

Diagnóstico del ritmo cardíaco durante la resucitación cardiopulmonar mediante técnicas de aprendizaje profundo

I. Isasi Liñero¹, L. López de Larrucea¹, U. Irusta Zarandona¹, E. Aramendi Ecenarro¹

¹BioRes, Departamento de ingeniería de comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {iraia.isasi, libe.lopezdelarrucea, unai.irusta, elisabete.aramendi }@ehu.eus

Resumen

Las compresiones torácicas inducen artefactos en el electrocardiograma (ECG) que dificultan el diagnóstico fiable del ritmo durante la reanimación cardiopulmonar (RCP). Esto obliga al rescatador a detener la terapia durante el análisis del ritmo comprometiendo la circulación y, por lo tanto, reduciendo la probabilidad de supervivencia del paciente. Este estudio propone un nuevo enfoque para la discriminación de ritmos desfibrilables y no desfibrilables durante compresiones torácicas. El método se basa en técnicas de aprendizaje profundo que han demostrado ser más precisas en múltiples problemas de clasificación en el ámbito de la ingeniería biomédica. Se usó una base de datos compuesta por 506 ritmos desfibrilables y 1697 no-desfibrilables adquiridos de pacientes en parada cardiorrespiratoria extrahospitalaria. El modelo fue entrenado y testeado mediante un esquema de validación cruzada de 10 iteraciones (10-fold CV). Este proceso se repitió 10 veces para caracterizar estadísticamente los resultados en términos de sensibilidad (SE), especificidad (SP) y precisión balanceada (BAC). Los resultados se compararon con los obtenidos utilizando técnicas de aprendizaje automático tradicionales. El método de aprendizaje profundo proporcionó un rendimiento similar al obtenido mediante el algoritmo tradicional con una SE, SP y BAC de 95.0 %, 96.1 % y 95.5%, respectivamente.

1. Motivación

Las compresiones torácicas representan una componente esencial de la reanimación cardiopulmonar (RCP) ya que tienen como objetivo mantener un flujo sanguíneo mínimo en los órganos vitales hasta la llegada de un desfibrilador externo automático (AED) [1]. Sin embargo, las compresiones a menudo hacen que el ECG no sea interpretable debido al artefacto de movimiento que produce errores de diagnóstico en el algoritmo *shock/no-shock* del AED. Por lo tanto, en la práctica actual, es necesaria la interrupción de las compresiones para analizar de forma fiable el ECG, lo que reduce el beneficio hemodinámico de la RCP y la probabilidad de supervivencia del paciente.

En los últimos años se han logrado importantes avances en el diagnóstico de ritmo durante compresiones. Las primeras estrategias en este ámbito de centraron en el uso de filtros adaptativos para eliminar el artefacto de compresión seguido de un algoritmo de decisión propio de un AED comercial para el diagnóstico *shock/no-shock* del ECG filtrado [2]. Sin embargo, estas soluciones no cumplían con los requisitos de rendimiento de la AHA (*American Heart Association*) para un diagnóstico fiable de ritmo. Los residuos del filtrado se identificaron como

principal factor de confusión en los algoritmos de diagnóstico comerciales, dado que éstos estaban originalmente diseñados para analizar ECG-s durante las pausas de compresión, es decir, sin artefacto [2]. De hecho, recientes estudios han demostrado que un análisis de ritmo fiable es posible si los algoritmos de decisión *shock/no-shock* se diseñan para clasificar el ECG en presencia de residuos de filtrado [3,4]. Este nuevo enfoque se centra en combinar características de discriminación *shock/no-shock* extraídas del ECG filtrado mediante técnicas de aprendizaje automático avanzadas como *Support Vector Machine* (SVM) [3] o *Random Forest* (RF) [4].

En el ámbito del aprendizaje supervisado, las técnicas de aprendizaje profundo han resultado ser muy eficientes para tareas de clasificación de señales biomédicas, aportando importantes avances incluso al campo de la resucitación cardiopulmonar [5,6]. En concreto, un estudio reciente ha demostrado que las técnicas de aprendizaje profundo presentan un mejor rendimiento que los algoritmos de aprendizaje automático tradicionales en el diagnóstico *shock/no-shock* del ritmo durante intervalos libres de compresiones torácicas [7]. Además, estas técnicas permiten que el propio algoritmo aprenda y seleccione las características de discriminación más relevantes que presenta la señal, eliminando así la necesidad de una extracción de características manual.

Sobre estos antecedentes, el objetivo de este trabajo es desarrollar un algoritmo de aprendizaje profundo para la discriminación de ritmos desfibrilables y no desfibrilables en presencia de los artefactos causados por la terapia RCP. El método se compone de un filtro adaptativo para suprimir el artefacto seguido de un algoritmo de diagnóstico *shock/no-shock* basado en técnicas de aprendizaje profundo.

2. Materiales

El conjunto de datos utilizado en este estudio fue obtenido por los servicios de emergencias de Londres, Estocolmo y Akershus (Noruega) entre Marzo de 2002 y septiembre de 2004. Las señales de ECG y de profundidad de compresión (PC) adquiridas mediante una versión modificada del desfibrilador Heartstart 4000 de Laerdal fueron exportadas a Matlab y remuestreadas a 250 Hz. Los instantes de las compresiones (t_k) se anotaron automáticamente en la señal de profundidad usando un detector de picos negativos con umbral de -1.5 cm.

Para este estudio se extrajeron automáticamente segmentos de 20 s con las siguientes características: un solo tipo de ritmo cardíaco durante todo el segmento, compresiones continuas durante un intervalo de 15 s y un intervalo de 5 s libre de compresiones ya sea antes o después del intervalo de compresiones (Figura 1). El intervalo de 15 s se utilizó para desarrollar y evaluar el algoritmo de decisión *shock/no-shock* mientras que los intervalos sin compresiones fueron usados por clínicos especialistas para anotar el ritmo subyacente del paciente (*ground truth*, GT).

La base de datos del estudio consta finalmente de 2133 segmentos pertenecientes a 272 pacientes de los cuales 506 son ritmos desfibrilables (principalmente fibrilaciones ventriculares, FV) y 1697 no-desfibrilables que incluyen pacientes en asistolia (AS) y con ritmos organizados (OR).

3. Métodos

3.1. Filtrado del artefacto RCP

Para suprimir el artefacto RCP se propone un filtro adaptativo RLS (*Recursive-Least-Squares*) para la eliminación de interferencias periódicas [8]. En resumen, durante las compresiones el artefacto se modela como una interferencia casi-periódica compuesta por 4 armónicos de frecuencia fundamental variable, $f_0(n)$:

$$s_{ct}(n) = A(n) \sum_{k=1}^4 a_k(n) \cos(k\omega_0(n)n) + b_k(n) \sin(k\omega_0(n)n) \quad (3)$$

$$f_0(n) = \frac{1}{t_k - t_{k-1}} \quad t_{k-1} < nT_s \leq t_k \quad (4)$$

donde la frecuencia angular discreta es $\omega_0(n) = 2\pi f_0(n)T_s$, T_s es el periodo de muestreo y $A(n)$ es una envolvente de amplitud que diferencia intervalos con compresiones ($A=1$) e intervalos sin compresiones ($A=0$).

El filtro RLS estima de forma adaptativa los coeficientes $a_k(n)$ y $b_k(n)$ y resta el artefacto estimado al ECG corrupto, $s_{cor}(n)$, para así obtener el ECG filtrado $\hat{s}_{ecg}(n)$, es decir, el ritmo subyacente del paciente. El ancho de banda del filtro RLS viene regulado por el factor de olvido λ , que se fijó a 0.999 [4].

3.2. Algoritmo de decisión *shock/no-shock*

El algoritmo de decisión *shock/no-shock* se diseñó para analizar tres ventanas consecutivas de 3 s dando un diagnóstico cada 9 s. El segmento de 9 s corresponde al intervalo 3-12 s del ECG filtrado (intervalo que se resalta en la Figura 1) y el diagnóstico se dio usando el criterio de voto por mayoría en las tres ventanas de análisis.

El algoritmo de aprendizaje profundo empleado en este trabajo para la decisión *shock/no-shock* se compone de dos bloques principales (Figura 2): una red neuronal convolucional (CNN) que se encarga de extraer características de alto nivel del ECG y una etapa de clasificación que haciendo uso de estas características produce una salida binaria. Los segmentos de 3 s se remuestrearon a 125 Hz para acelerar el proceso de entrenamiento y a su vez evitar el sobre ajuste de la red.

La CNN consta de tres sub-bloques de igual arquitectura. Utilizar varias capas convolucionales es una práctica común para extraer características cada vez más complejas. Cada CNN se compone de una capa convolucional que aplica 16 filtros de orden 16 al ECG seguido de una capa de normalización de *batch* y una función de activación lineal rectificadora (RELU). Se aplicó un mecanismo de *padding* para conservar la longitud de la señal a la salida de cada etapa convolucional. A continuación, se empleó una unidad de *max-pooling* para reducir a un tercio la dimensionalidad de las características extraídas por la CNN. Por último, la capa de *dropout* aplica una técnica de regularización que permite reducir el sobre ajuste de la red eliminando aleatoria y temporalmente neuronas de las capas internas de la red neuronal.

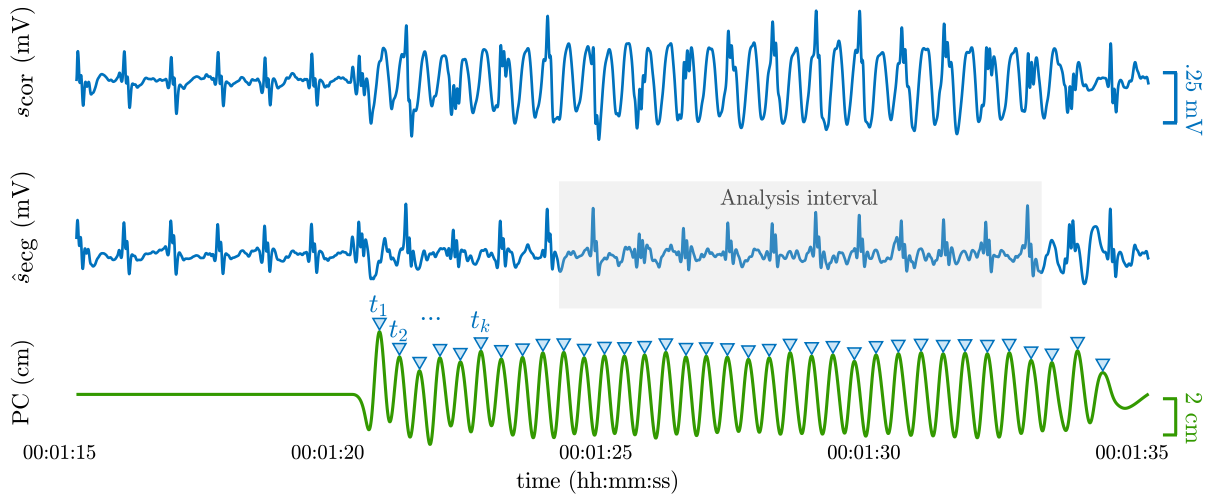


Figura 1. Un segmento de 20 s de duración que corresponde a un paciente con ritmo organizado (OR). Los primeros 5 s muestran el ritmo subyacente del paciente en un intervalo libre de compresiones en los primeros dos paneles. Los siguientes 15 s muestran el ECG artefactado en el primer panel y el ECG filtrado en el segundo panel. El panel inferior muestra la señal de profundidad de compresión (PC) con los instantes de compresión (t_k) marcados.

La última capa es la de clasificación. Mediante las características obtenidas se realiza la clasificación utilizando 10 neuronas de entrada y dos neuronas de salida para la clasificación *shock/no-shock*. Por último, se empleó la función *softmax* para normalizar los valores de las neuronas de salida de forma que las probabilidades sumen 1.

A modo de referencia se compararon los resultados con los obtenidos siguiendo un esquema de aprendizaje automático clásico implementado en [4] para el mismo propósito, el análisis *shock/no-shock* del ritmo durante la RCP. Este método se compone de un filtro adaptativo (RLS) para suprimir el artefacto de las compresiones, una etapa de extracción de características basada en la SWT (*Stationary Wavelet Transform*) y un clasificador *Random Forest* para el diagnóstico *shock/no-shock*. Se utilizaron las 25 características más seleccionadas por el *Random Forest* en [4] (Figura 6). Los hiperparámetros del RF, el número mínimo de observaciones por hoja y el número de características por división se fijaron a 3 y a 7, de acuerdo a los valores óptimos obtenidos en [4].

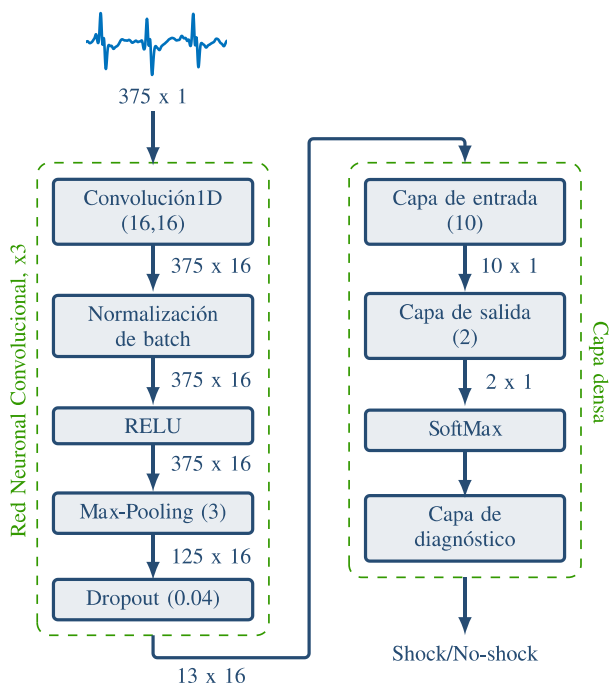


Figura 2. Arquitectura del algoritmo de decisión *shock/no-shock* basado en técnicas de aprendizaje profundo. Se indican las dimensiones de los datos a la salida de cada bloque.

3.3. Arquitectura y evaluación del modelo

Se utilizó un esquema de validación cruzada de 10 iteraciones (*10-fold CV*) para entrenar y testear el modelo. Las particiones se crearon de forma que las prevalencias de los ritmos coincidieran en la medida de lo posible con las prevalencias de los ritmos desfibrilables y no desfibrilables observados en la base de datos, y todos los segmentos asociados a un mismo paciente fueron asignados a la misma partición. Durante la fase de entrenamiento, se replicaron los segmentos de ritmos

desfibrilables (sin alteración) para reducir el desequilibrio entre clases (técnica de *data augmentation*).

Los diagnósticos del algoritmo de decisión *shock/no-shock* fueron comparados con las anotaciones de ritmo para obtener las siguientes métricas de rendimiento: proporción de ritmos desfibrilables correctamente identificados (Sensibilidad, SE), proporción de ritmos no-desfibrilables correctamente identificados (Especificidad, SP) y precisión balanceada (BAC) calculada como la media entre la SE y la SP. Este proceso se repitió sobre 10 particiones aleatorias para estimar la distribución estadística de las métricas de rendimiento.

4. Resultados

La tabla I muestra las distribuciones de las métricas de rendimiento obtenidas mediante el algoritmo propuesto (DL) y la solución basada en técnicas de aprendizaje automático (ML). En ambas soluciones se cumplen los requisitos de rendimiento de la AHA fijados para intervalos libres de compresiones: $SE > 90\%$ y $SP > 95\%$. Como se observa en la tabla I, el método basado en las técnicas de aprendizaje profundo presenta un BAC ligeramente superior al obtenido con la solución tradicional. Sin embargo, esta diferencia no es estadísticamente significativa, con un p-valor medio de 0.5 para el test de McNemar sobre las 10 particiones 10-fold CV de los datos. La función de costes en el test de McNemar fue modificada para triplicar la penalización de los ritmos desfibrilables, de forma que el test de McNemar comparase los clasificadores en términos de BAC.

	SE (%)	SP (%)	BAC (%)
DL	95.0 (94.4-95.1)	96.1 (95.7-96.4)	95.5 (95.1-95.7)
ML	92.7 (92.1-92.9)	97.5 (97.4-97.6)	95.1 (94.8-95.2)

Tabla 1. La mediana (rango interdecil) de las métricas de rendimiento.

La figura 7 muestra 4 ejemplos ilustrativos de dos segmentos correctamente clasificados (paneles a y b) y dos segmentos mal clasificados (paneles c y d) tanto para ritmos desfibrilables (paneles b y d) como para ritmos no desfibrilables (paneles a y c). Como se observa, los errores del clasificador se deben principalmente a residuos del filtrado. Los residuos que generan cambios bruscos en la amplitud del ECG (residuos con espigas) a menudo se interpretan como los complejos QRS de ritmos organizados en pacientes que se encuentran en FV (panel d), mientras que los residuos desorganizados de alta frecuencia son a veces mal interpretados como FV en pacientes con ritmos no-desfibrilables (panel c).

Aunque el método propuesto cumple con las especificaciones de la AHA, presenta un rendimiento similar a las técnicas de aprendizaje automático tradicionales. Sin embargo, los algoritmos de aprendizaje profundo mejoran su rendimiento conforme los datos de entrenamiento aumentan mientras que los algoritmos de aprendizaje automático no mejoran su precisión a partir de una determinada cantidad de datos. Estimamos que un

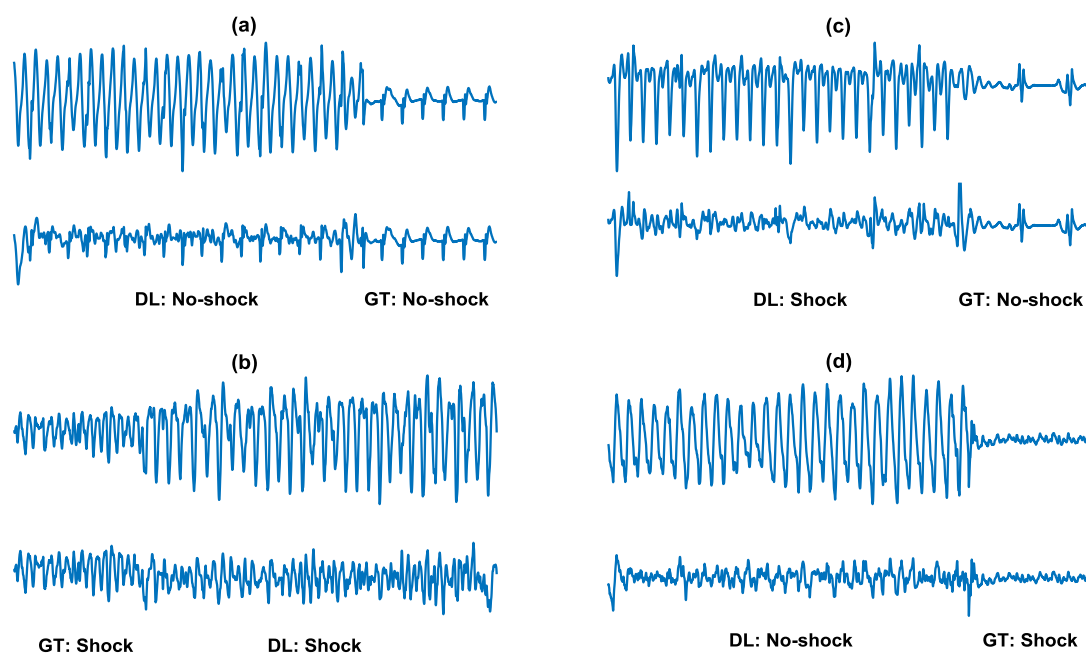


Figura 3. Cuatro segmentos de 20 s de duración (misma composición que la figura 1) que corresponden a dos pacientes en FV (b, d) y dos pacientes con ritmo organizado (a, c). Los primeros dos segmentos (a, b) son correctamente clasificados por el clasificador mientras que en los últimos dos segmentos (c, d) el algoritmo toma una decisión errónea.

aumento del 50% de la base de datos sería suficiente para que las técnicas de aprendizaje profundo superaran el rendimiento de las técnicas convencionales, lo cual es viable con las base de datos de las que ya disponemos. Además, las técnicas de aprendizaje profundo permiten eliminar la etapa de extracción de características, y nuestros resultados igualan ya a los obtenidos con las características que mayor precisión han proporcionado hasta la fecha.

5. Conclusiones

Este trabajo introduce un nuevo método para el diagnóstico *shock/no-shock* fiable del ritmo durante la RCP. Consiste en un filtro adaptativo RLS que suprime el artefacto de las compresiones seguido de un algoritmo de diagnóstico basado en técnicas de aprendizaje profundo. Este último se compone de 3 bloques convolucionales y una etapa de clasificación formada por una red neuronal clásica

Esta nueva estrategia para diagnosticar el ritmo durante compresiones podría incrementar el rendimiento y facilitar el desarrollo e implementación de las soluciones clásicas propuestas en la literatura, abriendo la posibilidad de que una RCP sin interrupciones sea implementable en la práctica real. La ausencia de interrupciones implica que el flujo sanguíneo generado por las compresiones torácicas no se vea comprometido durante la terapia RCP, incrementando la probabilidad de supervivencia del paciente.

Agradecimientos

Este trabajo ha recibido ayuda financiera del Ministerio de Ciencia, Innovación y Universidades, proyecto

RTI2018-101475-BI00, junto con el Fondo Europeo de Desarrollo Regional (FEDER). Ha recibido también financiación del Gobierno Vasco mediante la beca PRE-2019-2-0066 y la subvención a grupos IT-1229-19.

Referencias

- [1] Perkins GD et al. European Resuscitation council guidelines for resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation 2015. *Resuscitation*, vol 85, 2015, pp 81-89.
- [2] De Gauna S.R et al. Rhythm analysis during cardiopulmonary resuscitation: past, present, and future. *BioMed research international*, vol 2014, 2014: e386010.
- [3] Ayala U et al. A reliable method for rhythm analysis during cardiopulmonary resuscitation, *Biomed research international*, vol 2014, 2014: e872470.
- [4] Isasi I et al. Automatic Cardiac Rhythm Classification With Concurrent Manual Chest Compressions. *IEEE Access*, vol 7, 2019: e2935096.
- [5] Elola A et al. Deep Neural Networks for ECG-Based Pulse Detection during Out-of-Hospital Cardiac Arrest. *Entropy*, vol 21, 2019: e21030305.
- [6] Kwon J et al. Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes. *Resuscitation*, vol 139, 201, pp 84-91.
- [7] Picon A et al. Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia. *PloS one*, vol 14, 2019: e0216756.
- [8] Xiao Y et al. Fast RLS Fourier analyzers capable of accommodatin frequency mismatch. *Signal Process*, vol 87, 2007, pp 21.