

Evaluación de la Calidad en Registros Electrocardiográficos Procedentes de Pacientes con Fibrilación Auricular Mediante una Red Neuronal Convolutiva

A. Huerta Herraiz¹, A. Martínez Rodrigo¹, M.A. Arias Palomares², P. Langley³,
J.J. Rieta Ibáñez⁴, R. Alcaraz Martínez¹

¹ Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {alvaro.huerta, arturo.martinez, raul.alcaraz}@uclm.es

² Departamento de Arritmias Cardíacas, Hospital Virgen de la Salud, Toledo, España, maapalomares@secardiologia.es

³ Facultad de Ciencias e Ingeniería, Universidad de Hull, Reino Unido, p.langley@hull.ac.uk

⁴ BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politècnica de Valencia, Valencia, España, jjrieta@upv.es

Resumen

La fibrilación auricular (FA) es la arritmia cardíaca más común en la práctica clínica. Su aparición se asocia con un mayor riesgo de eventos cardiovasculares, siendo su detección temprana un reto todavía sin resolver. Esta arritmia suele comenzar con episodios asintomáticos y muy breves, necesiéndose por tanto para su detección sistemas de registro electrocardiográfico (ECG) con capacidad para la monitorización de largos periodos de tiempo. El principal problema que presentan estos sistemas es su fuerte exposición al ruido, por ser adquiridos en entornos no controlados. Por lo tanto, es necesario disponer de sistemas automáticos que evalúen la calidad del registro de ECG adquirido, como primer paso para una detección precisa y temprana de la FA. Así pues, en este trabajo se introduce un algoritmo basado en aprendizaje profundo, con el objetivo de discernir entre segmentos ECG de alta y baja calidad en registros de ECG de una sola derivación, adquiridos en pacientes con FA. El método se basa en la alta capacidad de aprendizaje de una red neuronal convolutiva, la cual se ha entrenado con señales de ECG sin preprocesar convertidas a imagen. Los resultados muestran una gran capacidad del método para discernir entre los segmentos de ECG de alta y baja calidad, alcanzando una exactitud del 95%. Además, el método solo confunde el 6% de los segmentos con FA como intervalos de baja calidad. Estos resultados mejoran en más de un 20% el rendimiento de la mayoría de los algoritmos previamente publicados para la evaluación de la calidad del ECG que también tratan con señales de FA.

1. Introducción

Actualmente, la fibrilación auricular (FA) es la arritmia cardíaca más común en la práctica clínica diaria, afectando aproximadamente a 37,5 millones de personas en todo el mundo [1]. Su prevalencia está relacionada directamente con la edad, y se prevé que alcance proporciones epidémicas a mediados de este siglo [2]. No es una enfermedad mortal por sí misma, pero aumenta el riesgo de accidentes

cerebrovasculares, así como otros trastornos cardiovasculares [3]. A la hora de realizar un diagnóstico temprano y acertado de esta arritmia, el principal inconveniente es la corta duración que presentan sus episodios iniciales, de los cuales se calcula que un 90% son asintomáticos [4]. Para intentar solucionar este problema, los pacientes se monitorizan con dispositivos portátiles o vestibles, capaces de capturar señales electrocardiográficas (ECG) de muy larga duración, mientras el paciente sigue realizando su vida diaria normal.

El principal inconveniente de los dispositivos portátiles es su sensibilidad al entorno en el que se registra la señal, la cual puede estar afectada por distintas perturbaciones, tales como artefactos de movimiento, interferencias de la red eléctrica, o ruido de alta frecuencia, entre otros [5]. A pesar de que la eliminación de estos ruidos en el ECG ha sido motivo de diferentes trabajos de investigación [6], el rendimiento de los métodos propuestos es todavía limitado. Esto se debe a que algunas componentes de frecuencia del ruido, así como los artefactos de movimiento se solapan con la forma de onda del ECG en los dominios del tiempo y frecuencia [7]. Además, en el caso de la FA, las regulares ondas P se reemplazan por otras mucho más irregulares y caóticas, llamadas ondas fibrilatorias, las cuales presentan un aspecto muy parecido al ruido [8]. Por lo tanto, la evaluación automática de la calidad del ECG es de vital importancia para una detección temprana y precisa de la FA, previniendo la mala interpretación de intervalos de ECG ruidosos y evitando una mala interpretación del registro, un mal diagnóstico y un mal tratamiento, con los consiguientes riesgos asociados.

Para discernir adecuadamente entre intervalos de ECG de alta y baja calidad en registros de muy larga duración, se requieren algoritmos automáticos que no necesiten supervisión durante su funcionamiento, de hecho, los análisis realizados automáticamente pueden ser más eficientes y son mucho más rápidos que los realizados manualmente, presentando una mayor coherencia en sus decisiones [9].

Así pues, hasta la fecha, se han publicado diversos métodos automáticos basados en características dependientes del tiempo y de eventos morfológicos [7, 10]. Sin embargo, el rendimiento de estos trabajos está estrechamente ligado con su habilidad para detectar los puntos más relevantes en el ECG, los cuales se ven muy deteriorados en presencia de ruido [11]. No obstante, recientemente se han presentado algunos métodos automáticos basados en redes neuronales, las cuales han obtenido un alto rendimiento a la hora de detectar tramos de alta calidad en registros de ECG. Con esta técnica, no es necesaria la extracción de características del ECG ni otro tipo de intervención externa [6, 12].

El principal problema de estos últimos trabajos es que solamente se han validado con registros de ECG obtenidos de personas sin patologías conocidas, evitando así un contexto más complejo, tal como el correspondiente a la presencia de FA, donde episodios de ritmo sinusal y de la arritmia se van alternando a lo largo del tiempo [10]. Otro inconveniente en estos trabajos es el relativo al aprendizaje de la red neuronal propuesta, la cual se ha entrenado desde cero con un número relativamente bajo de muestras, lo cual hace que pueda estar sobreentrenada. Por lo tanto, en el presente trabajo se propone el uso de una red neuronal preentrenada para discernir entre tramos de ECG de alta y baja calidad, extraídos de señales de muy larga duración y una sola derivación, pertenecientes a pacientes con FA. Concretamente, se ha empleado la bien conocida red AlexNet [13] bajo el concepto de aprendizaje de transferencia y con un pequeño ajuste para desempeñar la tarea descrita.

2. Base de datos

Para el desarrollo del presente trabajo y la validación del método propuesto, se ha empleado una base de datos compuesta por 36 registros de una sola derivación con una duración de entre una y dos horas. Las señales se adquirieron en pacientes con FA intermitente empleando un dispositivo Holter vestible, de la marca NuuboTM, trabajando con una frecuencia de muestreo de 250 Hz y una resolución de 12 bits en un rango dinámico de ± 5 mV. Los episodios de FA se detectaron mediante un método automático aplicado a las señales en bruto, sin ningún tipo de procesamiento [8]. Además, dos expertos etiquetaron de manera independiente las señales de ECG en tramos de alta y baja calidad. El criterio establecido para considerar un intervalo de baja calidad fue que ambos no pudiesen identificar con seguridad todos los picos R en el mismo. Finalmente, los registros se segmentaron en tramos de 5 segundos, y fueron agrupados en dos categorías, es decir, alta y baja calidad. El número total de tramos disponibles para el análisis fue de 46.709, de los cuales 9.663 pertenecían a la clase de baja calidad y 37.046 a la categoría de alta calidad. Cabe destacar que, del total de segmentos de alta calidad, 10.338 pertenecían a FA y 26.708 a ritmo sinusal normal.

3. Métodos

El método propuesto para evaluar la calidad del ECG se compone de dos etapas. En primer lugar, los segmentos de ECG de 5 segundos de duración se transforman en matrices

de color de dos dimensiones y, en segundo lugar, se entrena la red neuronal utilizando como entrada dichas imágenes. A continuación, se detallan en profundidad cada uno de los dos pasos.

3.1. Transformada Wavelet Continua

A través de la transformada Wavelet continua (TWC) se consiguió una representación en imagen de cada segmento de ECG de 5 segundos que contenía su información tiempo-frecuencia. Esta transformación ha obtenido muy buenos resultados a la hora de tratar con señales aperiódicas, no estacionarias y con transitorios como el ECG [14]. A grandes rasgos, se basa en la descomposición de la señal en una serie de escalas temporales, las cuales representan un determinado rango de frecuencia concreto en el plano tiempo-frecuencia. Esta transformación se consigue correlando una señal inicial con versiones escaladas y desplazadas de una función wavelet, denominada *wavelet madre*. De forma más precisa, para calcular la TWC de una señal $x(t)$ se emplea la siguiente expresión:

$$TWC(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (1)$$

donde $a, b \in \mathbb{R}, a \neq 0$ son, respectivamente, el factor de escalado y el parámetro de desplazamiento de la wavelet madre, y $*$ denota el operador complejo conjugado. Como resultado, se obtiene una matriz de dos dimensiones llamada escalograma en la que los coeficientes calculados se muestran en las escalas y posiciones correspondientes. Este tipo de representación es muy común en trabajos que utilizan redes neuronales artificiales cuya entrada es una imagen. Como ejemplo, en la Figura 1 se muestran dos segmentos de ECG de 5 segundos de duración, en el que se observa la señal original y su correspondiente escalograma.

3.2. La red AlexNet

Las redes neuronales convolucionales se forman mediante la combinación de diversas capas que emulan las conexiones neuronales del cerebro. Los tres tipos más comunes son las capas de convolución, las capas de agrupación o *pooling* y las capas totalmente conectadas. En primer lugar, las capas de convolución se encargan de extraer las características más relevantes de la señal o imagen de entrada. En segundo lugar, se utilizan las capas de agrupación para reducir las dimensiones espaciales del mapa de características, sin perder información relevante. Por último, las capas totalmente conectadas se utilizan para conectar todas las neuronas de la capa anterior con todas las neuronas de la actual. Cabe destacar que el número total de neuronas de la última capa totalmente conectada define el número de clases de salida.

Para el desarrollo del presente trabajo se empleó la conocida arquitectura AlexNet [13]. Se adaptó para discernir entre segmentos de ECG de alta y baja calidad y se reentrenó para que adquiriese el nuevo conocimiento requerido. La arquitectura original de AlexNet se compone de ocho capas con capacidad de aprender, cinco convolucio-

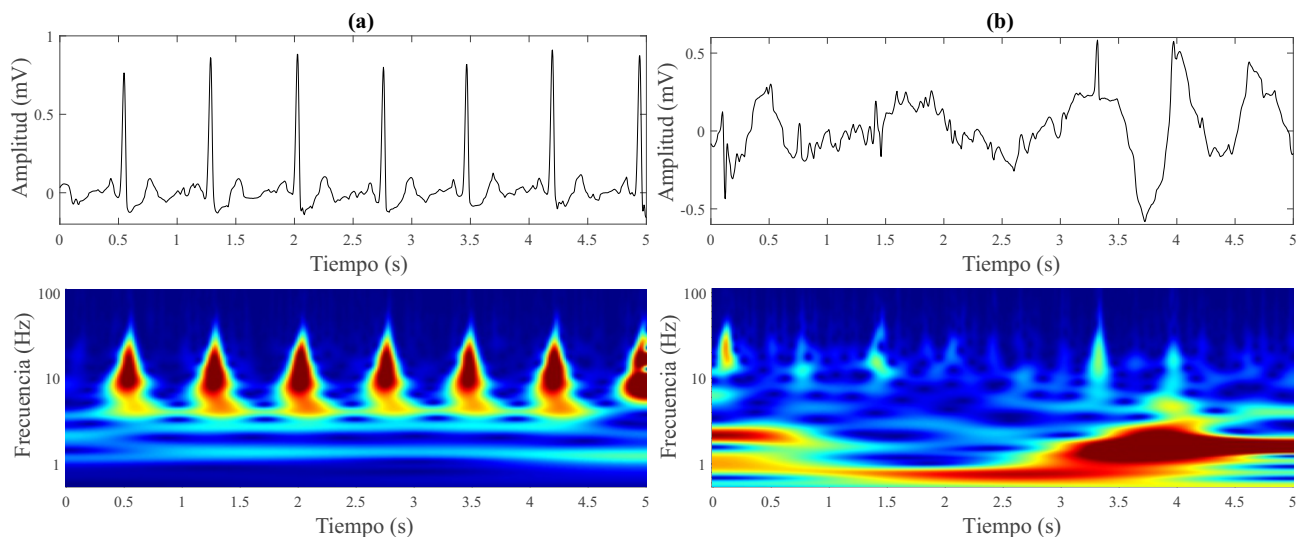


Figura 1. Ejemplo de ECG típico de alta (izquierda) y baja (derecha) calidad. Se representa la señal original junto con el escalograma obtenido mediante la TWC.

nales y tres totalmente conectadas, además de dos capas de agrupación. También, como función de activación emplea una función no lineal aplicada después de todas las capas convolucionales y totalmente conectadas. El modelo contiene dos funciones de *dropout* para reducir el sobreentrenamiento de la red.

Finalmente, la capa de salida de AlexNet se modificó para trabajar con dos clases y la red se re-entrenó con la base de datos descrita en la Sección 2, siendo necesarias 10 épocas con un tamaño de lote de 10 escalogramas. El proceso se desarrolló utilizando un algoritmo de descenso de gradiente estocástico con un momento de 0,9 y una tasa de aprendizaje de 0,0001 invariante durante todo el tiempo de entrenamiento.

3.3. Desarrollo experimental

A la vista de la base de datos descrita en la Sección 2, el número de segmentos de ECG de ambas clases se encuentra notablemente desbalanceado, pudiendo ocasionar resultados poco fiables o sesgados. Para solucionar este problema se ha planteado una validación basada en repeticiones. Así pues, la red neuronal se sometió a 20 ciclos de validación, en los que se utilizaron como datos de entrada 1.200 muestras seleccionadas al azar de cada uno de los grupos. Hay que destacar que el subconjunto de alta calidad estuvo siempre compuesto por 600 segmentos de ritmo sinusal normal y 600 intervalos de FA. Con cada uno de estos 20 conjuntos de 2400 muestras, el algoritmo propuesto se entrenó y validó 5 veces empleando un 80 % y un 20 % de las señales, respectivamente. Así pues, finalmente se condujeron 100 procesos de entrenamiento y test

En cada uno de estos procesos, se obtuvo el resultado de la clasificación en términos de sensibilidad (Se), especificidad (Es) y exactitud (Ex). Además también se obtuvo el máximo, el mínimo, la media y la desviación estándar (std) de estos tres índices de funcionamiento para todos los ciclos de validación. Cabe mencionar que la sensibilidad se definió como la tasa de segmentos de ECG de alta calidad

clasificados correctamente, la especificidad como el porcentaje de intervalos de baja calidad identificados correctamente y la exactitud como el número total de segmentos de ECG detectados correctamente. Además, se obtuvieron y promediaron las tasas de acierto para los tramos pertenecientes a ritmo sinusal normal (\mathcal{R}_{SR}) y FA (\mathcal{R}_{AF}), correctamente clasificados dentro del grupo de alta calidad.

4. Resultados

Los resultados obtenidos en cuanto a la media, std, máximo y mínimo de las métricas de funcionamiento evaluadas se encuentra en la Tabla 1. Se observa una exactitud en torno al 95 %, alcanzando un máximo del 96,88 %. También se aprecia una pequeña diferencia del 5 % entre sensibilidad y especificidad. En cuanto a las tasas de acierto dentro del grupo de ECG de alta calidad de intervalos de ritmo sinusal y de segmentos de FA, se observan valores superiores al 93 %. No obstante, se puede destacar que mientras que la tasa media de acierto de los segmentos de FA es de un 94 %, este valor sube hasta el 99,5 % para los segmentos de ritmo sinusal. Por último, en la tabla también se observan unos valores de std muy bajos para todos los índices, entre el 1 y el 3 %, por lo que no se aprecian diferencias significativas entre los diferentes ciclos de validación.

Índice	Se. (%)	Es. (%)	Ex. (%)	\mathcal{R}_{SR} (%)	\mathcal{R}_{AF}
Media	97.17	92.42	94.79	99.58	93.86
Std	1.57	3.14	1.41	0.94	2.75
Máximo	98.75	96.25	96.88	100	97.09
Mínimo	94.58	88.75	93.33	97.90	89.69

Tabla 1. Valores de media, std, máximo y mínimo obtenidos para los índices de funcionamiento de los 100 ciclos de validación conducidos.

5. Discusión

En este trabajo se ha presentado un novedoso algoritmo para discernir automáticamente entre tramos de alta y baja calidad en registros de ECG de una sola derivación obtenidos mediante un sistema portátil en pacientes con FA

intermitente. A pesar de que se han analizado las señales de ECG en bruto, se han alcanzado tasas de clasificación media comparables o incluso superiores a la mayoría de las técnicas anteriores que se han propuesto para evaluar la calidad en registros de ECG de una sola derivación [15].

Sin embargo, a la hora de compararse con estos métodos previos, hay que tener en cuenta ciertos aspectos, como por ejemplo que la base de datos empleada es diferente en cada trabajo. Además, mientras que en el presente estudio se desarrolló un robusto sistema de validación empleando más de 46.000 segmentos de ECG en 100 iteraciones, en la mayoría de los trabajos anteriores se ha realizado una validación con un solo conjunto de datos bastante limitado [15]. Por ejemplo, Zhao y Zhang [16] tan solo utilizaron 300 intervalos de ECG de 30 segundos de duración de una base de datos pública para el desarrollo de su trabajo. Claramente, con un análisis de la base de datos completa (más de 8.000 segmentos) se habrían obtenido resultados más robustos y una comparación más justa.

También hay que destacar que la mayoría de los métodos anteriores solamente se han validado con registros de ECG obtenidos de pacientes sin patologías [10]. Por lo tanto, considerar de manera conjunta segmentos de ritmo sinusal y FA dentro del grupo de alta calidad hace más compleja la tarea de discernir el grado de calidad que presenta el ECG que se procesa. De hecho, los pocos métodos que han considerado registros de FA para su validación han mostrado un descenso considerable en su capacidad de identificar la calidad del ECG, en comparación a cuando solo analizaban registros de pacientes sanos, obteniendo exactitudes en torno al 50 y 70 % [10, 11, 15]. Sin embargo, el presente trabajo ha demostrado una gran capacidad para discernir entre segmentos de FA de alta calidad y segmentos de ruido alrededor del 94 %, mientras que la exactitud global del método para discernir entre intervalos de ECG de alta y baja calidad también se ha mantenido en torno al 95 %.

Finalmente, cabe mencionar algunas limitaciones del estudio. Por un lado, la principal desventaja del método propuesto es la falta de una interpretación clara de los resultados en términos fisiopatológicos. En este sentido, los algoritmos basados en redes neuronales convolucionales se caracterizan por mantener ocultas las dependencias funcionales entre las entradas y las salidas [17]. Por otro lado, a pesar de que en el presente trabajo se ha analizado una base de datos con un elevado número de registros la cual se ha adquirido en un entorno totalmente realista, un estudio más exhaustivo incluyendo diferentes tipos de ruidos y artefactos sería muy útil para completar los resultados obtenidos.

6. Conclusiones

Se ha propuesto un método novedoso basado en una red neuronal convolucional pre-entrenada para discernir de manera fiable entre intervalos de alta y baja calidad en registros de ECG de una sola derivación, incluyendo episodios de ritmo sinusal normal y FA. El método se ha verificado de manera exhaustiva mediante múltiples ciclos de validación, obteniendo un rendimiento superior a la mayoría de métodos propuestos en trabajos previos.

Agradecimientos

Este trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM, AICO/2019/036 de la GVA y FEDER 2018/11744.

Referencias

- [1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke* Jan 2020;1–5.
- [2] Krijthe BP, Kunst A, Benjamin EJ, Lip GYH, Franco OH, Hofman A, Wittteman JCM, Stricker BH, Heeringa J. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J* Sep 2013;34(35):2746–51.
- [3] Khoo CW, Krishnamoorthy S, Lim HS, Lip GYH. Atrial fibrillation, arrhythmia burden and thrombogenesis. *Int J Cardiol* Jun 2012;157(3):318–23.
- [4] Stachon P, Ahrens I, Faber T, Bode C, Zirlirk A. Asymptomatic atrial fibrillation and risk of stroke. *Panminerva Med* Dec 2015;57(4):211–5.
- [5] Nagai S, Anzai D, Wang J. Motion artefact removals for wearable ECG using stationary wavelet transform. *Healthcare Technology Letters* 2017;4(4):138–141.
- [6] Yoon D, Lim HS, Jung K, Kim TY, Lee S. Deep learning-based electrocardiogram signal noise detection and screening model. *Healthcare Informatics Research* 2019;25(3):201–211.
- [7] Oster J, Behar J, Sayadi O, Nemati S, Johnson AE, Clifford GD. Semisupervised ECG ventricular beat classification with novelty detection based on switching kalman filters. *IEEE Trans Biomed Eng* 2015;62(9):2125–2134.
- [8] Ródenas J, García M, Alcaraz R, Rieta JJ. Combined non-linear analysis of atrial and ventricular series for automated screening of atrial fibrillation. *Complexity* 2017;2017.
- [9] Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR. Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine* 2018;161:1–13.
- [10] Behar J, Oster J, Li Q, Clifford GD. ECG signal quality during arrhythmia and its application to false alarm reduction. *IEEE Transactions on Biomedical Engineering* 2013;(6).
- [11] Satija U, Ramkumar B, Manikandan MS. Automated ECG noise detection and classification system for unsupervised healthcare monitoring. *IEEE J Biomed Health Inform* 05 2018;22(3):722–732.
- [12] Huerta A, Martínez-Rodrigo A, González VB, Quesada A, Rieta J, Alcaraz R. Quality assessment of very long-term ECG recordings using a convolutional neural network. In *IEEE E-Health and Bioengineering Conference*. 2019; 1–4.
- [13] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*. 2012; 1097–1105.
- [14] Shoeb A, Clifford G. Chapter 16—wavelets; multiscale activity in physiological signals. *Biomedical Signal and Image Processing* Springer New York NY USA 2005;.
- [15] Satija U, Ramkumar B, Manikandan MS. A review of signal processing techniques for electrocardiogram signal quality assessment. *IEEE Rev Biomed Eng* 2018;11:36–52.
- [16] Zhao Z, Zhang Y. SQI quality evaluation mechanism of single-lead ECG signal based on simple heuristic fusion and fuzzy comprehensive evaluation. *Front Physiol* 2018;9:727.
- [17] Mincholé A, Rodríguez B. Artificial intelligence for the electrocardiogram. *Nat Med* 01 2019;25(1):22–23.