

# Arquitecturas de aprendizaje profundo para la detección de pulso en la parada cardiaca extrahospitalaria utilizando el ECG

A. Elola<sup>1</sup>, E. Aramendi<sup>1</sup>, U. Irusta<sup>1</sup>, A. Picón<sup>2</sup>, E. Alonso<sup>1</sup>

<sup>1</sup> Departamento de ingeniería de comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, andoni.elola@ehu.eus

<sup>2</sup> TECNALIA Research & Innovation, Derio, Vizcaya, España

## Resumen

*La detección de la presencia de pulso durante la parada cardiorrespiratoria extrahospitalaria es crucial para la supervivencia del paciente. Se ha demostrado que la toma manual del pulso no es muy fiable y que consume demasiado tiempo, por lo que es necesario desarrollar métodos automáticos que ayuden en la identificación del retorno de la circulación espontánea del paciente en parada. En este trabajo se propone utilizar técnicas de aprendizaje profundo para la discriminación automática de ritmos con pulso (PR) y sin pulso (PEA) utilizando solamente información proveniente del ECG. Se ha utilizado una base de datos que contiene 3914 segmentos de 5 segundos (3372 PR y 1542 PEA), que se dividieron en dos bases de datos con pacientes disjuntos para la optimización y evaluación de los métodos. Los mejores resultados se han obtenido utilizando una red neuronal profunda que contiene dos etapas de convolución y una etapa recurrente para la extracción de características y a continuación un clasificador. El modelo se evalúa en términos de sensibilidad (SE, porcentaje de PRs correctamente detectados) y especificidad (SP, proporción de PEAs correctamente detectados). Sobre la base de evaluación se obtuvieron una SE/SP de 94.2%/91.0%, por lo que puede concluirse que la detección automática del pulso utilizando sólo el ECG es viable mediante técnicas de aprendizaje profundo.*

## 1. Motivación

La parada cardiorrespiratoria sigue siendo una causa importante de muerte en los países industrializados. La resucitación cardiopulmonar mediante compresiones torácicas y ventilaciones, junto con la desfibrilación temprana constituyen el tratamiento fundamental de la parada cardiorrespiratoria extrahospitalaria [1]. La detección de pulso es esencial tanto para identificar la parada, como para la detección de la recuperación de circulación espontánea. Las dificultades del personal lego y de los clínicos para identificar la presencia de pulso [2]–[4] ha impulsado el desarrollo de métodos automáticos de detección de pulso, tanto para el reconocimiento de la parada como para la identificación del RCE.

Los desfibriladores externos automáticos (DEAs) son aparatos diseñados para ser utilizados por personal mínimamente entrenado. Dichos dispositivos adquieren dos señales a través de los parches de desfibrilación: el electrocardiograma (ECG) y la impedancia torácica. En los últimos años se han desarrollado métodos para la detección de pulso mediante el procesado de ambas señales [5]–[7]. Desafortunadamente, la mayoría de los DEAs comerciales no adquieren la señal de impedancia torácica con suficiente

resolución para que los métodos propuestos sean aplicables. Sin embargo, el ECG se adquiere con suficiente resolución para realizar el análisis del ritmo cardiaco, y su potencial para la detección del pulso en el contexto de la parada cardiorrespiratoria no se ha explorado en profundidad.

Las técnicas de aprendizaje profundo han demostrado ser muy eficientes para tareas de clasificación de señales biomédicas [8], incluso durante la parada cardiorrespiratoria [9]. En este trabajo se evalúan diferentes arquitecturas de aprendizaje profundo para la discriminación de ritmos organizados con pulso (PR) y ritmos organizados sin pulso (PEA).

## 2. Base de datos

La base de datos se ha construido a partir de una base de 1015 episodios adquiridos de pacientes en parada cardiorrespiratoria extrahospitalaria en Dallas (Texas) por *DFW Center for Resuscitation Research*. Todos los episodios se adquirieron utilizando el dispositivo *Philips HeartStart MRx*, el cual adquiere el ECG con una frecuencia de muestreo de 250 Hz y una resolución de 1.03  $\mu$ V por bit menos significativo.

Un total de 1015 episodios contenían el ECG y la impedancia torácica, y basándose en la información clínica se separan en dos grupos: pacientes que han recuperado pulso y pacientes que no han recuperado pulso. En el primer grupo se dispone de la información del instante en el que el paciente ha recuperado pulso. Dicho instante fue anotado por el clínico en escena. Los pacientes que no recuperaron pulso y fueron transportados al hospital fueron descartados.

Los segmentos con una duración de 5 segundos y con complejos QRS discernibles fueron automáticamente detectados y se revisaron por tres ingenieros biomédicos. De los pacientes que recuperaron el pulso se extrajeron los segmentos PR, a partir del instante de recuperación de circulación espontánea anotado por el clínico. De los pacientes que no recuperaron el pulso se extrajeron los segmentos PEA. Entre dos PR consecutivos se deja una distancia mínima de 30 segundos, y entre dos PEA consecutivos una distancia mínima de 1 segundo.

La base de datos final contiene 279 pacientes (134 recuperaron el pulso), un total de 3911 segmentos (2372 PR y 1542 PEA). Estos datos se dividieron en dos

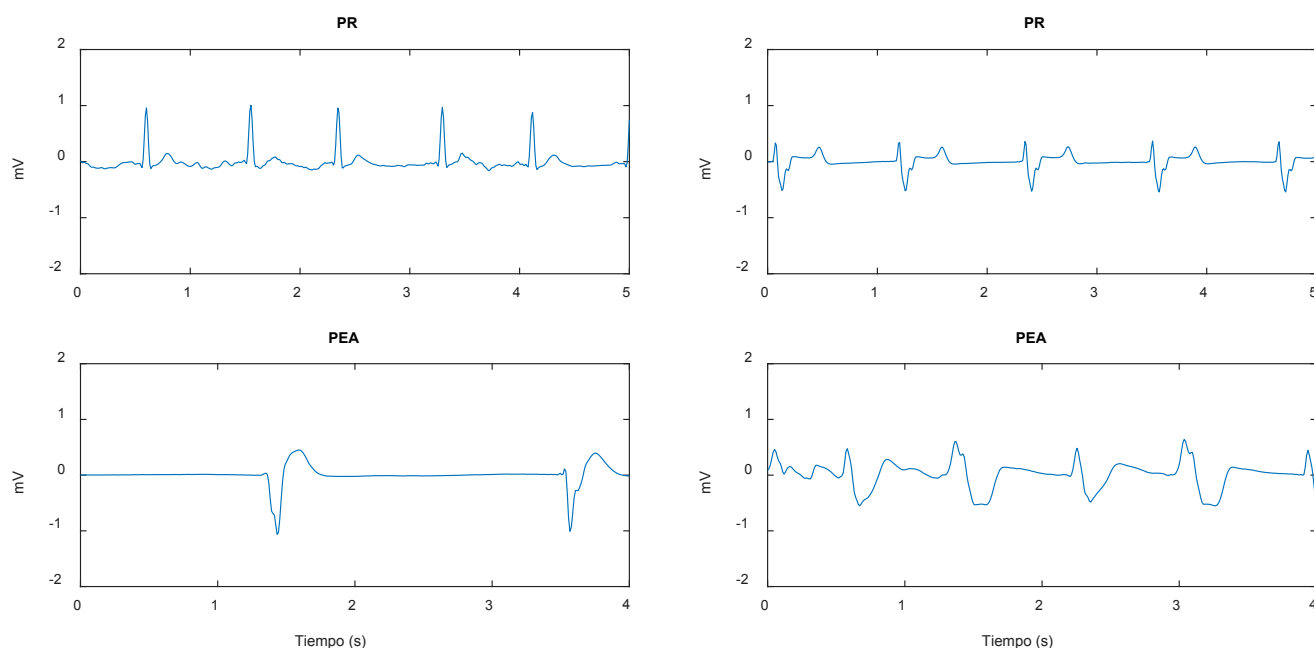


Figura 1. Ejemplos de ritmos PR y PEA

conjuntos con pacientes disjuntos: la base de entrenamiento (1871 PR y 1167 PEA) y la de testeo (501 PR y 375 PEA).

La Figura 1 muestra varios ejemplos de ritmos PR y PEA. En general, los ritmos PR se caracterizan por una mayor frecuencia cardíaca y unos complejos QRS más estrechos que en los ritmos PEA.

### 3. Métodos

La señal de ECG se filtró entre 0.5 y 30 Hz, para reproducir el ancho de banda de los desfibriladores actuales. Después, se remuestreó a 100 Hz.

En este trabajo se evalúan tres arquitecturas de aprendizaje profundo que se detallan a continuación.

#### 3.1. Arquitectura 1: CNN

La primera solución es una red convolucional (CNN), cuya estructura se resume en la Tabla 1.

Como método para conseguir incrementar la variabilidad de las señales disponibles, el primer paso es añadir ruido gaussiano de media cero y desviación estándar  $\sigma$  a la señal. Este proceso se da solamente durante la fase de entrenamiento y es útil para evitar el sobreajuste de la red.

Las capas convolucionales aplican a la señal  $n$  diferentes filtros de orden  $m$  para obtener descriptores de alto nivel del ECG. La segunda capa es la primera capa convolucional, que aplica 16 filtros de orden 4 y una función de activación RELU (unidad lineal rectificadora) en la salida.

La tercera capa realiza la compresión temporal del ECG para reducir el número de características. En nuestro caso hemos utilizado una capa *max-pooling*, la cual remuestrea la señal cogiendo el máximo valor de cada 3 muestras.

La cuarta capa, *dropout*, está presente sólo en la fase de entrenamiento. Esta capa anula ciertas unidades con cierta probabilidad, lo que permite que la red actúe como un conjunto de clasificadores simultáneos [10].

La quinta capa es otra capa convolucional. Utilizar varias capas convolucionales es una práctica común para extraer características cada vez más complejas. La salida se vuelve a remuestrear y se le aplica *dropout* otra vez (capas 6 y 7).

La última capa es la etapa de clasificación. Mediante las características obtenidas se realiza la clasificación utilizando una sola neurona con la función de activación de sigmoide. A la salida de esta neurona se obtiene la probabilidad de que el segmento sea PR.

# capa	Tipo	Parámetros
1	Ruido gaussiano	$\sigma=0.01$
2	Convolucional	$n=16, m=4$
3	Max-pooling	3
4	Dropout	0.4
5	Convolucional	$n=8, m=4$
6	Max-pooling	3
7	Dropout	0.4
8	Dense	-

Tabla 1. Estructura de la primera arquitectura propuesta

#### 3.2. Arquitectura 2: RNN

Las redes recurrentes (RNN) son capaces de exhibir el comportamiento temporal dinámico de una secuencia, por lo que son muy adecuadas para el procesamiento de ECG. La estructura de la arquitectura propuesta se resume en la Tabla 2.

Al igual que en la anterior solución, la primera capa se encarga de añadir ruido blanco a la señal durante el proceso de entrenamiento para evitar que se sobreajuste.

La segunda capa es una capa GRU (*Gated Recurrent Unit*), una versión simplificada del clásico LSTM [11] (*long short-term memory*) con una precisión similar, y son capaces de aprender dependencias temporales de largo plazo [12]. En nuestra solución hemos utilizado GRUs bidireccionales, de tal modo que puedan aprender las dependencias temporales hacia delante y hacia atrás. El único parámetro es la cantidad de unidades de cada una.

Tras la primera capa recurrente se aplica *dropout* y se aplica otra capa recurrente que reduce el número de características antes de la etapa de clasificación.

# capa	Tipo	Parámetros
1	Ruido gaussiano	$\sigma=0.01$
2	GRU bidireccional	24 + 24
3	Dropout	0.25
4	GRU bidireccional	4 + 4
5	Dropout	0.4
6	Dense	-

**Tabla 2.** Estructura de la segunda arquitectura propuesta

### 3.3. Arquitectura 3: CNN + RNN

La tercera arquitectura propuesta (Tabla 3) combina capas recurrentes y capas convolucionales. Primero se aplican varias capas convolucionales para la extracción de descriptores de alto nivel. Después se utiliza una GRU bidireccional para exhibir el comportamiento temporal de la señal, y por último las características extraídas por estas capas son utilizadas para clasificar el segmento como PR o PEA.

# capa	Tipo	Parámetros
1	Ruido gaussiano	$\sigma=0.01$
2	Convolutacional	$n=16, m=4$
3	Max-pooling	3
4	Dropout	0.4
5	Convolutacional	$n=8, m=4$
6	Max-pooling	3
7	Dropout	0.4
8	GRU bidireccional	4 + 4
9	Dropout	0.4
10	Dense	-

**Tabla 3.** Estructura de la tercera arquitectura propuesta

### 3.4. Parámetros de entrenamiento

Todos los experimentos se han desarrollado utilizando el paquete Keras [13]. Para evitar el sobreajuste, las capas

convolucionales y las capas recurrentes se han entrenado con la limitación de máxima norma, la norma de los pesos no podía exceder el valor de 4. Esta técnica es especialmente útil cuando se utiliza el *dropout* [10].

Todas las redes se han entrenado utilizando el optimizador Adam con los parámetros por defecto [13] y se han escogido 150 épocas. Para entrenar la red todos los pacientes se pesan equitativamente.

### 3.5. Evaluación de los modelos

Los modelos son evaluados en términos de sensibilidad (SE, probabilidad de detectar correctamente los ritmos PR), especificidad (SP, probabilidad de detectar correctamente los ritmos PEA) y precisión balanceada (BAC, media aritmética entre SE y SP). Para calcular todas las métricas de rendimiento todos los pacientes se pesan equitativamente.

## 4. Resultados

La SE, SP y BAC de cada modelo para la base de entrenamiento y la base de testeo se muestran en la Tabla 5. Puede observarse que los mejores resultados en términos de BAC para la base de testeo son obtenidos por la tercera arquitectura.

	Entrenamiento			Testeo		
	SE	SP	BAC	SE	SP	BAC
<b>CNN</b>	92.7	99.3	96.0	91.3	91.1	91.2
<b>RNN</b>	96.7	99.1	97.6	91.3	91.1	91.2
<b>CNN+RNN</b>	95.8	96.8	96.3	94.2	91.0	92.6

**Tabla 4.** Resultados obtenidos para las tres arquitecturas en la base de entrenamiento y en la base de testeo

## 5. Discusión y conclusiones

La detección de pulso sigue siendo una tarea difícil en la parada cardíaca extrahospitalaria tanto para el personal lego como para profesionales. En este trabajo se propone utilizar técnicas de aprendizaje profundo para la detección del pulso utilizando solamente el ECG. Esto permite el uso universal del algoritmo en los DEA, ya que la adquisición del ECG es necesaria para analizar el ritmo del paciente.

Se han probado tres arquitecturas diferentes: CNN, RNN y CNN+RNN. La tercera ha demostrado la mejor precisión en la base de testeo, siendo la diferencia de BAC 3.7% entre las bases de entrenamiento y testeo. Durante la fase de entrenamiento la red recurrente ha demostrado ser la más eficiente, pero el BAC baja 6.5 puntos en la base de testeo. Utilizar más datos, ajustar mejor los términos de regularización y utilizar técnicas de acrecimiento de datos podría permitir reducir la diferencia de eficiencia que hay entre la base de entrenamiento y la base de testeo.

Los resultados obtenidos mediante técnicas de aprendizaje profundo (SE/SP=94.2%/91.0%) superan los reportados por técnicas de aprendizaje automático (SE/SP=90.7%/87.5%) exclusivamente con el ECG [14]. El rendimiento es además comparable al reportado por métodos que combinaban parámetros del ECG y de la

impedancia [5], [6], para los que el BAC era del 92%. Nuestra solución sólo se basa en el ECG y eso permite su uso universal en DEAs.

## Agradecimientos

Este trabajo ha recibido apoyo económico conjunto del Ministerio de Economía y Competitividad Español y del Fondo Europeo de Desarrollo Regional (FEDER) a través del proyecto (TEC2015-64678-R), de la Universidad del País Vasco/Euskal Herriko Unibertsitatea mediante la ayuda a grupos de investigación GIU17/031, y del Gobierno Vasco a través de la beca PRE\_2017\_1\_0112.

## Referencias

- [1] J. Soar *et al.*, «European resuscitation council guidelines for resuscitation 2015: section 3. Adult advanced life support», *Resuscitation*, vol. 95, pp. 100–147, 2015.
- [2] B. Eberle, W. Dick, T. Schneider, G. Wisser, S. Doetsch, y I. Tzanova, «Checking the carotid pulse check: diagnostic accuracy of first responders in patients with and without a pulse», *Resuscitation*, vol. 33, n.º 2, pp. 107–116, 1996.
- [3] M. Ruppert *et al.*, «Checking for breathing: evaluation of the diagnostic capability of emergency medical services personnel, physicians, medical students, and medical laypersons», *Ann. Emerg. Med.*, vol. 34, n.º 6, pp. 720–729, 1999.
- [4] J. Tibballs y P. Russell, «Reliability of pulse palpation by healthcare personnel to diagnose paediatric cardiac arrest», *Resuscitation*, vol. 80, n.º 1, pp. 61–64, 2009.
- [5] E. Alonso *et al.*, «Circulation detection using the electrocardiogram and the thoracic impedance acquired by defibrillation pads», *Resuscitation*, vol. 99, pp. 56–62, 2016.
- [6] M. Risdal, S. O. Aase, J. Kramer-Johansen, y T. Eftesol, «Automatic identification of return of spontaneous circulation during cardiopulmonary resuscitation», *IEEE Trans. Biomed. Eng.*, vol. 55, n.º 1, pp. 60–68, 2008.
- [7] J. M. Ruiz *et al.*, «Circulation assessment by automated external defibrillators during cardiopulmonary resuscitation», *Resuscitation*, vol. 128, pp. 158-163, jul. 2018.
- [8] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, y U. R. Acharya, «Deep learning for healthcare applications based on physiological signals: a review», *Comput. Methods Programs Biomed.* pp. 1-13, 2018.
- [9] A. Picon *et al.*, *Detección de fibrilación ventricular mediante técnicas de aprendizaje profundo*. CASEIB 2017: XXXV Congreso anual de la Sociedad Española de Ingeniería Biomédica: Libro de actas, Bilbao 29 de Noviembre-1 de Diciembre, pp. 3-6, 2018.
- [10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, y R. Salakhutdinov, «Dropout: A simple way to prevent neural networks from overfitting», *J. Mach. Learn. Res.*, vol. 15, n.º 1, pp. 1929–1958, 2014.
- [11] S. Hochreiter y J. Schmidhuber, «Long short-term memory», *Neural Comput.*, vol. 9, n.º 8, pp. 1735–1780, 1997.
- [12] J. Chung, C. Gulcehre, K. Cho, y Y. Bengio, «Empirical evaluation of gated recurrent neural networks on sequence modeling», *ArXiv Prepr. ArXiv14123555*, 2014.
- [13] F. Chollet *et al.*, *Keras*. 2015.
- [14] A. Elola *et al.*, *Detección del retorno de la circulación espontánea en base al electrocardiograma*. CASEIB 2016: XXXIV Congreso anual de la Sociedad Española de Ingeniería Biomédica: Libro de actas, Valencia 23-25 de Diciembre, pp. 560-563, 2016 .