

Análisis de la Morfología de la Presión Arterial como Huella de la Salud Cardiovascular utilizando un enfoque de Aprendizaje Automático

Analysis of Blood Pressure Morphology as a Fingerprint of Cardiovascular Health using A Machine Learning Based Approach

Presentación: 04/10/2022 y 05/10/2022

Doctoranda:

Eugenia Ipar

Grupo de Investigación y Desarrollo en Bioingeniería, Facultad Regional Buenos Aires, Universidad Tecnológica Nacional
eipar@frba.utn.edu.ar

Director:

Leandro J. Cymberknop

Codirector:

Ricardo L. Armentano

Resumen

Se presentan resultados obtenidos de la predicción de varios parámetros claves, como la velocidad de onda de pulso, relacionados a la salud cardiovascular utilizando un modelo unidimensional (1-D) como parte del entrenamiento de los modelos de aprendizaje automático, así como también la verificación de dichas predicciones utilizando señales reales obtenidas de un protocolo de medición.

Palabras clave: Salud cardiovascular, aprendizaje automático, velocidad de onda de pulso.

Abstract

Results obtained from the estimation of several key parameters, such as pulse wave velocity, related to cardiovascular health using a one-dimensional (1-D) model as part of the machine learning models training, are presented, as well as the verification of these predictions using real signals obtained from a measurement protocol.

Keywords: Cardiovascular health, automatic learning, pulse wave velocity

Introducción

Realizar ejercicio diariamente y tener una dieta sana ayuda al sistema cardiovascular a envejecer lentamente. Sin embargo, hábitos de vida poco saludables y enfermedades impactan negativamente a esta condición, la cual se ve reflejada como un “endurecimiento” acelerado de las principales arterias, actuando independientemente del envejecimiento normal (Bikia, Papaiaoannou, Pagoulatou, & al., 2020). Por esa razón, la “edad cronológica”

(EC) de un individuo puede diferenciarse de la “edad arterial” (EA). En este sentido, la Asociación Americana del Corazón ha definido a la Salud Cardiovascular (SCV) como la presencia de cuatro comportamientos saludables (no fumador, actividad física regular, índice de masa corporal (IMC) menor a 25 kg/m² y una dieta consistente con las recomendaciones nacionales) y otros tres factores de salud (colesterol total, presión arterial y niveles de glucosa en ayuno) en valores ideales (Crichton, Elias, & Robbins, 2014).

Existe evidencia de correlación entre los parámetros de forma de onda de la presión arterial central (PAC) y la rigidez arterial (RA), ya que los mismos proveen información importante acerca de la estructura arterial sistémica y su funcionamiento permitiendo la evaluación de la reflexión y amplificación de las ondas (Kucharska-Newton, Stoner, & Meyer, 2019). De esta forma, ha sido demostrado que una mejor SCV, que comprende tanto los factores de salud como los comportamientos, se encuentra asociada a una menor RA (Crichton, Elias, & Robbins, 2014). Como resultado, la evaluación de los rangos normales de tanto propiedades hemodinámicas centrales como bio-mecánicas relacionadas con la SCV, como la velocidad de onda de pulso (VOP, estándar de oro para la estimación de la RA), el gasto cardíaco (GC), la PAC sistólica y diastólica (PAS y PAD, respectivamente) y al resistencia vascular sistémica (RVS), se vuelve extremadamente útil para la prevención de enfermedades cardiovasculares (ECV), especialmente para la población joven y de mediana edad. Se debe tener en cuenta que las enfermedades crónicas, las limitaciones funcionales y la mortalidad aumentan exponencialmente con la EC. En comparación con la EC, el concepto de EA es una medida propuesta que reflejaría con mayor precisión los cambios estructurales y funcionales (es decir, la rigidez vascular) que tienen lugar en el cuerpo a medida que envejece (Wolsk, y otros, 2017). Se puede demostrar la EA de un individuo en relación con la EC: Si la EA es mayor a la EC, se puede apreciar una discrepancia y se detectan signos de envejecimiento “prematureo” (Nilsson, 2016). Además, el uso de mediciones no invasivas y simples de realizar, particularmente aquellas obtenidas del análisis morfológico de las ondas de Presión Arterial (PA) pertenecientes a diferentes arterial superficiales, constituyen una herramienta valiosa a la hora de realizar protocolos de detección (Kullo & Malik, 2017).

Por otro lado, las técnicas de aprendizaje automático (AA, Machine Learning en inglés) involucran el estudio de modelos y algoritmos estadísticos que pueden aprender progresivamente de información (datos) y lograr un desempeño buscado para una tarea específica. Los modelos basados en inteligencia artificial (IA) posibilitan la utilización de una multitud de variables en un enfoque libre de hipótesis, así identificar similitudes y diferencias en series complejas. Estos modelos pueden adaptarse tanto a casos de clasificación como de regresión. La desventaja de estos modelos es la cantidad de información necesaria para obtener un buen rendimiento. Sin embargo, puede ser desafiante y costoso obtener dicha cantidad significativa de datos etiquetados de protocolos clínicos específicos, realizados por médicos capacitados (Kumar, Sindhu, Prashanthi, & Sulthana, 2020). En consecuencia, modelos unidimensionales (1-D) del SCV validados se convierten en una alternativa muy interesante para sobrepasar esta cuestión, dado que se pueden realizar simulación del SCV tanto en salud como en enfermedad. Los modelos computacionales 1-D proporciona un enfoque complementario para la investigación, ya que permite simular ondas de PA en diferentes condiciones de SCV. Estas ondas se encuentran influenciadas tanto por el corazón, como por la vasculatura, actuando sobre su morfología la RA y los reflejos de las propias ondas (Charlton, y otros, 2019). Es así, que se pueden realizar técnicas de AA para detectar signos de enfermedades cardiovasculares prematuras para facilitar a la prometedora medicina predictiva, preventiva y personalizada.

En virtud de lo expuesto, el objetivo principal de la tesis en desarrollo consiste en la utilización de técnicas de aprendizaje automático y profundo para la detección subclínica de enfermedades cardiovasculares, en términos de la morfología de onda de la PA en conjunto con datos propios del individuo (como la EC). Se prestará especial atención a aquellos indicadores dentro de un marco de referencia como el Riesgo Cardiovascular (RCV), la EA y la presencia de Calcio Coronario Arterial (CAC). A su vez, se considerará la asistencia que puedan brindar estos modelos en las etapas de procesamiento de señales cardiovasculares en virtud de alcanzar una mejor precisión de los parámetros característicos.

Desarrollo

Esta sección se dividirá en tres subsecciones que sintetizan el trabajo hasta la fecha desarrollado y que comprende los trabajos: **“Cardiovascular Age and Central Blood Pressure assessment based on Peripheral Blood Pressure Morphology and Artificial Neural Networks”**, presentado en el Congreso IEEE ArgenCon 2022; **“Blood Pressure Morphology as a Fingerprint of Cardiovascular Health: A Machine Learning Based Approach”**, presentado en el 4to Congreso Internacional de Informática Aplicada (ICAI) 2021.

Base de Datos

Se accedió a una base de datos con información de individuos sanos simulados por un modelo 1-D (Charlton, y otros, 2019). La misma consta con las señales de presión, flujo y área de distintos segmentos arteriales. La misma cuenta con 4374 sujetos virtuales cuyas edades se distancian cada 10 años, comenzando en los 25 hasta los 75 años. El objetivo de utilizar esta base de datos es el de poder evaluar y entrenar los modelos de AA con señales limpias (sin ruido) correspondientes a individuos catalogados como sanos, dado que sus valores hemodinámicos se encuentran dentro de los valores estándar óptimos, es decir, bajo un marco controlado. Con dicha información, se entrenó varios modelos de AA para estimar distintos parámetros hemodinámicos obtenidos a partir de la morfología de la PA tanto carótida como radial.

Además, se cuenta con una pequeña base de datos de señales de PA de la arteria radial obtenidos de un protocolo de medición realizado por parte del propio laboratorio GIBIO. La misma cuenta con los datos de 32 voluntarios, que cumplen con los comportamientos y factores de una buena SCV. Dado que la cantidad de datos de señales de sujetos reales no es la suficiente para formar parte del entrenamiento, se utilizaron como verificación de los modelos previamente entrenados. Es así como se puede ver la gran ventaja de los modelos 1D.

Procesamiento de Señales de Presión Arterial

Debido a que el propósito del presente trabajo es estimar parámetros en base a la morfología de la señal de PA, se procedió a extraer *features* características de la propia señal, tanto en amplitud como temporales, así como varios índices obtenidos a partir de dichas características. Parte de esta información se encuentra ya disponible en la base de datos virtual, además de que las señales ya se encuentran acondicionadas. Los *features* obtenidos se encuentran detallados en la Tabla 1, en conjunto con su significado. Además de su representación gráfica en la Figura 1.

Para la base de datos propia, se utilizó un filtro IIR pasa bajos de fase cero para minimizar la distorsión y eliminar el ruido de base. Luego, para la extracción de los *features* se analizaron las señales pulso a pulso utilizando un delineador desarrollado por (Aguirre, Grall-Maës, Cymberknop, & Armentano, 2021). En este proceso, se analizó el parámetro denominado como “*skewness*” (podría traducirse como asimetría estadística) para evaluar la calidad de la señal. Dicho parámetro es un indicador del grado de simetría que presenta la distribución de probabilidad de una señal. Cuanto mayor dicho parámetro, peor es la calidad de la señal, por ende, los datos extraídos de la misma no son fiables. Con este criterio, se descartaron 3 señales de la base de datos.

El resultado del procesamiento es un conjunto de datos utilizado como entrada y salida de varios modelos de AA, al que se le adjunta la EC de cada sujeto. Las salidas de los modelos se encuentran detalladas en la Tabla 2.

Features	Descripción
$P1pk, P2pk$	Puntos pico máximos luego y justo antes del pico de sistóle, antes del punto dicrótico de la señal
$P1in, P2in$	Puntos de inflexión de la onda reflejada y emitida, uno se produce más temprano que el otro
$PASr, PASC$	Pico de sistole en señal radial y carótida
$PADr, PASC$	Pico de diástole en señal radial y carótida
$A1c, A1r$	Índice de aumento radial y carótido. Es una medida indirecta de la RA y crece con la EC
AP	Índice de presión de aumento. Es la medida de la contribución que hace la reflexión de la onda a la PAS.
dP/dt	Punto pico de la primera derivada de la señal.
$SEVR$	Relación de viabilidad subendocárdica. Es la ratio de la integral presión-tiempo diástolica a sistólica. Refleja el equilibrio entre la perfusión coronaria y la carga arterial
AUC	Área bajo la curva de la señal
$LVET$	Tiempo de yección del ventrículo izquierdo. Intervalo de tiempo desde la apertura de la válvula aórtica hasta el cierre de la misma.
DTF	Fracción de tiempo diástolica. Es un ratio entre el tiempo de duración diástole y el tiempo de duración del pulso.

Tabla 1: Descripción de los predictores (o *features*) extraídos de la señal de PA

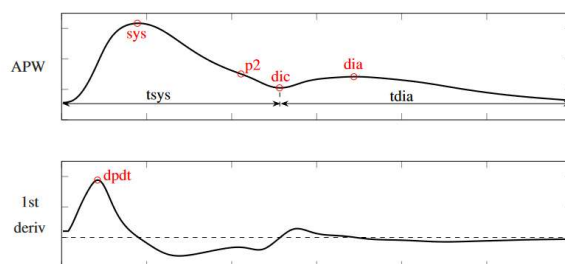


Figura 1: Representación de los predictores extraídos de la señal de PA

Features	Descripción
<i>GC</i>	Gasto Cardíaco.
<i>RVS</i>	Resistencia Vascular Sistémica
<i>PAS</i>	Pico de sístole aórtica (central)
<i>PAD</i>	Pico de diástole aórtica (central)
<i>VOPcf</i>	Velocidad de Onda de Pulso carótido-femoral.
<i>Ala</i>	Índice de presión de aumento aórtico.
<i>EC</i>	Edad Cronológica, como sustituto de la EV.

Tabla 2: Descripción de las salidas de los modelos

Estimación de Parámetros Hemodinámicos

Se realizaron dos etapas de estudio. La primera etapa consistió en la predicción del GC, RVS, PAS y PAD aórticas, VOPcf, Ala y EC, comparando los resultados de distintos modelos, aquellos son: Fully-Connected Neural Networks (NN), Gaussian Process (GP), Ensemble Methods (EM), Support Vector Machine (SVM) y K-Nearest Neighbors (KNN). Además, se comparó que señal proveniente de dos arterias (carótida y radial) otorgaba el mejor resultado con cada modelo. En esta etapa, se utilizaron exclusivamente los datos de la base de datos virtual tanto para entrenamiento como para verificación de los modelos, se realizó una separación aleatoria de los datos en un 80% del total para entrenamiento, y el 20% restante para verificación. Es de resaltar que, para la predicción de la EC, los modelos se utilizaron para clasificación de la misma. Para el restante de las salidas, se utilizó regresión. Los modelos de regresión se evaluaron utilizando el error cuadrático medio normalizado como no (RMSe, nRMSE) y el error porcentual absoluto medio (MAPE). Para clasificación, se comparó la precisión (accuracy, en inglés) de cada modelo. Además, se realizó el correspondiente análisis de Bland-Altman.

El propósito de la segunda etapa fue verificar los resultados de la primera, pero con señales de individuos reales. Se realizó sólo con Redes Neuronales, debido a que había dado resultados satisfactorios tanto para clasificación como regresión con señales provenientes de la arteria radial. Al comenzar, nos encontramos con una dificultad, que fue obtener los puntos de inflexión de las señales reales. Los mismo se obtienen de la segunda derivada de la señal. Si la señal es ruidosa, por más que se filtre, es difícil encontrarlos. De esta forma, se debió eliminar los puntos de inflexión (P1in, P2in, P1pk, P2pk) como entradas, y los índices obtenidos a raíz de ellos como AI, AP. Además, la base de datos real carecía de información para verificación de algunas salidas, como el GC, RVS y VOPcf, por lo cual las mismas se verificaron con nomogramas que estuviesen dentro de los rangos correctores, pero los valores de RMSE o MAPE no pudieron obtenerse. Dicha base de datos contaba sólo con las señales de la arteria radial, por lo tanto, se pudo verificar el modelo con dicha entrada, no con la señal carótida.

Resultados

Los resultados obtenidos en la primera etapa son muy satisfactorios, y sugieren que las señales de ambas arterias proveen información valiosa para la predicción de los parámetros de salida. Los mismos también sugieren que Gaussian Process y Neural Networks son los mejores modelos para dichas predicciones. La comparación de los valores de evaluación para cada arteria y modelo pueden observarse en la Tabla 3. La Figura 2 muestra los gráficos de dispersión y Bland-Altman para la salida de PAS. No se adjuntan los gráficos de todas las salidas por cuestión de espacio.

Modelo	Carótida						Radial					
	GC	RVS	PAS	PAD	VOPcf	Ala	GC	RVS	PAS	PAD	VOPcf	Ala
GP	0.118	4.405	0.087	0.067	0.127	1.815	0.093	6.675	0.459	0.541	0.872	1.764
NN	0.144	6.781	0.122	0.116	0.157	1.883	0.096	6.062	0.588	0.455	0.190	2.955
SVM	0.161	7.222	0.632	0.383	0.177	2.120	0.139	5.929	0.674	0.499	0.204	3.585
EM	0.326	14.507	0.518	0.479	0.379	2.681	0.326	14.347	1.323	0.815	0.410	5.016

Modelo	Precisión (%)	
	Carótida	Radial
NN	93.7	94.7
SVM	83.8	78.7
KNN	68.3	62.8
EM	81.4	77.6

Tabla 3: Valores de evaluación de los modelos. Tabla superior muestra los valores de RMSE para regresión. Tabla inferior los valores de precisión para clasificación.

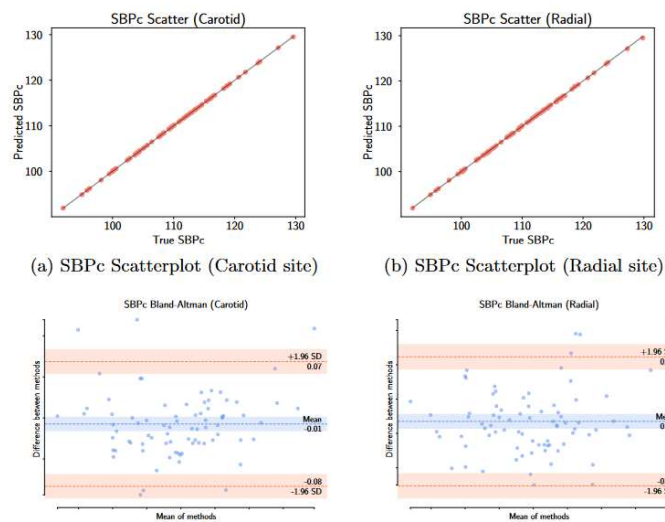


Figura 2: Gráfico de Dispersión y Bland-Altman para la PAS central en cada arteria en la etapa 1

Los resultados de la segunda etapa son satisfactorios, pero no concluyentes, debido a la poca cantidad de individuos de la base de datos propia. Además, no se pueden calcular los valores de evaluación de la mayoría de las salidas, debido a que no se cuenta con ese dato en la base. La Tabla 4 muestra los valores de evaluación para la PAS central, mientras que la Figura 3 muestra el gráfico de dispersión y Bland-Altman para la misma.

Métricas	PAS
RMSE	5.85
MAPE (%)	4.3

Modelo	Precisión (%)
NN	68.9

Tabla 4: Valores de evaluación de los modelos. Tabla izquierda muestra los valores de RMSE para PAS. Tabla derecha los valores de precisión para clasificación.

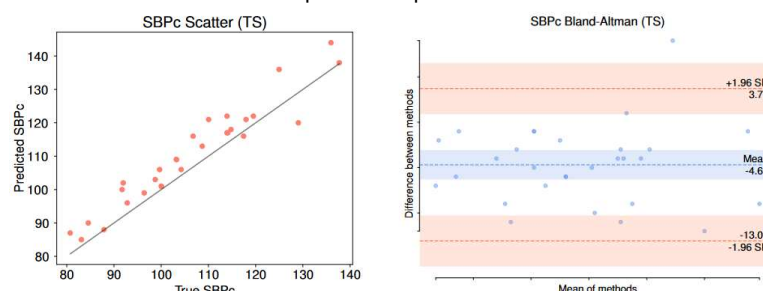


Figura 3: Gráfico de Dispersión y Bland-Altman para la PAS central en la etapa 2

Para las demás salidas, se realizaron nomogramas utilizando los datos de la base de sujetos virtuales con el fin de construir las líneas de los percentiles (utilizamos 97.5, 50 y 2.5). Luego, se graficaron sobre los mismos los valores de las predicciones realizadas, con el fin de verificar que estén dentro de los rangos correctos, dado que los voluntarios del protocolo se los considera saludables. En la Figura 4, se pueden visualizar los mismos para el GC y la VOPcf. Se pueden observar algunos “outliers”, los cuales corresponden a los mismos sujetos en ambos gráficos y se debe a que la señal obtenida es muy ruidosa, con lo cual los datos ingresados no son altamente confiables.

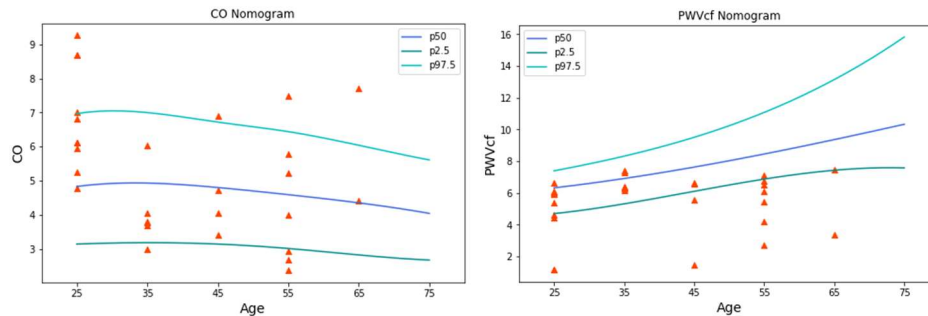


Figura 4: Nomogramas de GC (izquierda) y VOPcf (derecha), en conjunto con las predicciones realizadas como gráfico de dispersión.

Conclusiones

Los desarrollos hasta el día de la fecha responden parcialmente a los objetivos planteados por la tesis. Se implementaron modelos de aprendizaje automático que abordaron la estimación de parámetros cardiovasculares de carácter no invasivo a partir de la señal de presión arterial de una arteria, tanto en sujetos simulados como reales. Se profundizó en el procesamiento de señales de presión, y en la cuantificación de la calidad de las mismas. Además, se promovió la creación de bases de datos propias del laboratorio, incentivando a futuros protocolos clínicos de medición.

Referencias

- Aguirre, N. A., Grall-Maës, E., Cymberknop, L. J., & Armentano, R. L. (2021). A Delineator for Arterial Blood Pressure Waveform Analysis Based on a Deep Learning Technique. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, (pp. 56-59).
- Bikia, V., Papaioannou, T., Pagoulatou, S., & al., e. (2020). Noninvasive estimation of aortic hemodynamics and cardiac contractility using machine learning. *Scientific Report (10)* 15015.
- Charlton, P., Mariscal Harana, J., Vennin, S., Li, Y., Chowienczyk, P., & Alastruey, J. (2019). Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes. *American Journal of Physiology-Heart and Circulatory Physiology* 317(5).
- Crichton, G., Elias, M., & Robbins, M. (2014). Cardiovascular health and arterial stiffness: the Maiane-Syracuse Longitudinal Study. *Journal of Human Hypertension (28)*, 444-449.
- Kucharska-Newton, A., Stoner, L., & Meyer, M. (2019). Determinants of Vascular Age: An Epidemiological Perspective. *Clinical Chemistry* 1(65).
- Kullo, I., & Malik, A. (2017). Arterial ultrasonography and tonometry as adjuncts to cardiovascular risk stratification. *Journal of American College of Cardiology* 49(13), 1413-26.
- Kumar, N., Sindhu, G., Prashanthi, D., & Sulthana, A. (2020). Analysis and Prediction of Cardio Vascular Disease using Machine Learning Classifiers. *6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, (pp. 15-21).
- Nilsson, P. (2016). Vascular age: how can it be determined? What are its clinical applications. *Medicographia*, 37(4), 454-60.
- Wolsk, E., Bakkestrøm, R., Thomsen, J., Balling, L., Andersen, M., Dahl, J., . . . Gustafsson, F. (2017). The Influence of Age on Hemodynamic Parameters During Rest and Exercise in Healthy Individuals. *JACC Heart Fial (5)*, 337-346.