

Análisis de la Variabilidad del Ritmo Cardíaco para la Predicción de la Parada Cardíaca Extrahospitalaria Recurrente

A. Elola¹, E. Rueda¹, N. Amezaga¹, E. Aramendi¹, U. Irusta¹

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

Resumen

La parada cardíaca extrahospitalaria recurrente se produce cuando el paciente que tras recuperar el pulso al ser tratado por los servicios de emergencias, recae en una segunda parada prehospitalaria. Dicho fenómeno se ha asociado con menores probabilidades de supervivencia, y su predicción puede ser vital en la toma de decisiones sobre el tratamiento post parada. El objetivo de este trabajo es analizar la relación entre diferentes descriptores de la variabilidad del ritmo cardíaco calculados en el ECG y la parada recurrente. Para ello, se consideraron 98 pacientes (41 con parada recurrente) y se extrajeron 18 descriptores calculados en la ventana contigua a la recuperación de la circulación espontánea. Para la clasificación se utilizaron un clasificador Random Forest y se aplicaron 50 veces la validación cruzada con 5 grupos. Los mejores resultados se obtuvieron para ventanas de 2 minutos, con una mediana AUC (área bajo la curva) de 0.75. Los predictores más importantes para dicho modelo fueron los descriptores del diagrama de Poincaré, SD2 y la relación SD1/SD2, así como la frecuencia asociada al pico máximo en la banda HF (0.15-0.4 Hz).

1. Introducción

La parada cardiorrespiratoria extrahospitalaria es un problema de salud pública en los países industrializados. Los servicios de emergencias tratan más de 275000 pacientes al año debido a una parada a nivel europeo [1], y se estima que sólo se presencia un 60% de los casos [2]. La terapia básica incluye compresiones torácicas y ventilaciones para mantener un flujo mínimo de oxigenación en los órganos vitales, y la desfibrilación eléctrica, cuyo propósito es restaurar la funcionalidad del corazón. El objetivo principal de la terapia es la recuperación de la circulación espontánea (RCE). Una vez alcanzado este punto se comienza con la terapia de post-resucitación [3], la cual es uno de los eslabones clave para la supervivencia del paciente, e incluye el transporte del paciente al hospital.

Sin embargo, un porcentaje importante de pacientes, cercano al 24% [4], sufre una segunda parada con pérdida de pulso tras la RCE. Este fenómeno conocido como parada recurrente, PaR, a pesar de estar asociado con una peor probabilidad de supervivencia, no está asociado a unas causas conocidas y son pocos los métodos automáticos desarrollados para su predicción [5]. Un sistema que ayudara a cuantificar las probabilidades de PaR de un paciente permitiría optimizar el tratamiento post-parada suministrado. Por ejemplo, podría determinar el tratamiento farmacológico, o si el paciente debe ser transportado al hospital, ya que una parada dentro de la ambulancia dificulta y retrasa la terapia de resucitación.

En este estudio se aplican técnicas de aprendizaje automático para la predicción de la PaR una vez que el paciente haya recuperado el pulso. Para la extracción de características se utilizan parámetros de variabilidad del ritmo cardíaco calculados a partir de la forma de onda del electrocardiograma (ECG) y se analiza la importancia de cada uno con el fin de entender mejor qué es lo que influye en la PaR.

2. Métodos

2.1. Base de datos y proceso de anotación

La base de datos se construyó a partir de un registro de 1015 episodios adquiridos en pacientes en parada cardiorrespiratoria extrahospitalaria en Dallas (Texas) por DFW Center for Resuscitation Research. Todos los episodios se registraron utilizando el dispositivo Philips HeartStart MRx, el cual adquiere el ECG mediante los parches de desfibrilación con una frecuencia de muestreo de 250 Hz y una resolución de 1.03 μ V por bit menos significativo.

Gracias a la información clínica disponible se conoce si el paciente recuperó el pulso y el instante en el que el paciente recuperó el pulso (t_{RCE} anotado por el rescatador en escena). El criterio de inclusión exige un pulso mantenido al menos durante 2 minutos.

El criterio se cumplió en 98 casos (de los cuales 41 sufrieron una PaR), y se analizaron los 12 minutos posteriores a t_{RCE} . Si el paciente sufre una PaR en esos 12 minutos se anota como 1 (PaR), y en el caso contrario se anota como 0 (no hay PaR). La PaR se identifica cuando se observa un ritmo cardíaco maligno en el ECG (asistolia, actividad eléctrica sin pulso o fibrilación ventricular) o cuando el rescatador reinicia la terapia de reanimación con compresiones torácicas.

La Figura 1 muestra el ejemplo de un paciente que cumplió el criterio de inclusión ($t_{PaR}-t_{RCE}>2$ min). Dicho paciente recuperó la circulación espontánea tras recibir la terapia de resucitación, pero al cabo de dos minutos y medio muestra una fibrilación ventricular, que es una arritmia que causa la muerte si no se trata inmediatamente.

2.2. Extracción de características

Para la extracción de características, se consideró una ventana de duración t_w después de t_{RCE} tal y como se muestra en la Figura 1, y se extrajeron 18 descriptores de la variabilidad del ritmo cardíaco por ventana y paciente. El análisis se realizó para $t_w=120$ s y $t_w=60$ s.

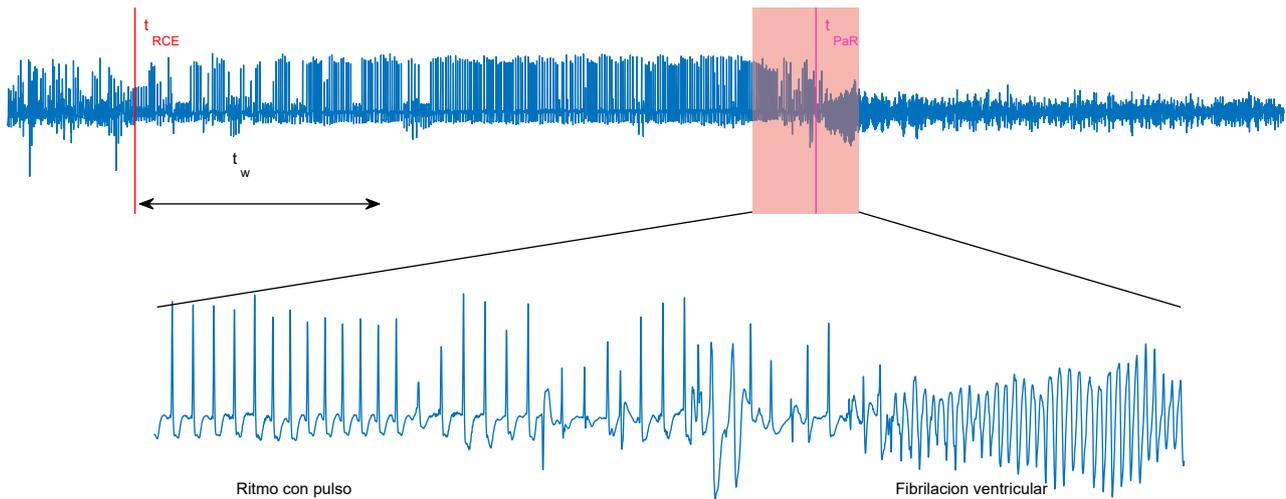


Figura 1. Forma de onda del ECG de 5 minutos y 30 segundos de un paciente que sufre una parada recurrente. El paciente recupera la circulación espontánea en el instante t_{RCE} y lo pierde en el instante t_{PaR} (aproximadamente dos minutos y medio más tarde)

La señal de ECG se filtró entre 0.5 y 40 Hz utilizando un filtro Butterworth de orden 4 y filtrado de fase cero, y se anularon aquellas ventanas con una varianza móvil superior a cierto umbral con el fin de facilitar la detección automática de los complejos QRS. Para el cálculo de la serie temporal RR, se utilizó el detector QRS de Hamilton-Tompkins para conocer los instantes de las ondas R.

Las características extraídas se pueden dividir en 3 grupos principales [6]:

- **Dominio temporal:** Media de los intervalos RR (v_1), desviación estándar (v_2), RMSSD (v_3 , *root mean squared of the successive differences*), coeficiente de variación (v_4 , definido como v_2/v_1) [7], nNN50 (v_5 , cantidad de intervalos RR que difieren más de 50 ms), pNN50 (v_6 , nNN50 dividido entre la cantidad de intervalos RR) y rango intercuartílico de la serie RR (v_7).
- **Dominio frecuencial:** El espectro de la serie temporal RR se estima utilizando el periodograma de Lomb-Scargle y se analizan dos bandas, LF (0.04-0.15 Hz) y HF (0.15-0.4 Hz). Las características extraídas son la potencia absoluta y relativa de la banda LF (v_8 y v_9) y de la banda HF (v_{10} y v_{11}), la relación v_8/v_{10} (v_{12}), frecuencia correspondiente al pico en la banda LF (v_{13}) y la frecuencia correspondiente pico máximo en la banda HF (v_{14}).
- **Descriptores no lineales:** Se calculan los descriptores que caracterizan la forma geométrica del diagrama de Poincaré, SD1 (v_{15}) y SD2 (v_{16}) asociados al eje menor y mayor de la elipse ajustadas, respectivamente, así como la relación SD1/SD2 (v_{17}). Por último, se remuestrea la serie temporal de los intervalos RR con una frecuencia de muestreo de 10 Hz (mediante la interpolación cúbica) y se calcula la entropía muestral con $m=1$ y $r=0.2$ (v_{18}).

2.3. Clasificación

Todas las características extraídas se utilizaron como entrada a un clasificador *Random Forest* de 100 árboles [8]. Dicho clasificador tiene una funcionalidad interna para medir las importancias de los descriptores de entrada. Para este trabajo se permutaron los valores de entrada de cada descriptor y se observó cómo variaba el *out-of-bag error* para caracterizar la importancia que tiene cada descriptor.

El clasificador se evaluó utilizando el AUC, es decir, el área bajo la curva ROC (*receiver operating characteristics*). Para entrenar los modelos se utilizó un esquema de validación cruzada, *5-fold cross-validation*. Los datos se partitionaron en 5 grupos, y en cada iteración se entrenó el clasificador utilizando 4/5 de los datos y se estimaron los diagnósticos en la parte restante. Cada vez que se entrenó el clasificador también se calcularon las importancias de los descriptores de entrada y se guardaron para su posterior análisis. Además, se volvió a entrenar el clasificador utilizando los N_f descriptores más importantes según el *Random Forest*. Este proceso se realizó 5 veces, de tal manera que toda la base de datos se utilizó para entrenar y testear el modelo. El cálculo del AUC se realizó comparando las anotaciones manuales (referencia) y las estimaciones del clasificador. Este ciclo se repitió 50 veces para estimar la distribución estadística del AUC.

3. Resultados

Los valores de mediana (percentil 10-percentil 90) del AUC calculados para $t_w=60$ s y $t_w=120$ s utilizando los 18 descriptores fueron 0.69 (0.71-0.75) y 0.75 (0.72-0.77), respectivamente. En la Figura 3 se muestran las curvas ROC para ambos valores de t_w , considerando en cada caso la iteración más próxima a la mediana del AUC.

La Figura 2 muestra las distribuciones de las importancias de los descriptores para ambos valores de t_w . Se puede apreciar que para ambos valores de t_w el descriptor más

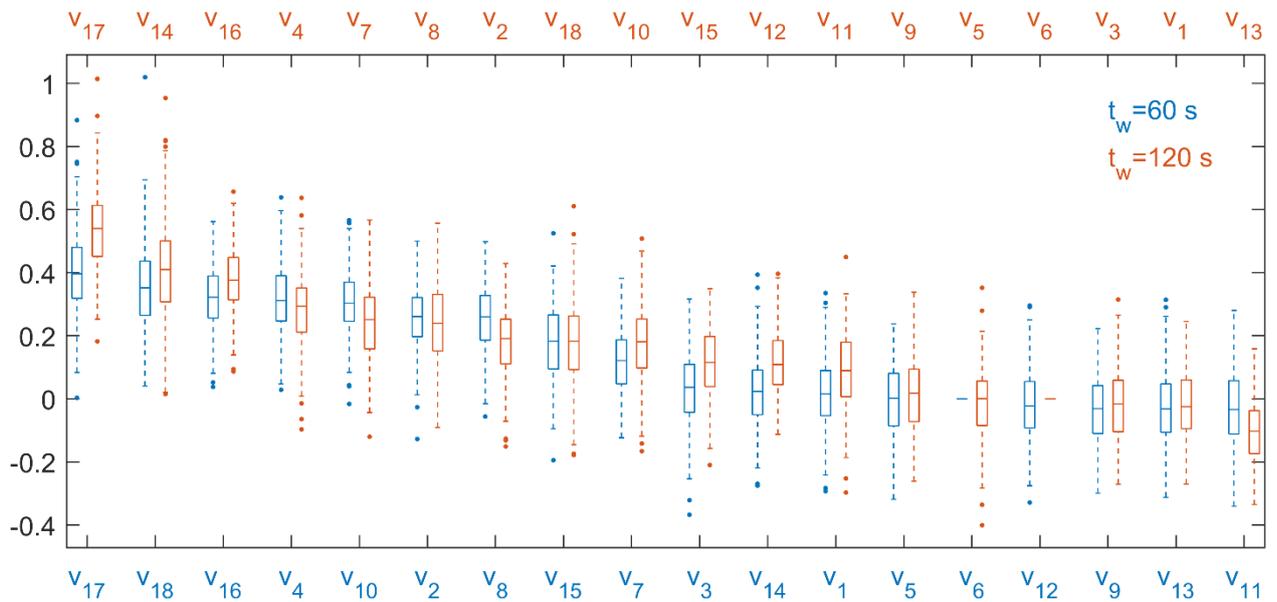


Figura 2. Distribuciones de las importancias asignadas a cada descriptor por el clasificador Random Forest utilizando ventanas de 60 s (azul) y 120 s (naranja). Los descriptors se ordenan según el valor de la mediana de su valor de importancia y se muestran para $t_w=60$ s en el eje inferior y para $t_w=120$ s en el eje superior

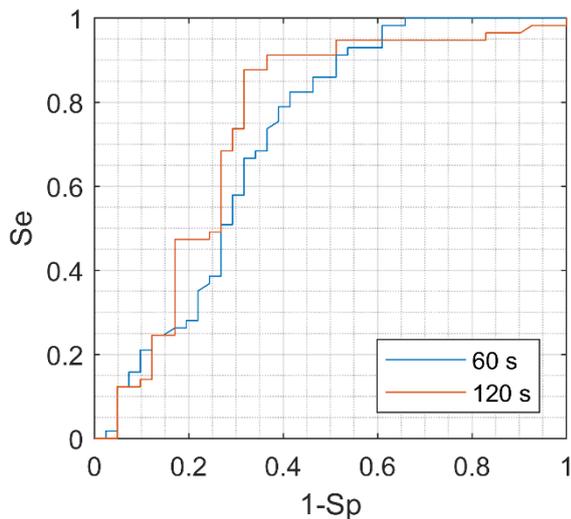


Figura 3. Curvas ROC para diferentes valores de t_w utilizando todos los descriptors. Para ambos casos se tiene en cuenta la iteración que más se acerca a la mediana del AUC

importante fue SD1/SD2. A pesar de que el segundo descriptor más importante es diferente para cada valor de t_w , el tercero y el cuarto son comunes, SD2 y el coeficiente de variación. Por otro lado, pNN50, la frecuencia pico en la banda LF y la potencia relativa de la banda LF están entre los 6 peores descriptors en ambos casos.

En la Figura 4 se muestra el comportamiento del clasificador en función del número de descriptors utilizados, N_f . La mediana del valor del AUC muestra que para $t_w=120$ no hay una ganancia significativa a partir de 10 descriptors ($AUC \approx 0.75$), pero para $t_w=60$ añadir más descriptors es perjudicial. El AUC máximo (≈ 0.73) en el segundo caso se obtiene cuando se entrena el clasificador con 5 descriptors, y desciende a 0.69 si se consideran todos los descriptors.

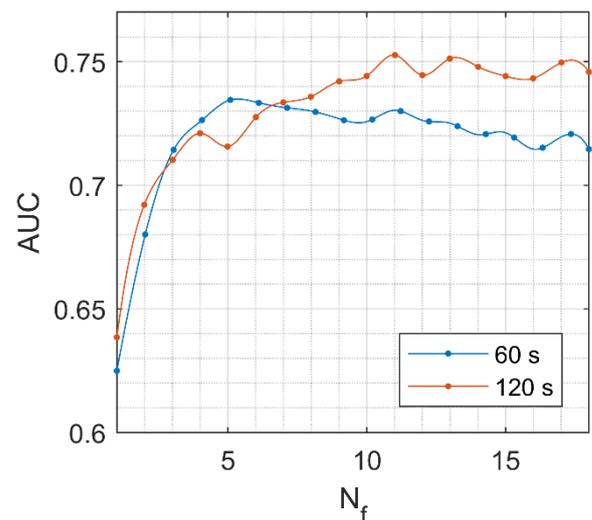


Figura 4. La mediana del valor del AUC para $t_w=60$ s y $t_w=120$ s en función del número de descriptors incluidos en el clasificador, N_f

4. Discusión y conclusiones

En este estudio se analiza la relación entre la variabilidad del ritmo cardíaco y la parada recurrente con el objetivo de predecir la PaR en un paciente que ha recuperado el pulso.

Hasta el momento, la bibliografía muestra un interés médico importante en el tema, y unas propuestas modestas de predicción basadas en relaciones numéricas calculadas manualmente sobre el ECG. Son pocos los métodos automáticos propuestos para este fin [9].

El mejor resultado, $AUC \approx 0.75$, se obtiene para un predictor basado en 11 descriptors de la variabilidad del ritmo cardíaco, calculados en una ventana de 120 s tras la

RCE, lo cual pone de manifiesto relación entre la variabilidad del ritmo cardíaco y la PaR.

La información más significativa se obtiene a partir de los descriptores de la forma geométrica del diagrama de Poincaré, y en general, la información que se obtiene a partir de la banda HF es más significativa que la información que se obtiene de la banda LF para el clasificador. Esto se puede deber al tamaño de la ventana, ya que la duración máxima es de dos minutos. Utilizar ventanas más grandes permitiría el análisis de más bandas y el uso de más descriptores que requieren series temporales RR más largas [6].

Para este estudio se consideraron sólo 98 pacientes, de los cuales 41 sufren una PaR. Tener pocos pacientes también limita a la hora de escoger el valor de t_w , ya que muchos pacientes sufren una parada recurrente entre dos y tres minutos después de la RCE. Analizar ventanas de mayor duración o la evolución temporal de ciertos descriptores en diferentes ventanas podría mejorar los resultados.

En este estudio sólo se consideran descriptores de la variabilidad del ritmo cardíaco. Añadir otros descriptores para caracterizar la forma de onda del ECG (amplitud, duración de los complejos QRS...) podría mejorar los resultados. De hecho, un estudio reciente ha demostrado que la forma de onda de un ECG con un ritmo con pulso contiene información sobre transiciones a un ritmo con pulso o sin pulso [10]. En la Figura 5 se muestra el ejemplo de un paciente que sufre una PaR (panel a), y en el panel b se muestra el ejemplo de un paciente que no sufre ninguna PaR. En el caso del panel a se puede apreciar que los complejos QRS son más inestables, lo que podría considerarse una posible característica de predicción de PaR.

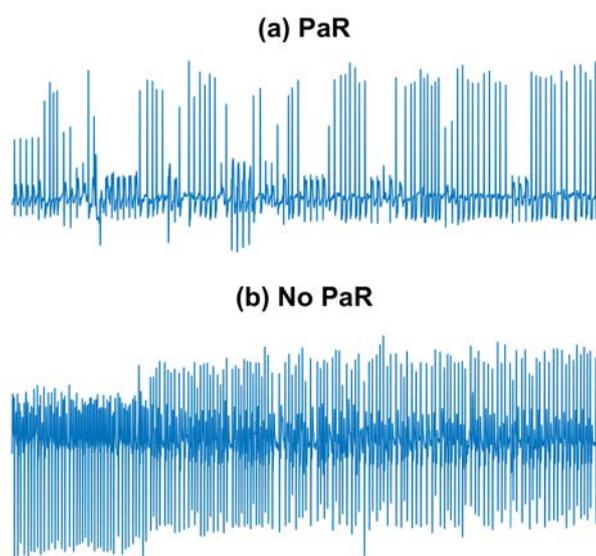


Figura 5. Ejemplos de ventanas de $t_w=60$ s asociados a dos pacientes de la base de estudio. El caso mostrado en el panel (a) corresponde a un caso con PaR, mientras que en el panel (b) se muestra un caso sin PaR

En resumen, se puede concluir que existe una relación entre la PaR y los descriptores de la variabilidad del ritmo cardíaco. Los resultados prometedores obtenidos con esta base de datos de 98 pacientes deberían ser contrastados con una base más amplia de registros.

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 y la beca PRE_2018_2_0260.

Referencias

- [1] C. Atwood, M. S. Eisenberg, J. Herlitz, y T. D. Rea, «Incidence of EMS-treated out-of-hospital cardiac arrest in Europe», *Resuscitation*, vol. 67, n.º 1, pp. 75–80, 2005.
- [2] J. J. de Vreede-Swagemakers *et al.*, «Out-of-hospital cardiac arrest in the 1990's: a population-based study in the Maastricht area on incidence, characteristics and survival», *J. Am. Coll. Cardiol.*, vol. 30, n.º 6, pp. 1500-1505, nov. 1997.
- [3] I. Lund-Kordahl, T. M. Olasveengen, T. Lorem, M. Samdal, L. Wik, y K. Sunde, «Improving outcome after out-of-hospital cardiac arrest by strengthening weak links of the local Chain of Survival; quality of advanced life support and post-resuscitation care», *Resuscitation*, vol. 81, n.º 4, pp. 422-426, abr. 2010.
- [4] A. Bhardwaj *et al.*, «Factors associated with re-arrest following initial resuscitation from cardiac arrest», *Resuscitation*, vol. 111, pp. 90-95, feb. 2017.
- [5] D. D. Salcido, M. L. Sundermann, A. C. Koller, y J. J. Menegazzi, «Incidence and outcomes of rearrest following out-of-hospital cardiac arrest», *Resuscitation*, vol. 86, pp. 19-24, ene. 2015.
- [6] F. Shaffer y J. P. Ginsberg, «An Overview of Heart Rate Variability Metrics and Norms», *Front. Public Health*, vol. 5, sep. 2017.
- [7] K. Tateno y L. Glass, «Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and deltaRR intervals», *Med. Biol. Eng. Comput.*, vol. 39, n.º 6, pp. 664-671, nov. 2001.
- [8] L. Breiman, «Random Forests», *Mach. Learn.*, vol. 45, n.º 1, pp. 5-32, oct. 2001.
- [9] D. Salcido, M. Sundermann, A. Koller, y J. Menegazzi, «Towards predicting the time and rhythm of rearrest after out-of-hospital cardiac arrest», *Resuscitation*, vol. 96, p. 10, nov. 2015.
- [10] E. Alonso, T. Eftestøl, E. Aramendi, J. Kramer-Johansen, E. Skogvoll, y T. Nordseth, «Beyond ventricular fibrillation analysis: Comprehensive waveform analysis for all cardiac rhythms occurring during resuscitation», *Resuscitation*, vol. 85, n.º 11, pp. 1541-1548, nov. 2014.