conceptos matemáticos. La mayor dificultad ha sido la gran cantidad de información que hay que procesar y el alto número de pruebas que ha habido que realizar para seleccionar la configuración. He tenido, además, la dificultad añadida de estar trabajando fuera de Zaragoza, por lo que la falta de medios y procesadores más potentes y de tiempo disponible me ha hecho retrasarme bastante. Por este motivo tengo que agradecer a mi director Juan Pablo Martínez el haberme facilitado las cosas y haberme prestado su atención y ayuda mientras ha estado en su mano.

12. BIBLIOGRAFÍA

- [1] Philip de Chazal, Member, IEEE, Maria O'Dwyer, and Richard B. Reilly, Senior Member, IEEE. "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features". IEEE Transactions on Biomedical Engineering, Vol. 51, NO. 7, JULY 2004.
- [2] Fayyaz-ul-Amir Afsar Minhas and Muhammad Arif. "Robust electrocardiogram (ECG) beat classification using discrete wavelet transform", Department of Computer and Information Sciences, Pakistan Institute of Engineering and Applied Sciences, Nilore, Islamabad, Pakistan. Doi: 10.1088/0967-3334/29/5/2003.
- [3] Juan Pablo Martínez, Rute Almeida, Salvador Olmos, Member, IEEE, Ana Paula Rocha, and Pablo Laguna, Member, IEEE. "A Wavelet- Based ECG Delineator: Evaluation on Standard Databases", IEEE Transactions on Biomedical Engineering, Vol. 51, No. 4, 4, April, 2004.
- [4] Leif Sörnmo, Pablo Laguna. "Bioelectrical Signal Processing in Cardiac and Neurological applications". Ed. Elsevier Academic Press. ISBN-13: 978-0-12-4375529.
- [5] Mariano Llamedo, Juan Pablo Martinez. "Heartbeat Classification Using Feature Selection Driven by Database Generalization Criteria". IEEE Transactions on Biomedical Engineering, Vol.58, NO. 3, March 2011.
- [6] G. B. Moody and R.G. Mark, "The MIT-BIH arrhythmia database on CD-ROM and software for use with it," in *Computers and Cardiology*. Los Alamitos, CA: IEEE Computer Society Press, 1990, pp.185-188.

13. ANEXOS

13.1 ANEXO1: RESULTADOS DEL CLASIFICADOR FS1_A1

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	38143	0	0	0	38143
S	1787	0	0	0	1787
V	1989	0	0	0	1989
F	365	0	0	0	365
					42284

13.2 ANEXO2: RESULTADOS DEL CLASIFICADOR FS1_A2

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	21663	1600	670	14210	38143
S	647	674	426	40	1787
V	15	121	1740	113	1989
F	16	3	4	342	365
					42284

13.3 ANEXO3: RESULTADOS DEL CLASIFICADOR FS1_A3

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	20299	1757	952	15135	38143
S	497	469	775	46	1787
V	11	584	1289	105	1989
F	16	2	8	339	365
					42284

13.4 ANEXO4: RESULTADOS DEL CLASIFICADOR FS1_A4

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	18444	1586	1395	16718	38143
S	371	322	1044	50	1787
V	11	535	1316	127	1989
F	36	2	7	320	365
					42284

13.5 ANEXO5: RESULTADOS DEL CLASIFICADOR FS1_A5

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	23421	1590	3686	9446	38143
S	425	286	1027	49	1787
V	19	363	1556	51	1989
F	66	31	146	122	365
					42284

13.6 ANEXO6: RESULTADOS DEL CLASIFICADOR FS1_D1

Etiqueta de Referencia	Etiqueta del algoritmo				sum
	n	s	v	f	
N	28525	1947	1266	6405	38143
S	652	635	477	23	1787
V	30	226	1711	22	1989
F	187	3	25	150	365
					42284

13.7 ANEXO7: RESULTADOS DEL CLASIFICADOR FS1_D2

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	sum	
N	26828	2183	1141	7991	38143	
S	659	588	519	21	1787	
V	34	299	1612	44	1989	
F	116	8	9	232	365	
					42284	

13.8 ANEXO8: RESULTADOS DEL CLASIFICADOR FS1_D3

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	sum	
N	23380	2437	1203	11123	38143	
S	653	635	459	40	1787	
V	35	418	1442	94	1989	
F	34	10	5	316	365	
					42284	

13.9 ANEXO9: RESULTADOS DEL CLASIFICADOR FS1_D4

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	sum	
N	21822	1649	1081	13591	38143	
S	635	492	614	46	1787	
V	25	373	1465	126	1989	
F	16	3	5	341	365	
					42284	

13.10 ANEXO10: RESULTADOS DEL CLASIFICADOR FS1_D5

Etiqueta de Referencia	Etiqueta del algoritmo				sum	sum
	n	s	v	f		
N	19314	1670	1065	16094	38143	
S	571	429	737	50	1787	
V	17	889	948	135	1989	
F	9	3	9	344	365	
						42284

13.11 ANEXO11: RESULTADOS DEL CLASIFICADOR FS1_ECG

Etiqueta de Referencia	Etiqueta del algoritmo				sum	sum
	n	s	v	f		
N	24607	1677	760	11099	38143	
S	626	760	359	42	1787	
V	19	149	1698	123	1989	
F	35	2	5	323	365	
						42284

13.12 ANEXO 12: RESULTADOS DEL CLASIFICADOR FS2_A1

Etiqueta de Referencia	Etiqueta del algoritmo				sum	sum
	n	s	v	f		
N	22956	1995	3539	9653	Nq	38143
S	62	370	1350	5	Sq	1787
V	15	743	1130	101	Vq	1989
F	13	5	4	343	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.13 ANEXO 13: RESULTADOS DEL CLASIFICADOR FS2_A2

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	21077		1497	3868	11701	Nq 38143
S	57		612	1109	9	Sq 1787
V	10		766	1091	122	Vq 1989
F	10		6	4	345	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.14 ANEXO 14: RESULTADOS DEL CLASIFICADOR FS2_A3

Etiqueta de Referencia	Etiqueta del algoritmo					Sum
	n	s	v	f	q	
N	16239		1519	5042	15343	Nq 38143
S	67		1111	598	11	Sq 1787
V	16		719	1118	136	Vq 1989
F	10		3	5	347	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.15 ANEXO 15: RESULTADOS DEL CLASIFICADOR FS2_A4

Etiqueta de Referencia	Etiqueta del algoritmo					Sum
	n	s	v	f	q	
N	16953		3833	2118	15239	Nq 38143
S	725		266	770	26	Sq 1787
V	138		990	708	153	Vq 1989
F	9		1	5	350	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.16 ANEXO 16: RESULTADOS DEL CLASIFICADOR FS2_A5

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	20069		2802	3566	11706	Nq 38143
S	591		276	901	19	Sq 1787
V	422		728	685	154	Vq 1989
F	12		2	7	344	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.17 ANEXO 17: RESULTADOS DEL CLASIFICADOR FS2_D1

Etiqueta de Referencia	Etiqueta del algoritmo					Sum
	n	s	v	f	q	
N	26511		1981	4573	5078	Nq 38143
S	343		379	881	184	Sq 1787
V	24		668	1229	68	Vq 1989
F	92		3	12	258	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.18 ANEXO 18: RESULTADOS DEL CLASIFICADOR FS2_D2

Etiqueta de Referencia	Etiqueta del algoritmo					Sum
	n	s	v	f	q	
N	25559		1618	5256	5710	Nq 38143
S	295		347	827	318	Sq 1787
V	27		637	1284	41	Vq 1989
F	34		5	21	305	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.19 ANEXO 19: RESULTADOS DEL CLASIFICADOR FS2_D3

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	24827		1617	4543	7156	Nq 38143
S	457		366	961	3	Sq 1787
V	24		775	1149	41	Vq 1989
F	15		3	30	317	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.20 ANEXO 20: RESULTADOS DEL CLASIFICADOR FS2_D4

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	22300		1351	3249	11243	Nq 38143
S	149		311	1318	9	Sq 1787
V	20		733	1133	103	Vq 1989
F	11		2	5	347	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.21 ANEXO 21: RESULTADOS DEL CLASIFICADOR FS2_D5

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	13415		3677	2020	19031	Nq 38143
S	632		259	855	41	Sq 1787
V	41		975	814	159	Vq 1989
F	6		3	3	353	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.22 ANEXO 22: RESULTADOS DEL CLASIFICADOR FS2_ECG

Etiqueta de Referencia	Algorithm Label					sum
	n	S	v	f	q	
N	23528	1634	3981	9000	Nq	38143
S	70	345	1366	6	Sq	1787
V	20	675	1207	87	Vq	1989
F	11	4	8	342	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.23 ANEXO 23: RESULTADOS DEL CLASIFICADOR FS3_A1

Etiqueta de Referencia	Algorithm Label					sum
	n	s	v	f	q	
N		0	0	0	38143 Nq	38143
S		0	0	0	1787 Sq	1787
V		0	0	0	1989 Vq	1989
F		0	0	0	365 Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.24 ANEXO 24: RESULTADOS DEL CLASIFICADOR FS3_A2

Etiqueta de Referencia	Etiqueta del algoritmo					Sum
	n	s	v	f	q	
N	23552	2176	2754	9661	Nq	38143
S	342	894	550	1	Sq	1787
V	15	151	1725	98	Vq	1989
F	16	6	6	337	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.25 ANEXO 25: RESULTADOS DEL CLASIFICADOR FS3_A3

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	18290	2480	2884	14489	Nq	38143
S	520	627	630	10	Sq	1787
V	18	571	1297	103	Vq	1989
F	11	5	10	339	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.26 ANEXO 26: RESULTADOS DEL CLASIFICADOR FS3_A4

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	17016	3641	1829	15657	Nq	38143
S	675	384	699	29	Sq	1787
V	130	607	1120	132	Vq	1989
F	21	1	12	331	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.27 ANEXO 27: RESULTADOS DEL CLASIFICADOR FS3_A5

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	22337	3342	4931	7533	Nq	38143
S	526	303	945	13	Sq	1787
V	422	396	1085	86	Vq	1989
F	27	54	146	138	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.28 ANEXO 28: RESULTADOS DEL CLASIFICADOR FS3_D1

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	29024	1838	4197	3084	Nq	38143
S	626	526	635	0	Sq	1787
V	23	93	1861	12	Vq	1989
F	236	2	46	81	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.29 ANEXO 29: RESULTADOS DEL CLASIFICADOR FS3_D2

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	28246	1642	5045	3210	Nq	38143
S	630	478	679	0	Sq	1787
V	23	114	1850	2	Vq	1989
F	220	6	53	86	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.30 ANEXO 30: RESULTADOS DEL CLASIFICADOR FS3_D3

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	26076	1633	4368	6066	Nq	38143
S	671	492	624	0	Sq	1787
V	39	289	1657	4	Vq	1989
F	70	5	32	258	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.31 ANEXO 31: RESULTADOS DEL CLASIFICADOR FS3_D4

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	23227		1556	3213	10147	Nq 38143
S		203	553	1030	1	Sq 1787
V		23	312	1584	70	Vq 1989
F		19	3	7	336	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.32 ANEXO 32: RESULTADOS DEL CLASIFICADOR FS3_D5

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	15379		4012	1262	17490	Nq 38143
S		679	468	605	35	Sq 1787
V		28	826	979	156	Vq 1989
F		14	2	10	339	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.33 ANEXO 33: RESULTADOS DEL CLASIFICADOR FS3_ECG

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	27413		1528	2615	6587	Nq 38143
S		334	779	674	0	Sq 1787
V		22	91	1812	64	Vq 1989
F		34	6	7	318	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.34 ANEXO 34: RESULTADOS DEL CLASIFICADOR FS4_A1

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	38143	0	0	0	0	Nq 38143
S	1787	0	0	0	0	Sq 1787
V	1989	0	0	0	0	Vq 1989
F	365	0	0	0	0	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.35 ANEXO 35: RESULTADOS DEL CLASIFICADOR FS4_A2

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	F	q	
N	27482	1142	3264	6255	0	Nq 38143
S	647	315	825	0	0	Sq 1787
V	14	366	1536	73	0	Vq 1989
F	25	0	10	330	0	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.36 ANEXO 36: RESULTADOS DEL CLASIFICADOR FS4_A3

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	24530	1306	4652	7655	0	Nq 38143
S	627	402	758	0	0	Sq 1787
V	20	446	1435	88	0	Vq 1989
F	23	2	11	329	0	Fq 365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ 42284

13.37 ANEXO 37: RESULTADOS DEL CLASIFICADOR FS4_A4

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	26594		1470	1607	8472	Nq	38143
S	797		284	704	2	Sq	1787
V	150		398	1326	115	Vq	1989
F	53		3	10	299	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.38 ANEXO 38: RESULTADOS DEL CLASIFICADOR FS4_A5

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	26521		2767	3018	5837	Nq	38143
S	500		388	899	0	Sq	1787
V	486		371	1027	105	Vq	1989
F	26		87	89	163	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.39 ANEXO 39: RESULTADOS DEL CLASIFICADOR FS4_D1

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	28908		1727	5115	2393	Nq	38143
S	734		557	496	0	Sq	1787
V	15		27	1944	3	Vq	1989
F	152		1	102	110	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.40 ANEXO 40: RESULTADOS DEL CLASIFICADOR FS4_D2

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	28588		1004	6494	2057	Nq	38143
S	740		526	521	0	Sq	1787
V	15		25	1947	2	Vq	1989
F	128		2	97	138	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.41 ANEXO 41: RESULTADOS DEL CLASIFICADOR FS4_D3

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	29116		909	4847	3271	Nq	38143
S	724		458	605	0	Sq	1787
V	18		36	1917	18	Vq	1989
F	15		2	46	302	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.42 ANEXO 42: RESULTADOS DEL CLASIFICADOR FS4_D4

		Etiqueta del algoritmo					
Etiqueta de Referencia	n	s	v	f	q	sum	
N	27143		1020	3211	6769	Nq	38143
S	176		287	1322	2	Sq	1787
V	16		145	1761	67	Vq	1989
F	20		0	11	334	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ	42284

13.43 ANEXO 43: RESULTADOS DEL CLASIFICADOR FS4_D5

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	26002	1367	2564	8210	Nq	38143
S	787	399	600	1	Sq	1787
V	32	309	1536	112	Vq	1989
F	28	2	15	320	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284

13.44 ANEXO 44: RESULTADOS DEL CLASIFICADOR FS4_ECG

Etiqueta de Referencia	Etiqueta del algoritmo					sum
	n	s	v	f	q	
N	29460	1077	3613	3993	Nq	38143
S	673	556	558	0	Sq	1787
V	28	153	1747	61	Vq	1989
F	51	2	14	298	Fq	365
Q	Qn	Qs	Qv	Qf	Qq	ΣQ
						42284