

Clasificación de los sonidos cardíacos usando ondículas y redes neuronales

Juan Peralta¹ , Liseth Carrión¹ , Juan B. Tenesaca¹ , Andrés Vázquez-Rodas² 

¹ Escuela de Electrónica y Telecomunicaciones, Universidad de Cuenca, Av. 12 de Abril & Agustín Cueva, Cuenca, Ecuador.

² Departamento de Ingeniería Eléctrica, Electrónica y Telecomunicaciones, Universidad de Cuenca, Av. 12 de Abril y Agustín Cueva, Cuenca, Ecuador.

Autor para correspondencia: andres.vazquezr@ucuenca.edu.ec

Fecha de recepción: 25 de agosto de 2017 - Fecha de aceptación: 29 de septiembre de 2017

RESUMEN

La auscultación de los sonidos cardíacos es un examen clínico que permite determinar si un paciente debe ser referido a un especialista. El fonocardiograma (PCG), por sus siglas en inglés, corresponde al registro de estos sonidos. El objetivo de este trabajo es la evaluación de un esquema fundamentado en dos algoritmos propuestos durante el desafío *PhysioNet* 2016, el primero basado en ondículas y el segundo en una red neuronal convolucional (RNC), para evaluar el desempeño en la clasificación de los sonidos cardíacos (normal/anormal). Los resultados obtenidos muestran un mejor equilibrio entre especificidad y sensibilidad con respecto al método de las ondículas, aunque su desempeño es inferior al método basado en RNC. El método propuesto tiene un menor costo computacional.

Palabras clave: Fonocardiograma, corazón, PCG, *PhysioNet*.

ABSTRACT

The auscultation of cardiac sounds is a clinical examination that allows to determine if a patient should be referred to a specialist. The phonocardiogram (PCG) corresponds to the recording of these sounds. The objective of this work is the evaluation of the combination of two of the proposed algorithms during *PhysioNet* 2016 challenge, the first is based on wavelets and the second on a neural convolutional network to evaluate the performance in the classification of cardiac sounds (normal/abnormal). The results show a better balance between specificity and sensitivity with respect to the wavelet method, although its performance is inferior to the method based on neural networks. The proposed method has a lower computational cost.

Keywords: Phonocardiogram, heart, PCG, *PhysioNet*.

1. INTRODUCCIÓN

La apertura y cierre de las válvulas del corazón se asocia con aceleraciones y desaceleraciones de la sangre, dando lugar a vibraciones de toda la estructura cardíaca. Estas vibraciones son audibles en la pared del pecho, y la auscultación de estos sonidos cardíacos específicos puede dar una indicación de la salud del corazón. La representación gráfica de estos sonidos se denomina fono cardiograma (PCG).

La segmentación de los ruidos cardíacos fundamentales (FHSS por sus siglas en inglés), se basa en la delineación de los sonidos: S1 y S2. El sonido S1 se produce al comienzo de la contracción ventricular isovolumétrica, cuando las válvulas mitral y tricúspide se cierran debido al rápido aumento de la presión dentro de los ventrículos. El sonido S2 se produce al comienzo de la diástole con el cierre de las válvulas aórtica y pulmonar. En la Figura 1, se presenta un diagrama que representa la ubicación de los sonidos cardíacos y su relación con el Electrocardiograma (ECG). La localización exacta de S1

y S2 es un requisito previo para la identificación de las regiones sistólica y/o diastólica, lo que permite la posterior clasificación de situaciones patológicas en estas regiones (Huiying, Sakari, & Iiro, 1997).

El desafío del Congreso Computing in Cardiology Challenge 2016, tuvo como objetivo fomentar el desarrollo de algoritmos para clasificar las grabaciones de sonidos del corazón recogidos de una variedad de entornos clínicos y no clínicos. El objetivo era identificar, a partir de un registro corto (10-60 s) y desde una única ubicación precordial, si el sujeto debía ser referido a un diagnóstico experto, centrándose sólo en la clasificación precisa de los sonidos cardíacos normales y anormales, especialmente cuando algunos sonidos cardíacos muestran una calidad de señal a ruido muy baja (Potes, Parvaneh, Rahman, & Conroy, 2016). Los mejores desempeños fueron obtenidos por Potes *et al.* (2016), usando técnicas de redes neuronales y una variante del clasificador AdaBoost. Otro enfoque fue el uso del clasificador basado en la transformada ondícula, a partir del cálculo de la entropía del PCG (Goda & Hajas, 2016).

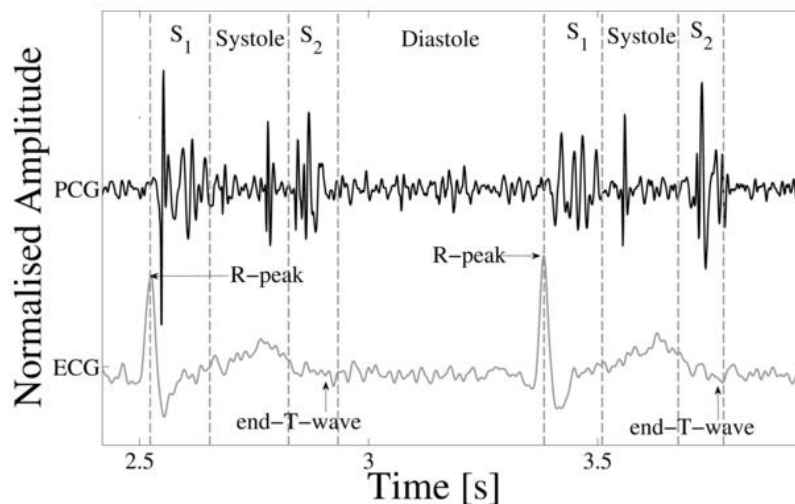


Figura 1. PCG y ECG Realizadas en simultáneo. Se muestran los cuatro estados del ciclo del corazón (sonido S1, sístole, sonido S2, y diástole), y para referencia las etiquetas del pico-R y el final de la onda-T.

El objetivo de este trabajo es la evaluación de un esquema basado en el uso de una red neuronal perceptrón multicapa en lugar de una red neural convolucional (CNN), tomando como entradas la entropía del PCG estimada mediante el método de la transformada de ondículas (Langley & Murray, 2016) en la clasificación de los sonidos cardíacos.

2. MATERIALES Y MÉTODOS

2.1. Base de datos

Los datos consistieron en ocho bases de datos de sonidos cardíacos, recopilados de forma independiente por siete equipos de investigación diferentes, procedentes de siete países, durante un período de más de una década. Esto resultó en un total de casi 30 horas de grabaciones que contienen 233,512 sonidos cardíacos de 116,865 latidos del corazón, en 4,430 grabaciones tomadas de 1,072 personas.

2.2. Algoritmos

Algoritmo basado en la transformada de ondículas

Una ondícula es una "pequeña onda" de duración limitada, es decir, su energía está concentrada en el tiempo alrededor de un punto, lo que proporciona una adecuada herramienta para el análisis de

fenómenos transitorios, no estacionarios, variables en el tiempo y aquellos que presenten discontinuidades. La implementación se basa en el uso de la transformada de ondículas como medio de separación de tipos de capturas de audio precordiales de los diferentes pacientes (Potes *et al.*, 2016). Este algoritmo calcula la entropía del audio empleado y separa el tipo de señal por medio de un umbral establecido como muestra la Figura 2. En la Figura 3 se muestra un ejemplo de procesamiento con este método, en los paneles superiores se muestra la señal original, en los paneles intermedios, la transformada ondículas y en los paneles inferiores el cálculo de la entropía (Potes *et al.*, 2016).



Figura 2. Diagrama de flujo para ilustrar los pasos de procesamiento del algoritmo (Langley & Murray, 2016).

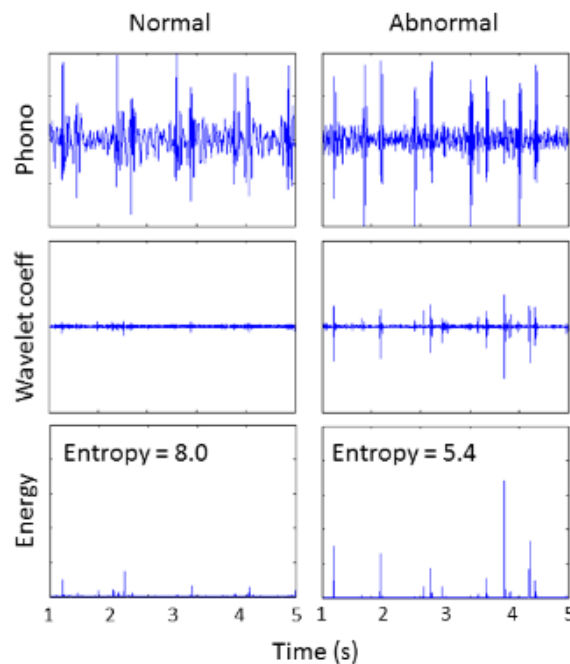


Figura 3. Ejemplo del procesamiento de la señal con la transformada de ondículas (Langley & Murray, 2016).

Algoritmo basado en redes neuronales

Cada PCG comprende más de un ciclo cardíaco (latido), y cada latido consta de cuatro estados de sonido del corazón (S1, sístole, S2 y diástole). Este algoritmo clasifica los registros de sonido en dos, usando un conjunto de clasificadores: AdaBoost y Red neuronal convolucional (Tschannen, Kramer, Marti, Heinzmann, & Wiatowski, 2016). Se emplea Ada Boost para clasificar PCGs en normales o anormales, usando características del dominio del tiempo y frecuencia. En el dominio temporal, se usan intervalos y amplitudes PCGs. En el dominio de la frecuencia se usa el espectro de potencia de cada estado del sonido del corazón, utilizando una ventana de Hamming y la transformada de Fourier de tiempo discreto.

En la Figura 4 se muestra la arquitectura de la CNN empleada. Las entradas de la red corresponden a la descomposición del ciclo cardíaco: S1, sístole, S2 y diástole. Las entradas a la red corresponden a cuatro series de tiempo, una por cada banda de frecuencia. Cada una de las CNN consta de tres capas: la capa de entrada seguida de dos capas de convolución. La salida de los 4-CNNs pasa a una red de perceptrón multicapa (MLP). La capa de salida calcula la puntuación de clase (es

decir, el valor de probabilidad, CNN_ABN) del sonido anormal del corazón. Finalmente, en la salida de los dos clasificadores, tanto de ABN y CNN se combinan usando una regla de decisión para producir el resultado de clasificación final: normal o anormal, de acuerdo a un umbral de decisión (Potes *et al.*, 2016).

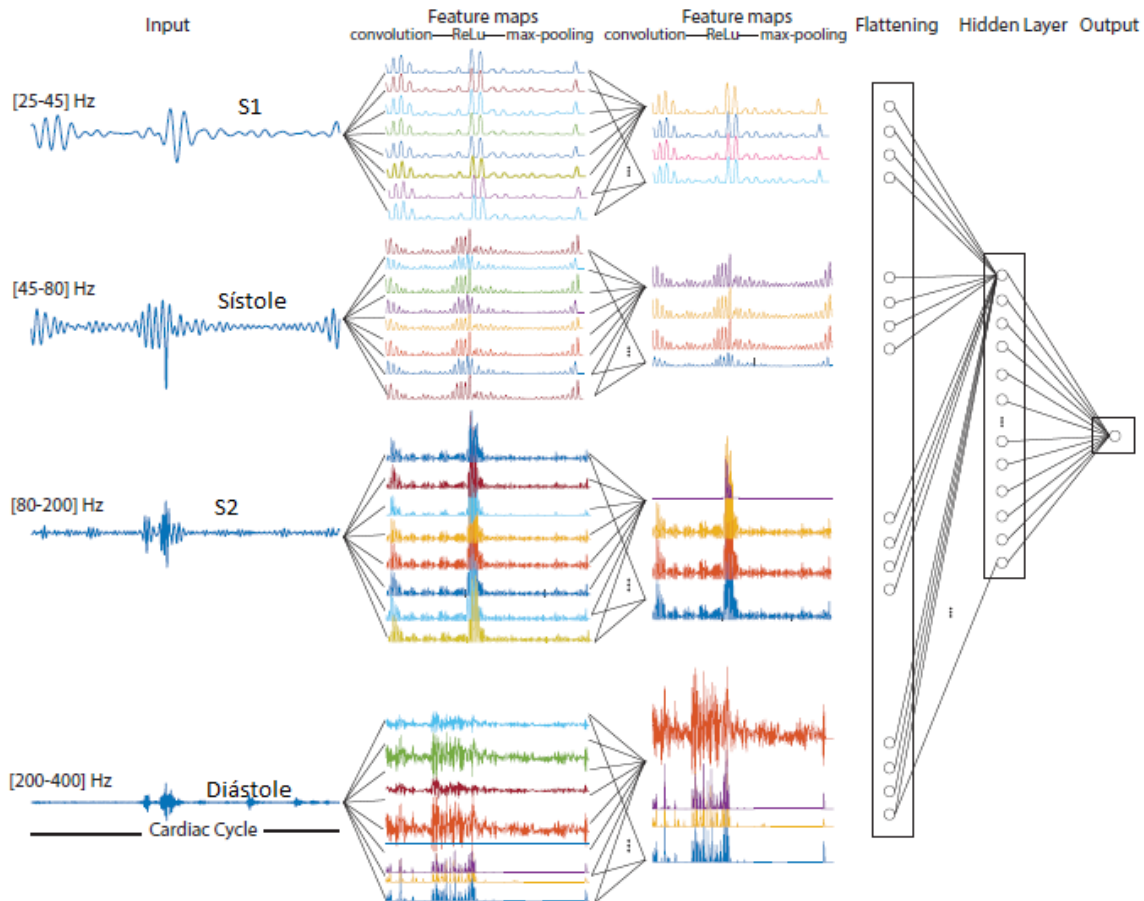


Figura 4. Arquitectura de una red neuronal convolucional (CNN) (Potes *et al.*, 2016).

Algoritmo propuesto

Se propone una metodología usando una red neuronal de perceptrón multicapa (MLP), las entradas de la red corresponden a la entropía de la transformada de ondículas del clip de sonido. Este esquema reduce los costos computacionales de la red neuronal convolucional (CNN). Tomando como punto de partida el algoritmo planteado por Langley & Murray (2016), que consta de tres etapas: la señal, los coeficientes de la transformada de ondículas y la energía, a partir de la cual se calcula la entropía.

Para ciertos casos, el cálculo de la entropía total de la energía de la transformación de ondículas es igual a un valor no existente debido a la división por cero. Por esto, se optó por dividir la energía total de la señal en 5 partes y realizar el cálculo de la entropía en cada división, de esta manera la pérdida de información fue menor.

Se usó el cálculo de la entropía para la validación y entrenamiento de la red, el esquema de perceptrón multicapa usado fue de dos capas, con tres neuronas en la primera y dos en la segunda. Finalmente se usó validación cruzada para entrenar la red con la base de datos. El diagrama de flujo del algoritmo propuesto se expone en la Figura 5.



Figura 5. Diagrama de bloques del algoritmo propuesto para el análisis del sonido cardiaco.

3. RESULTADOS Y DISCUSIÓN

Las primeras cuatro etapas del algoritmo, correspondientes a la transformada de ondículas y el cálculo de energía y entropía, se realizó usando Matlab versión R2013B. La matriz de entropía de cada sección de energía del PCG, junto con la base de datos que clasifica el sonido en normal o anormal, proporcionada por PhysioNet (Langley & Murray, 2016), fueron los datos de entrada a la red neuronal implementada, usando Weka versión 3.8.

En la siguiente etapa se buscó la configuración para la red neuronal que logró maximizar los valores de sensibilidad y especificidad, para evitar el sobre entrenamiento de la red se utilizó el método de validación cruzada. Se realizaron pruebas con configuraciones de redes neuronales con un número de capas intermedias de 2 a 5 y con un número de neuronas de 3 a 50, obteniendo un porcentaje de clasificación correcta entre un mínimo de 38.00% a un máximo de 61.78%. Adicionalmente se realizó una prueba de clasificación mediante el uso de máquinas de vectores de soporte, con el mismo se obtuvieron valores de probabilidad de clasificación correcta, que variaron entre un mínimo de 40.00% y un máximo de 58.00%. Los mejores resultados se obtuvieron con un modelo de entrenamiento de validación cruzada, los resultados se pueden observar en la Tabla 1.

Tabla 1. Resultados de los tres tipos de algoritmos.

Algoritmo	Especificidad	Sensibilidad	Distancia al punto óptimo
Entropía Wavelet	0.2800	0.9338	0.723
Ada Boost-Red Neuronal	0.7781	0.9424	0.229
Propuesto	0.3466	0.8874	0.663

Se observó que, si bien el algoritmo basado en redes neuronales tiene mejores resultados de clasificación que el algoritmo propuesto y que el basado en la transformada de ondículas, resulta mucho más complicado en su ejecución computacional, esto se demuestra al constatar el tiempo de ejecución y la memoria RAM necesaria para la clasificación de 301 sonidos cardiacos, tal como se expone en la Tabla 2.

Tabla 2. Comparación de eficiencia entre algoritmos.

Algoritmo	Tiempo de ejecución (s)	Cantidad de memoria RAM (MB)	Clasificaciones por segundo
Entropía de Ondículas	4.202	9.5	71.942
Ada Boost-Red Neuronal	1334.23	62.1	0.2256
Propuesto	7.429	17.3	40.51

4. CONCLUSIONES

El método de la transformada de ondículas refleja diferencias importantes en las características de las grabaciones en tiempo y frecuencia, con un rendimiento eficaz únicamente con alta sensibilidad y baja especificidad. El algoritmo propuesto presenta un equilibrio entre sensibilidad y especificidad, mejora el desempeño de los resultados obtenidos mediante el método de ondículas propuesto por Philip

Langley. Sin embargo, el desempeño del algoritmo propuesto es todavía inferior al algoritmo original de Potes, pero con menor complejidad computacional, ya que utiliza un número menor de datos de entrada, y con menor complejidad en el preprocesamiento de estos. Esto es importante si se desea implementar este algoritmo en dispositivos de bajo procesamiento.

Los trabajos futuros deberían enfocarse en pre-procesar el PCG para mejorar la relación señal a ruido y en adicionar otras características en la entrada del perceptrón multicapa, manteniendo al mínimo la complejidad computacional para una eventual implementación del algoritmo en dispositivos móviles.

AGRADECIMIENTOS

A Sara Wong, docente del Departamento de Eléctrica, Electrónica y Telecomunicaciones de la Universidad de Cuenca, quien asesoró con la evaluación de este proyecto.

REFERENCIAS

- Goda, M. Á., Hajas, P. (2016). Morphological determination of pathological PCG signals by time and frequency domain analysis. *Computing in Cardiology Conference (CinC)*, 1133-1136. <https://doi.org/10.23919/CIC.2016.7868947>
- Huiying, L., Sakari, L., Iiro, H. (1997). A heart sound segmentation algorithm using wavelet decomposition and reconstruction. *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4, 1630-1633. <https://doi.org/10.1109/IEMBS.1997.757028>
- Langley, P., Murray, A. (2016). Abnormal heart sounds detected from short duration unsegmented phonocardiograms by wavelet entropy. *Computing in Cardiology Conference (CinC)*, 545-548. <https://doi.org/10.23919/CIC.2016.7868800>
- Potes, C., Parvaneh, S., Rahman, A., Conroy, B. (2016). Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. *Computing in Cardiology Conference (CinC)*, 621-624. <https://doi.org/10.23919/CIC.2016.7868819>
- Tschannen, M., Kramer, T., Marti, G., Heinzmann, M., Wiatowski, T. (2016). Heart sound classification using deep structured features. *Computing in Cardiology Conference (CinC)*, 565-568. <https://doi.org/10.23919/CIC.2016.7868805>