



Ordenamiento y selección de parámetros para la clasificación de lesiones de mama en ultrasonografía

C. Urunaga ^{1*}, W.C.A. Pereira ², A. Kauati ³

¹ Centro Latino-americano de Tecnologías Abiertas/FPTI, Foz do Iguaçu, Brasil

² Programa de Engenharia Biomédica/COPPE/UFRJ, Rio de Janeiro, Brasil

³ Programa de Pós-Graduação em Eng. Elétrica e Computação/UNIOESTE, Foz do Iguaçu, Brasil

* urunagacristhian@gmail.com

Background, Motivación y Objetivo. La ultrasonografía de una lesión de mama (BUS) se considera el complemento más importante a la mamografía, para pacientes con masas palpables, mamas densas o mamografías no concluyentes (Zonderland et al, 1999, DOI:10.1148/radiology.213.2.r99nv05413). Por consiguiente, varios métodos computacionales han sido propuestos como segunda opinión para determinar la necesidad de la realización de una biopsia de la lesión, en este contexto surgen los Sistemas Auxiliares al Diagnóstico, del inglés: Computer Aided Diagnosis systems (CAD). Estos sistemas son utilizados cada vez más en el análisis de imágenes médicas, siendo posible encontrar aplicaciones en la detección de cáncer de mama (Flores, Pereira e Infantosi, 2015, DOI:10.1016/j.patcog.2014.06.006; Cheng et al, 2010, DOI:10.1016/j.patcog.2009.05.012; Drukker, Sennett e Giger, 2009, DOI:10.1109/TMI.2008.928178).

Este trabajo tiene como principal objetivo el estudio de la técnica de Información Mutua (MI) para seleccionar del espacio n-dimensional de atributos, aquellos de mayor relevancia y mínima redundancia. Además de implementar alguna técnica de clasificación para distinguir la lesión en dos clases: benigna y maligna.

Methods. El banco de datos utilizado fue adquirido del Programa de Ingeniería Biomédica COPPE/UFRJ. El cual consiste en 641 imágenes de lesiones de mama obtenidas con un ultrasonógrafo Sonoline Sienna® (Siemens, Germany) utilizando un transductor lineal de 7,5 MHz, que captura las imágenes directamente de la señal de video de 8 bits. Existen 413 lesiones benignas y 228 malignas. Cada imagen fue segmentada manualmente por un médico especialista y también fueron extraídos 22 atributos de cada una de las imágenes (Flores, 2013, DOI:10.1109/CEWIT.2013.6713755).

Para disminuir la redundancia entre los atributos son utilizados dos esquemas: El primero utiliza un criterio de diferencia de información mutua (MID) mientras que el segundo aplica el criterio de cociente de información mutua (MIQ). Usando ambos esquemas se obtiene un conjunto de datos ordenados de acuerdo al criterio de mínima redundancia-máxima relevancia (MRMR). A partir del cual fue utilizado un subconjunto de 7 atributos obteniendo para MID: [número substancial de protuberancias y depresiones (NSPD), esqueleto



XXVI Congresso Brasileiro de Engenharia Biomédica

Armação de Búzios – RJ – Brasil

October 21st to 25th, 2018

normalizado por la elipse equivalente (ENS), orientación, índice de lobulación (LI), distancia proporcional (PD), cruzamiento de la longitud radial normalizada(NRL cruzamiento), factor de forma] y para MIQ: [número substancial de protuberancias y depresiones (NSPD), esqueleto normalizado por la elipse equivalente (ENS), cruzamiento de la longitud radial normalizada(NRL cruzamiento), factor de forma, orientación, índice de lobulación (LI), distancia proporcional (PD)], luego fue implementada la técnica de clasificación LDA, para la cual la base de datos fue aleatoriamente dividida en 80% para entrenamiento y 20% para pruebas.

Results. Los resultados de de sensibilidad, especificidad y exactitud se encuentran el la **Tabla 1.**

Discusión y Conclusión. De acuerdo a los resultados obtenidos por (Ding y Peng, 2005, DOI: 10.1142/S0219720005001004), la diferencia principal entre ambas condiciones es que MIQ aplica una mayor penalidad en los atributos redundantes que MID. Sin embargo, en el trabajo realizado no se ha utilizado alguna técnica o metodología para determinar el número de atributos necesarios para representar las propiedades observadas en el conjunto de datos, por lo que es necesario como siguiente paso explorar la dimensionalidad intrínseca del mismo, para determinar el número de atributos que se seleccionarán.

Tabla 1. Clasificación basado en el análisis ROC.

	MIQ	MID
Área bajo la curva ROC	0.90	0.92
Sensibilidad (%)	83.99	82.67
Especificidad (%)	91.83	92.32
Exactitud (%)	89.06	88.90

Acknowledgment.

CNPq 311.650/2017-1

FAPERJ E-26/203.041/2015

CAPES/PROEX

Keywords. Mama, ultrasonografía, información mutua.