

# 1D Convolutional Neural Network for Detecting Ventricular Heartbeats

A. Suárez-León, and J. Núñez, *Member IEEE*

**Abstract**—This paper shows a novel approach for detecting ventricular heartbeats using a 1D Convolutional Neural Network (1D-CNN). The algorithm input is the raw ECG signal, i.e., no signal pre-processing nor feature extraction are involved. The output of the 1D-CNN is filtered using a combination of linear and nonlinear filters to produce the final output. The MIT-BIH arrhythmia database was used for both algorithm training/tuning and evaluation. The assessment methodology followed the inter-patient paradigm, where the algorithm was trained and evaluated using independent subsets. The performance of the proposed method was evaluated for two tasks; QRS detection, and heart-beat classification. QRS detection resulted in a sensitivity of 99.0% and a positive predictivity of 96.5%. The performance assessment of the ventricular ectopic beat detection resulted in a sensitivity of 85.8% and a positive predictivity of 64.5%. Although there is still room for improvement, the results suggest that convolutional neural networks are a promising approach for building heartbeat classifiers.

**Index Terms**—ECG, 1D-CNN, Heartbeat classifier.

## I. INTRODUCCIÓN

LAS enfermedades cardiovasculares (ECV) constituyen la primera causa de muerte a nivel mundial [1]. Según datos de la Organización Mundial de la Salud, alrededor de 17.65 millones de personas fallecieron en 2016 debido a ECV, lo que representa el 31.8% del total de decesos en el mundo. La muerte súbita es el cuadro clínico que después de presentarse los primeros síntomas conlleva a la pérdida de conciencia y luego a la muerte en un plazo inferior a las 24 horas [2]. Más del 25% de los fallecimientos por ECV corresponden a casos de muerte súbita cardiaca (MSC) [3]. Con frecuencia, los casos de MSC experimentan taquicardia ventricular seguida de fibrilación ventricular [4]. El evento que provoca la MSC en estos casos es la falla del sistema de activación eléctrica del corazón.

El electrocardiograma superficial (ECG) es el registro de la actividad eléctrica del corazón tomado en la superficie del cuerpo. Debido a la sencillez de la prueba, bajo costo y alto valor diagnóstico, es el método más usado en la práctica clínica para evaluar la función cardiovascular. El impulso eléctrico que da origen al latido en el ECG se genera en el nodo sinoauricular (nodo SA) que es el marcapasos natural del corazón [5]. En ciertas condiciones, por ejemplo, por fallo o enfermedad del nodo SA, otras regiones del corazón pueden iniciar el impulso cardiaco. Una contracción ventricular prematura (CVP) es un latido que ocurre prematuramente

(ectópico) y tiene como origen los ventrículos. La presencia ocasional de CVP aislados en el registro de ECG es benigna y no constituye riesgo para el paciente. No obstante, la aparición en patrones periódicos, frecuente o sostenida de CVP puede derivar en taquicardia ventricular, fibrilación ventricular y muerte súbita. Por esta causa, es de particular interés la detección automatizada, precisa y robusta de CVP, fundamentalmente, en registros ambulatorios, donde el número de latidos es elevado.

El tema de clasificación automática de latidos de ECG se ha tratado principalmente en el contexto de detección automática de arritmias [6], pero también es de interés en análisis de turbulencia [7] y como paso previo en los análisis de variabilidad de la frecuencia cardiaca [8] y de la repolarización ventricular [9], [10].

En la metodología de entrenamiento y evaluación de algoritmos de clasificación de latidos de ECG existen dos aproximaciones principales. El método independiente del paciente, emplea un subconjunto de registros para construir el clasificador y otro diferente para la evaluación. En la alternativa dependiente del paciente se emplea parte de los datos del registro para construir el clasificador. Estas aproximaciones reportan un rendimiento superior a las alternativas independientes del paciente. Sin embargo, requieren de la anotación de una fracción de los datos del registro bajo análisis lo que no siempre se presenta como una alternativa atractiva o incluso viable para el personal médico.

En este estudio se toma un enfoque independiente del paciente. Entre los trabajos con esta aproximación se encuentra el estudio de De Chazal *et al.* [11], que emplea un clasificador lineal y características temporales y morfológicas para distinguir entre las clases de latidos recomendadas en el estándar ANSI/AAMI EC57:1998 [12]. En Llamedo y Martínez [13] se presenta un método que emplea selección de características basado en la capacidad de generalización del algoritmo de clasificación. Otro estudio combina características estadísticas, temporales y morfológicas en la etapa de extracción. Para la clasificación se emplea un conjunto de clasificadores SVM y la regla de decisión por mayoría de votos [14].

Un elemento común que se observa en estos estudios es que la construcción del mapa de características es un proceso separado al entrenamiento del algoritmo de clasificación. Esto es, el experto selecciona *a priori* y evalúa las características a emplear. Sin embargo, con el desarrollo de los métodos de aprendizaje profundo (*deep learning*, DL) el enfoque cambia, centrándose en métodos de aprendizaje automático que incorporan la extracción de características en el proceso de construcción del algoritmo [15].

Los principios del DL se fundamentan en redes neuronales

A. A. Suárez-León, trabaja en el Departamento de Ingeniería Biomédica de la Facultad de Ingeniería en Telecomunicaciones, Informática y Biomédica, Universidad de Oriente, Santiago de Cuba, Cuba e-mail: [haal@uo.edu.cu](mailto:haal@uo.edu.cu)

José R. Núñez trabaja en la Universidad de la Costa, CP 080002, Barranquilla, Colombia e-mail: [jnunez22@cuc.edu.co](mailto:jnunez22@cuc.edu.co)

artificiales que proporcionan una forma viable de representar el conocimiento a través de mapas de características jerárquicos. La idea central es que la construcción de estos mapas forma parte del proceso de entrenamiento de la red neuronal. Esto significa que las operaciones antes indispensables en aplicaciones como reconocimiento de señales e imágenes (por ejemplo, el filtrado y la extracción de características) pueden ser sustituidas por un número de capas que codifican estas acciones. Todo ello, con la ventaja adicional de que las operaciones y características se aprenden directamente de los datos [16].

Los métodos de DL se han aplicado con éxito en problemas de reconocimiento de voz [17] y clasificación de textos, señales e imágenes [18], [19], [20]. En el contexto del análisis del ECG se han realizado estudios de clasificación de latidos usando redes recurrentes y clustering [21], redes convolucionales [22], [23], *autoencoders* y redes convolucionales [24], *Deep Belief Networks* (DBN) [25] y *Long Short-Term Memory* (LSTM) bidireccionales [26]. Otros estudios se centran en problemas de un nivel de abstracción superior, como el reconocimiento de diferentes tipos de arritmias [27], [28], [29], [30].

## II. OBJETIVO

El objetivo de este estudio es proponer un clasificador basado en redes convolucionales (1D-CNN) para la detección de latidos de origen ventricular. El método que se propone tiene las siguientes características:

- 1) Se ha adoptado un enfoque independiente del paciente, de forma que se puede comparar con los estudios que utilizan esta aproximación y además evaluar la capacidad de generalización del algoritmo.
- 2) La entrada del algoritmo se asume como la señal de ECG sin procesamiento previo.

El resto del documento se estructura como sigue: la sección III muestra la metodología propuesta para la construcción del algoritmo, en la sección IV se presentan los resultados de la evaluación, en el apartado V se discuten los resultados y finalmente en la sección VI se presentan las conclusiones.

## III. MATERIALES Y MÉTODOS

El estándar ANSI/AAMI EC57:1998 [12] define cinco clases de latidos: normales (N), ectópicos supraventriculares (S), ectópicos ventriculares (V), fusión de normales + ventriculares (F) y los de origen desconocido o artificiales (Q). Debido a que el propósito de este estudio es la detección de latidos de origen ventricular, es necesaria una distribución de clases diferente. Con este fin y para los propósitos de este estudio, se define la clase positiva ( $V'$ ) que agrupa a los latidos de origen ventricular, esto es, la unión de las clases (V)entricular y (F)usión. Similarmente, en la clase negativa (N+S) se agrupan los latidos que no tienen un origen ventricular y está formada por las clases (N)ormal y (S)upraventricular. La clase Q del estándar no se considera en este estudio.

El esquema general del detector propuesto se presenta en la figura 1. La red convolucional clasifica en una ventana de  $M$  muestras de señal de forma continua.

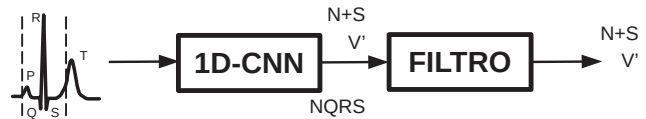


Fig. 1. Esquema general del algoritmo de detección de latidos de origen ventricular usando redes convolucionales y filtrado.

Para determinar si la ventana actual de longitud  $M$  contiene un latido válido se define una subventana de  $m$  muestras centrada en la ventana actual. Para un latido arbitrario, si la posición del pico de la onda R se encuentra dentro de la subventana, se considera un latido válido y corresponde a alguna de las clases, N+S o  $V'$ . En otro caso, la red genera como salida una tercera clase (NQRS) que determina la ausencia de latido válido en la ventana de análisis actual, figura 2.

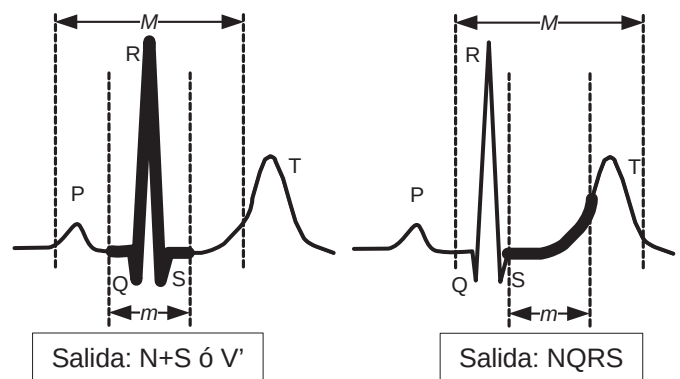


Fig. 2. La salida para el latido actual es igual a la clase (N+S o  $V'$ ) sólo si el pico de la onda R se encuentra dentro de la ventana interior de longitud  $m$ , en otro caso la salida es NQRS.

### A. Modelo de Red Convolucional

Una aproximación justa al diseño de la arquitectura de la red convolucional involucraría el uso de técnicas de validación cruzada u otras clases de meta-algoritmos en la fase de entrenamiento/validación. Desafortunadamente, afrontar el coste computacional que acarrea el uso de estas técnicas no resulta viable a partir de los recursos computacionales al alcance de los autores. Por esta razón, en este trabajo se emplearon criterios heurísticos para determinar la estructura de la red convolucional. Teniendo en cuenta que las redes convolucionales han tenido un gran éxito en la clasificación de imágenes, parte de este paradigma ha guiado la estructura de la red. Los elementos fundamentales que se tomaron como referencia para la definición de la arquitectura abordan por una parte las características de la señal y por otra parte la relación complejidad del modelo frente al costo computacional. Teniendo en cuenta estos elementos se establecen las siguientes premisas:

- El tamaño de la entrada se toma igual a 128 muestras (0.356 s a 360 Hz) lo que equivale a un poco más de 3 veces la duración de un complejo QRS típico.

- El uso de varios niveles de representación de la información, que se traduce en una arquitectura de varias capas donde la salida de cada bloque representa un nivel de abstracción diferente.
- Un costo computacional entre moderado y bajo en el sentido de que el tiempo de cómputo empleado en la evaluación de un segundo (360 muestras) de señal se encuentre suficientemente por debajo de 1 s en una estación de trabajo moderna de gama media (ver sección de software y hardware más adelante).

Siguiendo los puntos anteriores, se optó por un diseño con un número relativamente grande de filtros (64) de baja dimensión (entre 3 y 4) para las capas de convolución. Esta decisión se justifica en que los filtros grandes tienen una respuesta transitoria más dilatada, lo que se traduce en dos efectos indeseables: por una parte, la respuesta transitoria gana mayor influencia y por otra parte, en caso de no usarse *padding* en las capas convolucionales (como es el caso), el tamaño de la salida con respecto a la entrada decrece más rápidamente.

La red convolucional propuesta tiene la estructura que se muestra en la figura 3. El modelo está compuesto por 5 bloques CDS (Convolución-Detector-Submuestreo) en cascada, una capa de redimensionado y dos capas densas con sus correspondientes capas de activación. Los bloques CDS se componen de 3 capas: convolución o correlación 1D, detector o función no lineal y submuestreo (*pooling*). Cada capa de convolución 1D se compone de 64 filtros de longitud 3 (para los bloques 1, 3, 4 y 5) ó 4 (para el bloque 2). Todas las capas de submuestreo utilizan la función max y el tamaño de la región de submuestreo es 2 para todos los bloques excepto el bloque 5 que es 4. El bloque de submuestreo provee invarianza a traslación, en este caso se ha empleado la función max pues se asume que para la derivación dada el pico de la onda R (que es a su vez la característica que más destaca en el trazado del ECG) es un máximo.

La función de activación de las capas de detectores es la lineal rectificadora (*Rectified Linear Unit*, ReLU). Se ha elegido la función ReLU debido a que estudios recientes han mostrado que el uso de esta función de activación resulta en gradientes grandes y consistentes, lo que es ventajoso para el algoritmo de entrenamiento en la solución del problema de optimización implícito [16].

Las longitudes de los vectores de entrada a cada bloque CDS siguen la secuencia: 128(126)- 63(60) - 30(28) - 14(12) - 6(4). El número entre paréntesis es la longitud del vector a la salida de la capa de convolución y por ende el vector que se submuestra en este bloque. De esta forma, la estructura de la red garantiza que a la salida de cada bloque CDS se conserve la información relevante. Sólo la capa de remuestreo hace decrecer de forma significativa el vector de entrada al bloque siguiente.

A la salida de los bloques CDS una capa de redimensionado construye el vector de entrada a la estructura densa que le sigue. La capa densa 1 tiene 128 neuronas y la función de activación es ReLU. La capa densa final tiene 3 neuronas (igual al número de clases) y la función de activación es softmax. La Tabla I resume la estructura del modelo de red convolucional propuesto.

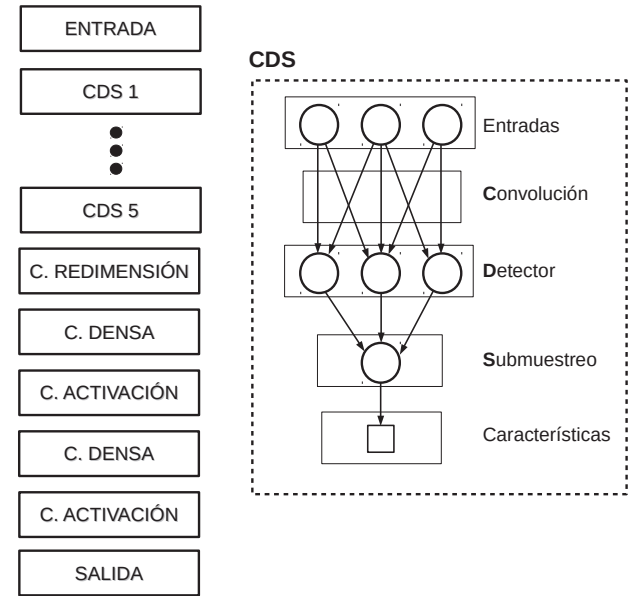


Fig. 3. Modelo de red convolucional propuesto.

TABLA I  
ESTRUCTURA DE LA RED CONVOLUCIONAL PARA CLASIFICACIÓN DE LATIDOS DE ORIGEN VENTRICULAR

No	Tipo	Observaciones
1	Conv-1D	No. Filtros:64, Tamaño filtro: 3
2	Detector	Función de Activación:ReLU
3	Submuestreo	Tamaño Región Submuestreo: 2
4	Conv-1D	No. Filtros:64, Tamaño filtro: 4
5	Detector	Función de Activación:ReLU
6	Submuestreo	Tamaño Región Submuestreo: 2
7	Conv-1D	No. Filtros:64, Tamaño filtro: 3
8	Detector	Función de Activación:ReLU
9	Submuestreo	Tamaño Región Submuestreo: 2
10	Conv-1D	No. Filtros:64, Tamaño filtro: 3
11	Detector	Función de Activación:ReLU
12	Submuestreo	Tamaño Región Submuestreo: 2
13	Conv-1D	No. Filtros:64, Tamaño filtro: 3
14	Detector	Función de Activación:ReLU
15	Submuestreo	Tamaño Región Submuestreo: 4
16	Redimensión	Operación: $64 \times 1 \times 1 \rightarrow 64 \times 1$
17	Densa 1	No. Pesos: $64 \times 128$
18	Activación 1	Función de Activación:ReLU
19	Densa 2	No. Pesos: $128 \times 3$
20	Activación 2	Función de Activación:softmax

## B. Filtro

Debido a que la salida de la red convolucional es continua y la detección/clasificación de un latido es un evento discreto, se requiere de cierto post-procesamiento a la salida de la red. El filtrado permite determinar: (1) la ocurrencia del latido y (2) la clase a la que pertenece. Es posible definir varias aproximaciones a este problema en dependencia de la aplicación. En lo sucesivo se ha adoptado un enfoque adecuado para un problema de procesamiento *offline*, de modo que se asume que se cuenta con el registro completo antes de aplicar el algoritmo.

La salida de la red para cada registro es una matriz de  $N \times 3$  donde  $N$  es el número de muestras del registro y cada columna corresponde a las clases de salida. Esta matriz es vectorizada

asignando a cada columna un coeficiente del 0 al 2 y sumando todas las columnas. En este esquema, 0 corresponde a la clase N+S, 1 a la clase V' y 2 a la clase NQRS. En el caso ideal, el vector obtenido es una secuencia de 2 donde ocasionalmente aparecen fragmentos de longitud  $m$  con valores 0 ó 1 en dependencia de la clase a la que pertenezca el latido. En la práctica, la salida de la red no es perfecta y es necesario un algoritmo más robusto. Para ello, se propone un método que realiza los siguientes pasos: (1) detecta las regiones donde la salida es distinta de 2, lo que corresponde a una detección de QRS y (2) halla la moda en estas zonas, lo que corresponde a clasificar el latido.

La figura 4 muestra las salidas de cada una de las etapas intermedias del proceso de filtrado. Primero, el resultado de la vectorización se convierte en un vector binario donde cada muestra contiene un 1 cuando el valor de la salida de la red es distinto de 2 y un 0 en otro caso. A continuación se aplica un filtro FIR de media móvil y fase cero con una ventana de tamaño igual a la mitad del número de detecciones (valores distintos de 2) expresado en porcentaje. Después, la señal se eleva al cuadrado y se aplica un umbral según la expresión (1).

$$\hat{s}_i = \begin{cases} s_i & s_i > t \\ 0 & s_i \leq t \end{cases}, \quad t = \left[ \frac{\lfloor L_F/2 \rfloor - 1}{L_F} \right]^2 \quad (1)$$

Donde  $s_i$  son las muestras del vector después de aplicar la función cuadrática,  $L_F$  es la longitud del filtro y  $\lfloor \cdot \rfloor$  representa la operación de redondeo al entero inferior. Este proceso de filtrado lineal, elevar al cuadrado y aplicar un umbral tiene como efecto agrupar las regiones donde se ha detectado un latido y al mismo tiempo eliminar falsas detecciones aisladas, ver figura 4.

A continuación, en la señal resultante cada región válida se expande en cuatro muestras, dos a cada lado. Este paso realiza un OR lógico con versiones de la misma señal desplazadas una muestra a la derecha y una muestra a la izquierda. Este proceso se realiza dos veces, lo que garantiza que un área donde se ha detectado un posible latido tenga siempre un tamaño mínimo de 5 muestras. En el siguiente paso se aplica otro filtro FIR fase cero con función transferencia equivalente  $H(z) = (1 - z^{-2})^2$ . Finalmente en la señal resultante se exploran las transiciones de -1 a 1 y de 1 a -1 para segmentar el inicio y el final del complejo QRS.

### C. Conjunto de Evaluación: MIT-BIH Arrhythmia Database

La base de datos *MIT-BIH Arrhythmia* [31], [32] contiene 48 registros anotados de aproximadamente 30 min. Cada registro contiene dos canales. Excepto el registro 104 que incluye las derivaciones V5 y V2, en todos los registros está presente la derivación MLII. El otro canal varía en dependencia del registro, aunque siempre es una derivación precordial. La frecuencia de muestreo es de 360 Hz, 11 bits de resolución y un rango dinámico de 10 mV. La base de datos incluye una variedad de fenómenos poco frecuentes pero clínicamente relevantes.

Por otra parte, varios de los registros de la serie 200 fueron seleccionados de manera que las características del ritmo, la

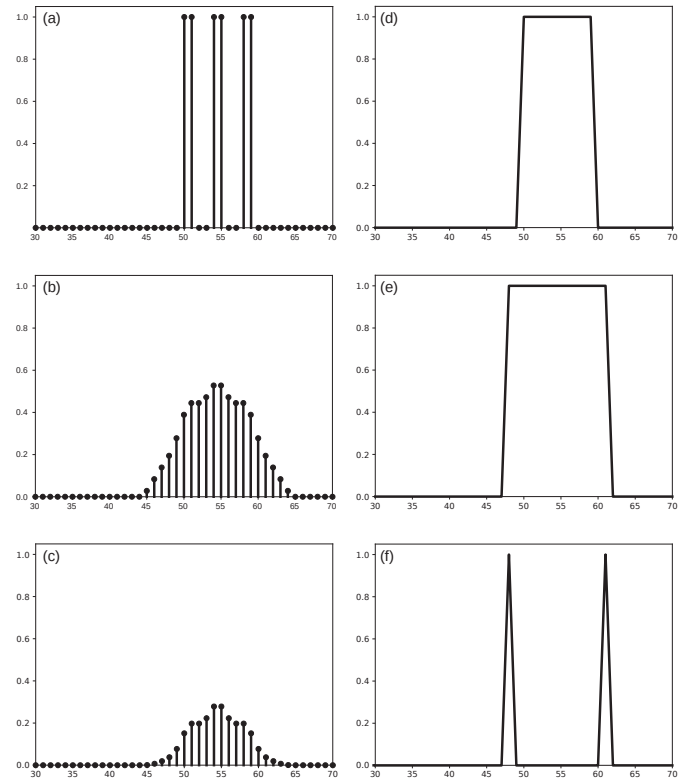


Fig. 4. Pasos del proceso de filtrado a la salida de la red convolucional, (a) es la salida de la red para un posible latido, observar que la salida de la red puede ser discontinua, (b) es la salida filtrada con un filtro FIR de media móvil y fase cero, (c) es la señal filtrada elevada al cuadrado, (d) es el resultado de aplicar el umbral  $t$ , (e) es el resultado después de expandir la zona de detección en dos muestras a cada lado y (f) es el resultado de aplicar el segundo filtro FIR. La posición del QRS puede estimarse como el punto medio entre los dos picos de la señal (f) y la clase, como la moda de la salida de la red en este intervalo.

morfología del complejo QRS, o la calidad de la señal en dichos registros constituyan una significativa dificultad para los detectores automáticos de arritmias.

Del total de 48 registros, cuatro (102, 104, 107, 217) corresponden a pacientes con marcapasos, que no se consideran en la evaluación de algoritmos de clasificación. De modo que la base de datos queda con un total de 44 registros. Para el entrenamiento, validación y evaluación de algoritmos independientes del paciente se sigue el esquema propuesto en [11]. La mitad de los registros (22) se emplean para entrenamiento-validación, el resto se emplea en evaluación. La distribución en los respectivos conjuntos se muestra en la Tabla II. Para el análisis de la señal se usa la derivación MLII pues se encuentra en todos los registros de los conjuntos DS1 y DS2.

Existe un notable desbalance entre clases en ambos grupos, esto acarrea un impacto negativo en el desarrollo del entrenamiento y en el rendimiento durante la evaluación. Esto es así porque las clases no tienen igual número de ejemplos, lo que puede polarizar la salida de la red dando una falsa percepción de buen rendimiento.

Para minimizar este efecto se introducen pesos diferentes para cada clase en la función de costo a evaluar por el algoritmo de entrenamiento. De este modo se penaliza más

TABLA II  
REGISTROS DE LA MIT-BIH Arrhythmia Database TOMADOS PARA ENTRENAMIENTO-VALIDACIÓN (DS1) Y PARA EVALUACIÓN (DS2)

DS1
101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 208, 215, 220, 223, 230.

DS2
100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234.

fuertemente al modelo por clasificar incorrectamente alguna de las clases N+S o V'. Los pesos en la función de costo se eligen como el inverso de la probabilidad de ocurrencia (frecuencia relativa) de las clases N+S y V'. Estas probabilidades se pueden calcular a partir del tamaño de la ventana interior (de longitud  $m$ ), la longitud de la ventana de análisis  $M$ , el número total de latidos en el registro ( $N$ ), el número promedio de latidos en cada registro ( $\bar{n}_{DS1}$ ) y la probabilidad de cada clase ( $p_C$ ) en el conjunto de datos de entrenamiento-validación según la ecuación (2).

$$w_C = \frac{1}{p_C} \cdot \frac{1}{p_L} = \frac{1}{p_C} \left( \frac{N - M}{m \cdot \bar{n}_{DS1}} \right) \quad (2)$$

Donde  $p_L$  es un estimado de la probabilidad de un complejo QRS válido,  $C$  representa la clase (N+S o V'),  $N = 650000$  y  $\bar{n}_{DS1} = 2319$ . Los valores de  $p_C$  para DS1 son  $p_{N+S} = 0.92$  y  $p_{V'} = 0.08$  respectivamente y se tomaron,  $M = 128$  y  $m = 10$ . Esto resultó en los siguientes valores de pesos  $w_{N+S} = 31$ ,  $w_{V'} = 370$  y  $w_{NQRS} = 1$ . Como algoritmo de entrenamiento se usó ADADELTA [33] con los parámetros por defecto. El número de épocas se fijó en 10. El conjunto de entrenamiento-validación fue dividido en 90% para entrenamiento y 10% para validación.

#### D. Hardware/Software

Para el entrenamiento del clasificador se utilizó una estación de trabajo con un procesador Intel® Core™ i3-7100 con 4 GB de memoria RAM ejecutando Linux Mint 19.1 Tessa 64-bits. Como lenguaje de programación se usó Python 3.6 con Keras [34] y Theano [35].

## IV. RESULTADOS

Para evaluar el rendimiento de la detección de QRS se emplean la sensibilidad (Se) y el valor predictivo positivo (P+). Las métricas para evaluar el rendimiento del algoritmo en la clasificación son las anteriores (Se y P+) y como referencia se añade la exactitud (Acc). Todos los resultados se muestran para el conjunto de evaluación (DS2). Los principales resultados de la evaluación se muestran en las Tablas III y IV. La Tabla III muestra los resultados en la detección del complejo QRS. Esta detección está implícita en la propia red, se separa de la clasificación sólo como una forma de presentar los resultados. Sin embargo, una consecuencia de usar el paradigma del DL es que la red aprende ambas tareas al mismo tiempo. La Tabla IV muestra los resultados de la clasificación. Finalmente, la

Tabla V muestra una comparación con algunos de los estudios que siguen la aproximación independiente del paciente.

TABLA III  
RESULTADOS DE LA DETECCIÓN DE QRS POR REGISTRO EN EL CONJUNTO DE EVALUACIÓN, FP: FALSOS POSITIVOS, FN: FALSOS NEGATIVOS, TP: VERDADEROS POSITIVOS, Se(%): SENSIBILIDAD Y P+(%): VALOR PREDICTIVO POSITIVO

Reg.	FP	FN	TP	Se(%)	P+(%)
100	0	0	2273	100	100
103	1	0	2084	100	100
105	32	27	2575	99.0	98.8
111	0	10	2114	99.5	100
113	1664	0	1795	100	51.9
117	0	1	1534	99.9	100
121	0	55	1808	97.0	100
123	0	1	1517	99.9	100
200	14	4	2597	99.8	99.5
202	0	4	2134	99.8	100
210	0	49	2602	98.2	100
212	3	1	2749	100	99.9
213	28	0	3251	100	99.1
214	63	9	2258	99.6	97.3
219	0	0	2154	100	100
221	0	4	2423	99.8	100
222	0	276	2207	88.9	100
228	0	53	2024	97.4	100
231	0	0	1571	100	100
232	0	5	1775	99.7	100
233	0	6	3081	99.8	100
234	0	0	2753	100	100
Total	1805	505	49279	99.0	96.5

TABLA IV  
RESULTADOS DE LA CLASIFICACIÓN POR REGISTRO, TP: VERDADEROS POSITIVOS, FP: FALSOS POSITIVOS, FN: FALSOS NEGATIVOS, TN: VERDADEROS NEGATIVOS, Se(%): SENSIBILIDAD, P+(%): VALOR PREDICTIVO POSITIVO Y Acc(%): EXACTITUD

Reg.	TP	FP	FN	TN	Se(%)	P+(%)	Acc(%)
100	1	0	0	2272	100	100	100
103	0	0	0	2084	-	-	100
105	25	152	30	2368	45.5	14.1	92.9
111	1	1017	0	1096	100	0.1	51.9
113	0	3	0	1792	-	0.0	99.8
117	0	2	0	1532	-	0.0	99.9
121	1	0	0	1807	100	100	100
123	3	0	0	1514	100	100	100
200	788	23	38	1748	95.4	97.2	97.7
202	15	6	6	2107	71.4	71.4	99.4
210	119	11	52	2420	69.6	91.5	97.6
212	0	23	0	2725	-	0.0	99.2
213	319	10	263	2659	54.8	97.0	91.6
214	170	423	90	1575	65.4	28.7	77.3
219	59	0	6	2089	90.8	100	99.7
221	390	0	2	2031	99.5	100	99.9
222	0	3	0	2204	-	0.0	99.9
228	356	17	15	1636	96.0	95.4	98.4
231	2	0	0	1569	100	100	100
232	0	1	0	1574	-	0.0	99.9
233	834	8	7	2226	99.2	99.0	99.5
234	2	1	1	2749	66.7	66.7	99.9
Total	3085	1700	510	43777	85.8	64.5	95.5

## V. DISCUSIÓN

De la Tabla III, se observa que el algoritmo tiene un rendimiento aceptable en la detección del QRS en 20 de los 22 registros. Los peores resultados en cuanto a detección de

TABLA V  
RENDIMIENTO DEL CLASIFICADOR EN LA DETECCIÓN DE LATIDOS DE ORIGEN VENTRICULAR, COMPARACIÓN CON OTROS ESTUDIOS

Estudio	Se(%)	P+(%)
De Chazal <i>et al.</i> [11]	77.7	81.9
Llamedo y Martínez [13]	81.0	87.0
Este estudio	85.8	64.5
Mondéjar-Guerra <i>et al.</i> [14]	94.7	93.9

QRS se producen en el registro 113 para los falsos positivos y el 222 para el caso de los falsos negativos. En el registro 113 el rendimiento se degrada significativamente debido a que el algoritmo detecta incorrectamente el pico de la onda T de la mayoría de los latidos como si se tratase de una contracción ventricular prematura. La causa de este comportamiento es que en este registro la onda T es alta y estrecha (en términos generales alcanza casi el 40% de la onda R que la precede) por lo que es similar a una CVP. Una posible solución a este problema sería introducir un período refractario después de la detección de cada QRS, lo que evitaría la detección incorrecta de la onda T. Otra variante puede ser extender la zona de análisis (ventana de longitud  $M$ ), esto permitiría incluir mayor información y diferenciar más fácilmente entre una onda T y un latido ectópico.

Con respecto al elevado número de falsos negativos en el registro 222, la razón puede ser la presencia de ritmo nodal (de escape) en el registro. La mayoría de los ejemplos de este tipo de ritmo (212/229 92.6%) se encuentran en el registro 222. De esta forma sólo menos del 8% de los latidos de esta clase presentes en la base de datos se encuentran disponibles para entrenamiento. Es por ello que no es sorprendente que este tipo de latido no sea reconocido correctamente pues se encuentra marginalmente representado en el conjunto de entrenamiento con sólo 16 ejemplos distribuidos entre los registros 108, 124 y 201. En consecuencia la red convolucional no recibe el número suficiente de ejemplos de esta clase de latidos durante el entrenamiento.

En el caso de la detección de latidos de origen ventricular, los peores rendimientos se alcanzan en los registros 111 y 214. Ambos registros incluyen el total de latidos del tipo bloqueo de rama izquierda. Este tipo de latido, pese a considerarse normal desde el punto de vista que tiene como origen el nodo SA, en general muestra variaciones morfológicas importantes con respecto a los latidos normales de corazones que no presentan bloqueo. Otro factor que debe señalarse es la diversidad morfológica en los cuatro registros de la base de datos que incluyen esta clase de latidos, ver figura 5.

Cuando se comparan los ejemplos de latidos con bloqueo de rama izquierda en el conjunto de entrenamiento-validación (registros 109 y 207) con los del conjunto de evaluación (111 y 214) se observan notables diferencias. Esto lleva a la conclusión de que pese a que este tipo de latido se encuentra representado en el conjunto de entrenamiento, las diferencias en la morfología entre ambos conjuntos son lo suficientemente significativas como para impactar negativamente en el rendimiento del clasificador. Una posible solución a este problema puede ser ampliar el conjunto de entrenamiento con

latidos de otras bases de datos. En otra vertiente, puede implementarse una alternativa de aprendizaje activo aprovechando la transferencia del aprendizaje característica de las redes convolucionales. Esto permitiría incluir algunos ejemplos de los latidos del registro actual en el entrenamiento.

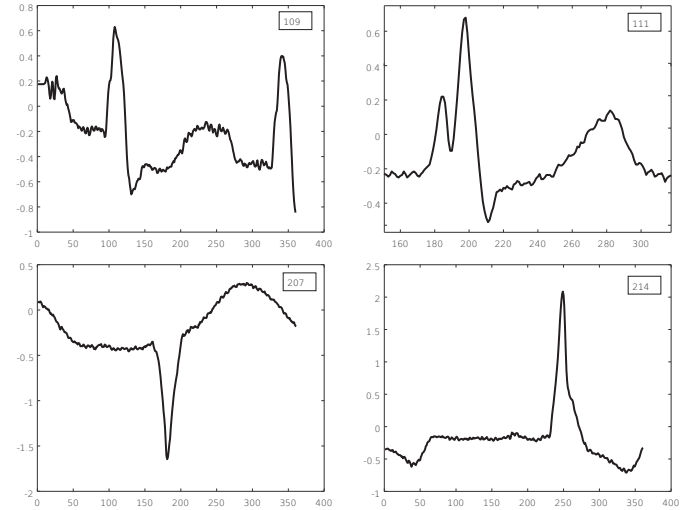


Fig. 5. Diversidad morfológica entre los registros de la base de datos con latidos que presentan bloqueo de rama izquierda; a la izquierda se muestran latidos típicos de los registros 109 y 207 pertenecientes al conjunto DS1 (entrenamiento-validación), a la derecha se muestran los latidos típicos de los registros 111 y 214 pertenecientes al conjunto DS2 (evaluación); todos pertenecen a la misma clase (N+S).

Pese a las limitaciones mencionadas arriba, el algoritmo tiene una sensibilidad global del 85.8%. Por otra parte, el valor predictivo positivo se ve impactado por el número de falsos positivos en 111 y 214. En la Tabla V se muestra esta situación, donde el algoritmo propuesto es segundo en sensibilidad pero último en valor predictivo positivo. Sin embargo, se debe tener en cuenta que todos los algoritmos citados parten de una localización más o menos precisa del complejo QRS y aplican sofisticadas técnicas de extracción de características. Nada de lo anterior es necesario en el método propuesto, por el contrario, la red construye su propio mapa de características en base a la entrada sin filtrar de la señal de ECG. Hasta el conocimiento de los autores, es la primera vez que se propone un algoritmo de clasificación de latidos de ECG que no incorpore ninguna de las características siguientes: filtrado en sus diferentes variantes, segmentación y/o detección de puntos característicos y extracción y/o selección de características. Además, la evaluación del algoritmo no se sustenta en resultados locales obtenidos de fragmentos específicos de la señal, sino que todas las muestras de todos los registros pueden ser evaluadas como entrada del clasificador.

## VI. CONCLUSIONES

Este estudio se ha centrado en la detección de latidos de origen ventricular a partir de la señal de ECG (cruda) usando redes convolucionales. Como resultado se ha construido, entrenado y evaluado un modelo de red convolucional 1D. Como referencia para la construcción y evaluación del modelo se ha tomado la base de datos *MIT-BIH Arrhythmia Database*

que es la metodología estándar. Los resultados mostraron que a pesar de evaluarse un modelo relativamente simple, este es capaz de detectar y clasificar latidos ventriculares con un rendimiento aceptable salvo en algunos casos desfavorables.

Desde el punto de vista del costo computacional, la evaluación de un segundo de señal toma un tiempo medio menor a 600 ms en el ordenador en el cual se entrenó el modelo. Este resultado muestra que, en principio, podría usarse en un sistema de clasificación en tiempo real, asumiendo que se modifique la estrategia de filtrado propuesta.

El algoritmo propuesto no es independiente de la frecuencia de muestreo de la señal de ECG que se esté procesando. Lo anterior implica que cualquier intento de uso en una señal que no esté muestreada a esta frecuencia debe considerar el uso de remuestreo con la consiguiente carga computacional que ello implica. Otro aspecto a señalar es que el método propuesto está optimizado para una derivación por lo que no se garantiza que se obtengan los mismos resultados en caso de que se use con otro canal. Sin embargo, el enfoque independiente del paciente que se ha empleado implica que, en principio, no se requiere de entrenamiento adicional en caso de que se considere emplear el modelo con diferentes pacientes y/o registros.

Entre las principales opciones que se pueden manejar para mejorar el rendimiento se encuentran: la inclusión de modelos con redes recurrentes que permitan aprovechar la información temporal presente en el ECG, incrementar el tamaño del modelo añadiendo nuevas capas y/o bloques CDS, implementar estrategias de aprendizaje activo aprovechando la capacidad de transferencia de aprendizaje disponible en las redes convolucionales, entre otras. Los resultados, aunque no superan aún las técnicas más depuradas del estado del arte, son prometedores pues aún tienen un margen significativo de mejoría. Lo que en suma apunta a que las redes convolucionales se vislumbran como una alternativa razonable al problema de detección y clasificación de latidos en el ECG.

## REFERENCIAS

- [1] World Health Organization, "Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva: 2018," 2019.
- [2] N. T. Srinivasan and R. J. Schilling, "Sudden cardiac death and arrhythmias," *Arrhythmia & electrophysiology review*, vol. 7, no. 2, p. 111, 2018.
- [3] S. G. Priori *et al.*, "2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: The Task Force for the Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death of the European Society of Cardiology (ESC) Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC)," *EP Europace*, vol. 17, pp. 1601-1687, 08 2015.
- [4] B. A. Koplán and W. G. Stevenson, "Ventricular tachycardia and sudden cardiac death," in *Mayo clinic proceedings*, vol. 84, pp. 289-297, Elsevier, 2009.
- [5] J. E. Hall, *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences, 2015.
- [6] E. J. da S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "ECG-based heartbeat classification for arrhythmia detection: A survey," *Computer Methods and Programs in Biomedicine*, no. 127, pp. 144-164, 2016.
- [7] A. Bauer, M. Malik, G. Schmidt, P. Barthel, H. Bonnemeier, I. Cygankiewicz, P. Guzik, F. Lombardi, A. Müller, A. Oto, *et al.*, "Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: International Society for Holter and Noninvasive Electrophysiology Consensus," *Journal of the American College of Cardiology*, vol. 52, no. 17, pp. 1353-1365, 2008.
- [8] M. Marek and C. A. John, "Task Force of the European Society of Cardiology: Heart rate variability standards of measurement, physiological interpretation, and clinical use," *Eur Heart J*, vol. 17, pp. 354-381, 1996.
- [9] E. Pueyo, P. Smetana, P. Caminal, A. B. De Luna, M. Malik, and P. Laguna, "Characterization of QT interval adaptation to RR interval changes and its use as a risk-stratifier of arrhythmic mortality in amiodarone-treated survivors of acute myocardial infarction," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 9, pp. 1511-1520, 2004.
- [10] J. Ramírez, M. Orini, J. D. Tucker, E. Pueyo, and P. Laguna, "Variability of ventricular repolarization dispersion quantified by time-warping the morphology of the T-waves," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1619-1630, 2016.
- [11] P. DeChazal, M. O'Dwyer, and R. Reilly, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1196-1206, jul 2004.
- [12] *Testing and Reporting Performance Results of Cardiac Rhythm and ST-Segment Measurement Algorithms*, American National Standard, ANSI/AAMI/ISO EC57, 1998-(R)2012. 2012.
- [13] M. Llamado and J. P. Martínez, "Heartbeat Classification Using Feature Selection Driven by Database Generalization Criteria," *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 616-625, March 2011.
- [14] V. Mondéjar-Guerra, J. Novo, J. Rouco, M. Penedo, and M. Ortega, "Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers," *Biomedical Signal Processing and Control*, vol. 47, pp. 41-48, 2019.
- [15] L. Deng and D. Yu, "Deep Learning: Methods and Applications," Tech. Rep. MSR-TR-2014-21, May 2014.
- [16] N. Ketkar *et al.*, *Deep Learning with Python*. Springer, 2017.
- [17] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury, *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal processing magazine*, vol. 29, 2012.
- [18] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in neural information processing systems*, pp. 649-657, 2015.
- [19] J. H. Tan, Y. Hagiwara, W. Pang, I. Lim, S. L. Oh, M. Adam, R. San Tan, M. Chen, and U. R. Acharya, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Computers in biology and medicine*, vol. 94, pp. 19-26, 2018.
- [20] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [21] C. Zhang, G. Wang, J. Zhao, P. Gao, J. Lin, and H. Yang, "Patient-specific ecg classification based on recurrent neural networks and clustering technique," in *Proceedings of the IASTED International Conference Biomedical Engineering (BioMed 2017)*, 2017.
- [22] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in *2016 6th international conference on IT convergence and security (ICITCS)*, pp. 1-5, IEEE, 2016.
- [23] P. Rajpurkar, A. Y. Hannun, M. Haghanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv preprint arXiv:1707.01836*, 2017.
- [24] M. A. Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340-354, 2016.
- [25] M. Huanhuan and Z. Yue, "Classification of Electrocardiogram Signals with Deep Belief Networks," in *2014 IEEE 17th International Conference on Computational Science and Engineering (IEEE, ed.)*, 2014.
- [26] O. Yildirim, "A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification," *Computers in Biology and Medicine*, vol. 96, pp. 189-202, 2018.
- [27] A. Y. Hannun, P. Rajpurkar, M. Haghanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, pp. 65-69, 2019.
- [28] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Information Sciences*, vol. 405, pp. 81-90, 2017.
- [29] Z. Xiong, M. K. Stiles, and J. Zhao, "Robust ECG signal classification for detection of atrial fibrillation using a novel neural network," in *2017 Computing in Cardiology (CinC)*, pp. 1-4, IEEE, 2017.

- [30] O. Yildirim, P. Plawiak, R.-S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Computers in Biology and Medicine*, 2018.
- [31] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [32] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
- [33] M. D. Zeiler, "ADADELTA: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.
- [34] F. Chollet *et al.*, "Keras." <https://keras.io>, 2015.
- [35] Theano Development Team, "Theano: A Python framework for fast computation of mathematical expressions," *arXiv e-prints*, vol. abs/1605.02688, May 2016.



**Alexander Alexeis Suárez León** Ingeniero en Automática, Máster y Doctor en Ingeniería Biomédica. Perteneció al Departamento de Ingeniería Biomédica de la Universidad de Oriente, Santiago de Cuba, Cuba. Entre sus áreas de interés se encuentran: el aprendizaje automático, el procesamiento de señales biomédicas, los sistemas embebidos y el hardware programable.



**José R. Núñez** Ingeniero Electricista y Máster en Automatización Industrial del departamento de Energía de la Universidad de la Costa, Barranquilla, Colombia. Sus áreas de interés son: Sistemas de Automatización Industrial, Generación, Transmisión y Distribución de Energía Eléctrica, Sistemas Domóticos e Inmóticos.