

Algoritmo basado en redes neuronales recurrentes para la detección de ritmos desfibrilables

X. Jaureguibeitia¹, U. Irusta¹, E. Aramendi¹, B. Chicote¹, D. Alonso², C. Corcuera²,
J. Veintemillas², A. Larrea², M. Olabarria²

¹ BioRes, Dpto. de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {xabier.jaureguibeitia, unai.irusta, elisabete.aramendi, beatriz.chicote}@ehu.eus

² Emergentziak-Osakidetza, Servicio Vasco de Salud (Osakidetza) Bilbao, España, {danielfrancisco.alonsomoreno, carlos.corcuerabergado, josefrancisco.veintemillasvidal, andima.larrearedin, mikel.olabarriaibarreche}@osakidetza.eus

Resumen

La desfibrilación es crítica para la supervivencia en episodios de parada cardiorrespiratoria extra-hospitalaria (PCREH), por lo que los desfibriladores externos automáticos (DEA) incorporan algoritmos para la detección de ritmos desfibrilables. En este trabajo se emplearon 852 casos de PCREH con 4216 ventanas de análisis del ritmo del DEA. Seis expertos anotaron ECG como desfibrilable / no-desfibrilable y el diagnóstico mayoritario se empleó como referencia. Se obtuvieron 12185 segmentos ECG de 4s, que se partitionaron por paciente en subconjuntos para desarrollar (70%) y evaluar (30%) un algoritmo de análisis de ritmo. El algoritmo se compone de una etapa de descomposición subbanda y una red neuronal recursiva. El mejor modelo presentó una sensibilidad (Se) de 97.6% y una especificidad (Sp) de 98.3%, cumpliendo holgadamente con las especificaciones mínimas para uso clínico del 90% de Se y 95% de Sp.

1. Motivación

La parada cardiorrespiratoria extra-hospitalaria (PCREH) representa un problema de salud pública de primera magnitud, con una tasa de incidencia estimada de unos 55 casos por cada 100.000 habitantes y año, y ratios de supervivencia muy reducidos, de en torno al 10% [1]. Los primeros minutos de un episodio de PCREH son críticos para el éxito de la resucitación. Arritmias ventriculares letales tales como la fibrilación ventricular (FV) o la taquicardia ventricular (TV) se presentan con prevalencias por encima del 70%, y las tasas de supervivencia pueden alcanzar el 50% cuando se aplica terapia de desfibrilación [2]. Estos valores decrecen rápidamente con el tiempo. Por cada minuto que transcurre sin que una descarga sea aplicada, la probabilidad de supervivencia se reduce en un 3-4%, o hasta en un 10-12% si tampoco se administra terapia de resucitación cardiopulmonar (RCP) [3]. De forma similar, las arritmias ventriculares eventualmente degeneran en asistolia (AS) o actividad eléctrica sin pulso (AESP), ritmos inmunes a la desfibrilación y con peor pronóstico clínico en general. Atendiendo a todo esto, y considerando un tiempo medio de respuesta de varios minutos para los servicios médicos de emergencia (SME) [4], la intervención por parte de testigos presenciales y el acceso rápido a un desfibrilador externo automático (DEA) pueden resultar determinantes para la supervivencia del paciente.

La terapia de desfibrilación no puede ser administrada de forma arbitraria. Una descarga aplicada a un paciente con pulso efectivo puede desencadenar por sí misma un evento de parada. Aun cuando la patología del paciente se corresponde con PCREH, la desfibrilación de una AESP o una asistolia no solo es ineficaz, sino que conlleva también posibles daños en el miocardio y pausas innecesarias de la terapia RCP, lo que merma las opciones de supervivencia [5,6]. Cualquier intento de desfibrilación debe, por tanto, ir precedido de un análisis del estado cardíaco del paciente. En este sentido, y de cara a su empleo por parte de personas sin formación médica, los DEAs incorporan algoritmos automáticos para el análisis de ritmo y la discriminación de arritmias desfibrilables.

El desarrollo de estos algoritmos ha sido objeto de estudio durante décadas. La American Heart Association (AHA) estableció los mínimos de precisión exigibles en un 90% de sensibilidad para detección de FV (75% para TV) y un 95% de especificidad para ritmos no desfibrilables [7]. La literatura sobre detección de FV es extensa [8,9], pero pocos trabajos se han realizado empleando datos de casos reales de PCREH. En los últimos años, las técnicas de aprendizaje automático han permitido el uso conjunto de las características más relevantes de la bibliografía y su evaluación sobre datos PCREH [10]. Más recientemente aún, las técnicas de aprendizaje profundo han permitido abordar el problema sin mayor información que la propia señal ECG [11].

Este trabajo emplea datos reales de episodios PCREH y plantea un abordaje de clasificación novedoso, basado en redes neuronales recursivas y descomposición subbanda mediante transformada wavelet estacionaria (*stationary wavelet transform*, SWT).

2. Materiales

El trabajo se sustenta en los registros de 852 episodios de PCREH tratados por los SME de Osakidetza entre 2013 y 2015. En todos los casos se emplearon DEAs del modelo LP1000 (Stryker/PhysioControl, Redmond, WA, EEUU), que registra la señal ECG con ancho de banda 0.5–21Hz y una frecuencia de muestreo de 125Hz. Los datos de los DEAs se convirtieron a formato MATLAB (MathWorks, Natick, MA, EEUU) y se desarrolló una herramienta para

la visualización y anotación de los casos. Se identificaron todas las ventanas de análisis automático de ritmo de los DEAs. El tipo de ritmo observado en cada ventana fue anotado de forma independiente por 6 clínicos expertos en PCREH. Los ritmos se identificaron como VF o TV, para los casos desfibrilables; y para los no desfibrilables como AS o ritmo organizado (ORG), exigiendo a estos últimos un mínimo de 12 latidos por minuto. El consenso entre diagnósticos fue elevado, con kappa de Fleiss de 0.93 para la distinción entre ritmos desfibrilables y no desfibrilables. El diagnóstico mayoritario se estableció como referencia (*ground truth*). Se excluyeron aquellas ventanas con ritmos intermedios o no consensuados, transiciones de ritmo o calidad de señal especialmente degradada. Se obtuvieron 4216 ventanas, 498 de ritmos desfibrilables y 3718 de no desfibrilables. Por último, cada ventana fue dividida en segmentos no solapados de 4s, orientando así el diseño y evaluación de los modelos a pausas razonables en la terapia RCP. Como beneficios adicionales, se consiguieron tanto un aumento del volumen de datos como la normalización de la duración de los segmentos ECG, muy convenientes para las técnicas de aprendizaje profundo empleadas. Se obtuvieron un total de 12185 segmentos, 2244 de ritmos desfibrilables y 9941 de no desfibrilables. El desglose de ventanas y segmentos por tipo de ritmo se presenta en la Tabla 1.

	Ventanas		Segmentos	
	N	%	N	%
FV	461	10.93	2107	17.29
TV	37	0.88	137	1.12
AS	2418	57.35	5687	46.67
ORG	1300	30.83	4254	34.91
Defib.	498	11.81	2244	18.42
No defib.	3718	88.19	9941	81.58

Tabla 1. Base de datos. Distribución de ritmos

3. Métodos

3.1. Descomposición SWT

Los experimentos preliminares del estudio revelaron una mejor convergencia de la red neuronal recurrente cuando la entrada incluía múltiples canales. En este sentido, se optó por realizar una descomposición SWT de la señal ECG, lo que posibilitaría también el empleo de técnicas de cancelación de ruido.

La SWT se basa en la aplicación recursiva de dos filtros $h(n)$ y $g(n)$, de norma unitaria y longitud L , que se derivan de una wavelet madre específica, y satisfacen las condiciones de ortogonalidad interna y cruzada necesarias para constituir un par de filtros espejo en cuadratura (*quadrature mirror filter*, QMF). La respuesta frecuencial de estos filtros es simétrica respecto de $\pi/2$, la frecuencia central de la banda a descomponer.

Para una señal de entrada $s(n)$, la información de banda alta se obtiene por convolución con $g_1(n) = g(n)$ y se identifica como coeficientes de detalle de primer nivel

$d_1(n)$. De forma similar, la información de banda baja se recoge en los denominados coeficientes de aproximación de primer nivel $a_1(n)$, del filtrado con $h_1(n) = h(n)$. Tal y como se muestra en la Figura 1, niveles sucesivos de descomposición pueden ser obtenidos aplicando el mismo procedimiento a los coeficientes de aproximación del nivel previo, de modo que $d_j(n) = a_{j-1}(n) * g_j(n)$ y $a_j(n) = a_{j-1}(n) * h_j(n)$. Los coeficientes de los filtros $h_j(n)$ y $g_j(n)$ se obtienen de los nivel inmediatamente anterior mediante sobremuestreo de factor 2 por inserción de ceros, tal que para L longitud de $h(n)$ y $g(n)$:

$$f_j(n) = \begin{cases} f\left(\frac{n}{2^{j-1}}\right) & n = 2^{j-1}k; k = 1, 2, \dots, L \\ 0 & \text{En otro caso} \end{cases} \quad (1)$$

para $j \geq 1$ y $f \in \{h, g\}$.

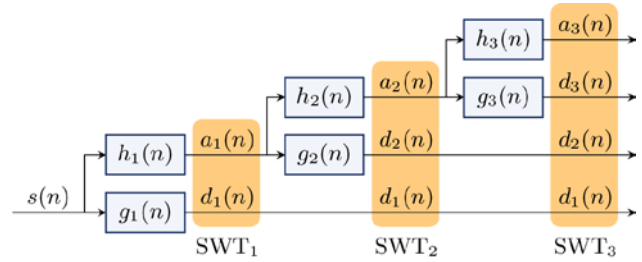


Figura 1. Diagrama SWT para 3 niveles de descomposición

En lo específico al modelo que se presenta, la señal de ECG fue submuestreada con factor 2 y luego descompuesta en 5 niveles SWT. El submuestreo inicial responde al filtrado interno del DEA, que elimina toda información en esas bandas. Los coeficientes de detalle obtenidos constituyen los canales de entrada a la red neuronal. La componente de aproximación a_5 , que contiene la información ≤ 1 Hz, no se consideró relevante y fue descartada. Los filtros de la SWT correspondieron a una wavelet Daubechies-2.

3.2. Cancelación de ruido

La cancelación de ruido se abordó empleando el método de umbralización universal [12], con el valor del umbral dado por:

$$\delta = \frac{\sqrt{2 \ln L}}{0.6745} \text{med}\{|d_1 - \text{med}\{d_1\}|\} \quad (2)$$

donde L es la longitud de la señal, en muestras. El ruido se cancela mediante *soft-thresholding*.

La estimación del nivel de ruido se realiza típicamente sobre el primer nivel de detalle de la SWT (d_1 en la ecuación (2)), asumiendo que este contiene principalmente ruido e información muy localizada sin apenas efecto en el cálculo de la mediana. Debido al submuestreo aplicado, este primer nivel se corresponde con la banda 16.125-31.25Hz y contiene información relevante de señal. Como alternativa, se obtuvieron los coeficientes d_1 de la señal original, sin aplicar submuestreo, los cuales corresponden a la banda de frecuencias 31.25-62.5Hz.

3.3. Red Neuronal Recurrente

El esquema general de la arquitectura de red propuesta se representa en la Figura 2. La red consta de 3 bloques: dos

bloques recurrentes bidireccionales basados en unidades GRU (*Gated Recurrent Unit*) [13] y un bloque final de clasificación. El primer bloque GRU recibe las secuencias de entrada: los 5 canales de descomposición SWT, con 250 muestras de longitud cada uno, y genera a su salida 48 secuencias de la misma longitud. El segundo bloque trabaja sobre estas secuencias, generando una única característica de salida por cada una de las 8 unidades que lo componen. Cada bloque GRU incluye una capa de *dropout*. Durante el entrenamiento, estas capas seleccionan de forma aleatoria una proporción p de unidades, que son desactivadas. La disposición variable de unidades fuerza a la red a distribuir la información, evitando así patrones muy específicos y reduciendo el sobreajuste. El bloque final de clasificación consta de una única neurona, conectada a cada una de las salidas del segundo bloque GRU. Esta neurona aplica una función lineal a las características obtenidas, y alimenta una función de activación sigmoide, obteniendo un valor entre 0 y 1 que puede interpretarse como una estimación de probabilidad. El diagnóstico del clasificador se obtiene entonces por umbralización simple, con umbral de 0.5.

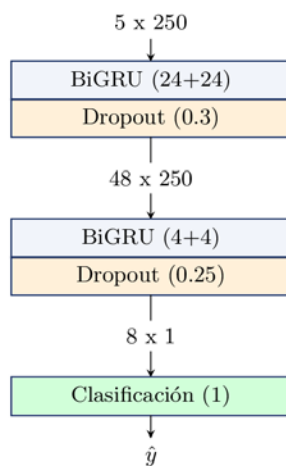


Figura 2. Arquitectura de la red neuronal recurrente

3.4. Evaluación del modelo

El conjunto de ventanas de análisis fue particionado por pacientes en grupos de entrenamiento (70%) y validación (30%). Esta división se realizó por muestreo aleatorio estratificado, conservando para cada grupo la proporción global de ritmos. Se crearon 100 particiones diferentes a fin de evaluar la variabilidad estadística del modelo. Cada segmento i fue etiquetado con $y_i \in \{0,1\}$ en función del ritmo anotado. La clase positiva $y = 1$ fue asignada a los ritmos desfibrilables (FV, TV), mientras que $y = 0$ se asoció a los no desfibrilables (AS, ORG).

A fin de evitar un sesgo hacia las ventanas más largas, con más segmentos asociados, cada segmento fue caracterizado con un peso w , igual al inverso del total de segmentos en su correspondiente ventana de origen. Estos pesos se utilizaron tanto para el entrenamiento de la red neuronal como para el cálculo de las métricas de rendimiento. Se definieron también dos modelos diferentes en función del balanceo de clases: un modelo balanceado (B), para el que se emplearon pesos adicionales con el fin de equiparar la

importancia de ambas clases, y un modelo empírico (E), sin balanceo.

Para cada uno de los experimentos se determinaron sobre el grupo de validación el número de verdaderos positivos (VP) y negativos (VN), así como los falsos positivos (FP) y negativos (FN). Se obtuvieron así la sensibilidad (Se), especificidad (Sp), precisión (Acc) y precisión balanceada (BAC):

$$Se = \frac{VP}{VP + FN} \quad Acc = \frac{VP + VN}{VP + FP + VN + FN}$$

$$Sp = \frac{VN}{VN + FP} \quad BAC = \frac{Se + Sp}{2}$$

4. Resultados

Los resultados obtenidos para las métricas de rendimiento se muestran en la Tabla 2.

	Modelo B	Modelo E
Se	97.6 (94.8 – 98.8)	95.5 (92.2 – 97.5)
Sp	98.3 (97.1 – 99.2)	99.1 (98.5 – 99.5)
Acc	98.2 (97.2 – 98.8)	98.6 (98.2 – 99.0)
BAC	97.7 (96.9 – 98.4)	97.4 (95.8 – 98.1)

Tabla 2. Valor mediano (rango interdecil) de las métricas de rendimiento sobre las 100 particiones utilizadas.

El modelo empírico presentó valores de especificidad y precisión significativamente más altos, a cambio de una caída en sensibilidad, resultado previsible dado el fuerte desbalanceo de clases.

Se analizó también la precisión de los modelos para los subtipos anotados dentro de las clases. Los resultados de este análisis se recogen en la Tabla 3.

	Modelo B	Modelo E
FV	98.8 (97.0 – 99.7)	97.1 (94.7 – 98.5)
TV	81.9 (61.7 – 96.1)	75.5 (54.4 – 90.3)
AS	99.0 (98.0 – 99.6)	99.6 (99.2 – 99.9)
ORG	97.1 (95.4 – 98.6)	98.3 (97.1 – 99.2)

Tabla 3. Precisión mediana (rango interdecil) para los ritmos concretos diferenciables dentro de las clases.

Se vuelve a observar cómo el modelo empírico favorece la detección de ritmos no desfibrilables. Destacan, por otro lado, los valores para AS y TV. Si bien son mayores en el modelo empírico, las métricas para AS se mantienen elevadas para ambos modelos. Esto es probablemente debido a la fuerte presencia de la AS en la base de datos (ritmo mayoritario, 57% de las ventanas). El caso opuesto sucede para la TV, para la que solo se dispuso de 37 ventanas (ritmo minoritario, <1% de las ventanas).

A modo ilustrativo, en la Figura 3 se muestran una serie de ejemplos, uno para cada tipo de ritmo, de diagnósticos acertados y errados por ambos modelos (la partición fue escogida al azar).

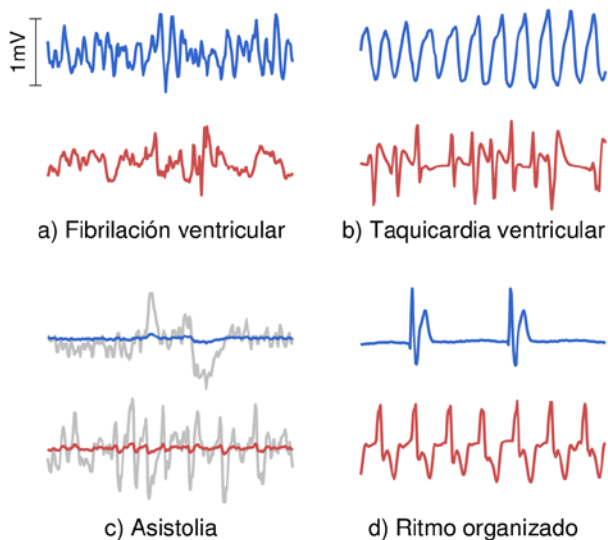


Figura 3. Clasificaciones correctas (azul) e incorrectas (rojo) para cada tipo de ritmo, en segmentos de 4s. Se ofrece una versión ampliada ($\times 10$) de la AS para apreciar el detalle.

5. Conclusiones

El algoritmo cumplió holgadamente las especificaciones generales de rendimiento marcadas por la AHA (90% de Se y 95% de Sp). Falló, no obstante, en especificaciones menores relacionadas con la TV. Si bien el valor medio se mantuvo en el 75% de Se exigido, el valor al 90% de confianza no alcanzó el 67%. Este problema es fácilmente atribuible a la escasa disponibilidad de segmentos de TV, y se antoja sencillo de subsanar mediante técnicas de *data augmentation*. Tampoco fue posible evaluar la Sp mínima de 99% para ritmo sinusal normal (subtipo de ORG), pues estos ritmos no se presentan durante PCREH. Si bien el número de registros empleado se mostró suficiente para generar modelos efectivos, un mayor volumen de datos podría ayudar a reducir la variabilidad de los resultados.

En términos más generales, el modelo es aún susceptible de ciertas mejoras. El filtrado aplicado por el DEA impidió un uso ideal de la umbralización universal, con lo que otros métodos de cancelación de ruido podrían ser explorados. De forma similar, la fuerte presencia de AS, caracterizada por niveles muy bajos de señal, impidió el uso de los métodos habituales de normalización, algo muy deseable en redes recursivas dado el uso extensivo de funciones sigmoide y tangente hiperbólica. No se realizó análisis alguno de las posibles wavelets en las que basar la SWT, aunque estudios anteriores sugieren que no es relevante [14]. Tampoco se analizó la importancia relativa de los diferentes niveles de detalle. La arquitectura de la propia red queda también abierta a un ajuste más fino de los parámetros y a técnicas de regularización, así como a la optimización con técnicas de validación cruzada.

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), y del Gobierno Vasco mediante la subvención a grupos de investigación IT-1229-19.

Referencias

- [1] Berdowski J, Berg RA, Tijssen JGP, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: Systematic review of 67 prospective studies. *Resuscitation*, vol 81, 2010, pp 1479-1487.
- [2] Berdowski J, Blom MT, Bardai A. Impact of onsite or dispatched automated external defibrillator use on survival after out-of-hospital cardiac arrest. *Circulation*, vol 124, 2011, pp 2225-2232.
- [3] Perkins GD, Handley AJ, Koster RW, et al. European Resuscitation Council Guidelines for Resuscitation 2015 Section 2. Adult basic life support and automated external defibrillation. *Resuscitation*, vol 95, 2015, pp 81-99.
- [4] Daya MR, Schmicker RH, Zive DM, et al. Out-of-hospital cardiac arrest survival improving over time: Results from the Resuscitation Outcomes Consortium (ROC). *Resuscitation*, vol 91, 2015, pp 108-115.
- [5] Xie J, Weil MH, Sun S, et al. High-energy defibrillation increases the severity of postresuscitation myocardial dysfunction. *Circulation*, vol 96, 1997, pp 683-688.
- [6] Cheskes S, Schmicker RH, Christenson J, et al. Perishock pause: an independent predictor of survival from out-of-hospital shockable cardiac arrest. *Circulation*, vol 124, 2011, pp 58-66.
- [7] Kerber RE, Becker LB, Bourland JD, et al. Automatic External Defibrillators for Public Access Defibrillation: Recommendations for Specifying and Reporting Arrhythmia Analysis Algorithm Performance, Incorporating New Waveforms, and Enhancing Safety. *Circulation*, vol 95, 1997, pp 1677-1682.
- [8] Amman A, Tratnig R, Unterkofler K. Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators. *Biomedical Engineering Online*, vol 4, 2005.
- [9] Aramendi E, Irusta U, Pastor E, et al. ECG spectral and morphological parameters reviewed and updated to detect adult and paediatric life-threatening arrhythmia. *Physiological Measurement*, vol 31, 2010, pp 749-761.
- [10] Figuera C, Irusta U, Morgado E, et al. Machine Learning Techniques for the Detection of Shockable Rhythms in Automated External Defibrillators. *PLoS one*, vol 11 2016.
- [11] Picón A, Irusta U, Álvarez-Gila A, et al. Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia. *PLoS one*, vol 14, 2019.
- [12] Donoho DL. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, vol 41, 1995, pp 613-627.
- [13] Cho K, van Merriënboer B, Gulcehre C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translations. *arXiv* 1406.1078v3, 2014.
- [14] Isasi I, Irusta U, Elola A, et al. A Machine Learning Shock Decision Algorithm for Use During Piston-Driven Chest Compressions. *IEEE transactions on bio-medical engineering*, vol 66, 2019, pp 1752-1760.