

Ensayos

Redes Bayesianas para la clasificación de masas en mamografías

Resumen

Actualmente en México, el cáncer de mama ocupa el primer lugar como causa de muerte, por cáncer, entre las mujeres. En este trabajo se explora el uso de las redes Bayesianas en la clasificación de masas en imágenes de mamografías. Dichas redes son modelos gráficos probabilísticos que han sido utilizadas en la Medicina, debido a su capacidad para capturar el conocimiento de expertos y de trabajar en dominios inciertos. Aquí, se presenta una comparación del desempeño de modelos del tipo Naive Bayes, Naive Bayes aumentado a árbol (TAN), Clasificador Bayesiano K-dependiente (KDB) y Naive Bayes aumentado a bosque (FAN). Estos modelos fueron construidos con un conjunto de 67 características de ubicación, intensidad, forma y textura; y evaluados con el método Leave-one-out cross validation. Las características mencionadas fueron extraídas de un conjunto de datos público denominado mini-MIAS. Los resultados experimentales muestran que modelos del tipo TAN y FAN, construidos con un subconjunto de once características, logran un desempeño de 0.81 en exactitud, 0.65 en sensibilidad y 0.89 en especificidad. Por lo tanto, se considera que estos modelos, los cuales permiten dependencias entre variables (características), son adecuados y prometedores para la clasificación automatizada de masas.

Abstract

Nowadays, breast cancer is the deadliest form of cancer among women in Mexico. In this study the performance of Bayesian networks for the classification of masses in mammographic images is analyzed. These networks are probabilistic graphical models that have been applied in medicine due to their ability to represent expert knowledge and reasoning under uncertainty. Here, the performance comparison of Naïve Bayes, Tree Augmented Naïve Bayes (TAN), K-dependence Bayesian classifier (KDB), and Forest Augmented Naïve Bayes (FAN) models is presented. These models were built with a set of 67 features of location, intensity, shape, and texture; and evaluated with the leave-one-out cross validation method. The previously mentioned features were extracted from a public database called mini-Mias. The experimental results have shown that TAN and FAN models with a subset of only eleven features have achieved a performance of 0.81 in accuracy, 0.65 in sensitivity and 0.89 in specificity. Therefore, these models which allow dependencies between variables (features) are considered suitable and promising methods for automated mass classification.

Résumé

Actuellement, le cancer du sein occupe la première place chez les femmes en tant que cause de décès, par le cancer, au Mexique. Le travail présenté dans ce document examine l'utilisation des réseaux Bayésiens dans la classification des masses en images de mammographie. Ces réseaux sont des modèles graphiques probabilistes qui ont été utilisés en Médecine, pour leur capacité à enregistrer le savoir des experts et travailler sur des domaines incertains. On présente ici une comparaison de l'accomplissement de modèles Bayésien Naïf, le réseau Bayésien Naïf augmenté par un Arbre (TAN), le Classifieur Bayésien K-dépendant (KDB) et le réseau Bayésien Naïf augmenté par une forêt (FAN). Ces modèles ont été construits avec un ensemble de 67 caractéristiques de position, d'intensité, de forme et de texture ; et ils ont été évalués avec la méthode Leave-one-out cross validation. Les caractéristiques mentionnées ont été extraites d'un ensemble de données publiques dénommées mini-MIAS. Les résultats expérimentaux montrent que les modèles de type TAN et FAN, construits avec un sous-ensemble de onze caractéristiques, réussissent à obtenir un accomplissement de 0.81 en exactitude, 0.65 en sensibilité et 0.89 en efficacité. Cependant, on considère que ces modèles, qui permettent des dépendances entre variables (caractéristiques), sont adéquats et prometteurs pour la classification automatique de masses.

Rolando Pedro Gabriel, Verónica Rodríguez López, Raúl Cruz Barbosa

Palabras clave: Clasificación, Cáncer de Mama, Mamografía, Modelos Gráficos.

Introducción

El cáncer de mama es un cáncer que se forma en los tejidos de la mama, por lo general en los conductos y los lobulillos (Instituto Nacional del Cáncer, 2013). Es considerado un problema de salud mundial y en México, a partir del 2006, ocupa el primer lugar como causa de muerte

Universidad Tecnológica de la Mixteca

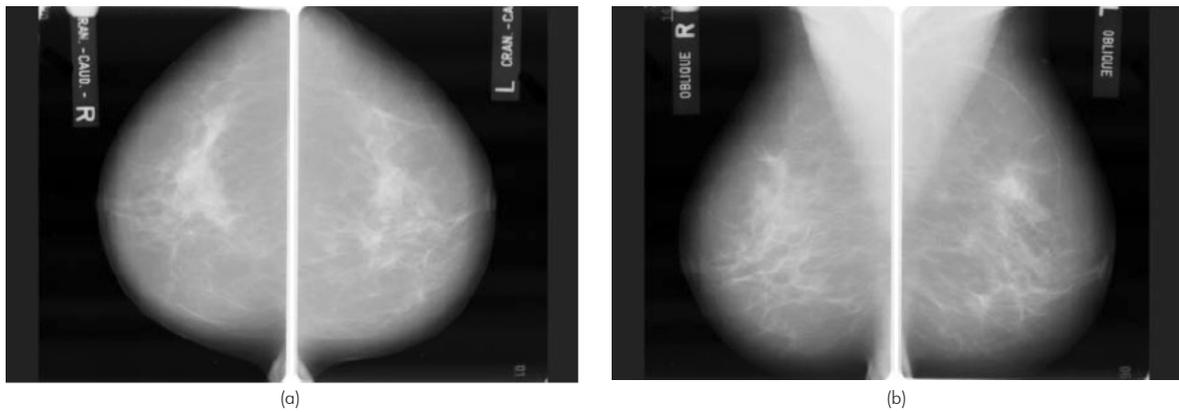


Fig. 1. Mamografía de la mama derecha: (a) proyección cráneo-caudal y (b) proyección medio-lateral-oblicua.

por cáncer entre las mujeres (Cárdenas Sánchez, et al., 2013). Hasta el momento, la mamografía sigue siendo considerada el estudio fundamental para la detección temprana de este tipo de cáncer (Leod & Verma, 2013).

La mamografía es un examen de la glándula mamaria que consiste en la obtención de diversas proyecciones radiográficas de la mama. Un examen mamográfico típico consiste en la obtención de dos proyecciones de cada mama (ver Fig. 1): cráneo-caudal (CC) y medio-lateral-oblicua (MLO)¹. La proyección CC permite la evaluación del tejido mamario y/o localización de lesiones de los cuadrantes interno (inferior en la placa) y externo (superior en la placa) de la mama. Mientras que con la MLO, se pueden evaluar el tejido mamario y/o localización de lesiones en cuadrantes superiores e inferiores de la mama (Instituto Nacional del Cáncer de Argentina, 2012).

Las masas son uno de los signos más importantes de la presencia del cáncer y que un radiólogo busca al analizar mamografías. Este tipo de lesiones son estructuras tridimensionales que tienen, total o parcialmente, bordes externos convexos y son evidentes en al menos dos proyecciones de la mama (Pant, 2011). Los radiólogos para formular la interpretación de una mamografía deben caracterizar estas lesiones de acuerdo a su ubicación, forma, márgenes y densidad. La forma de una masa puede ser redonda, oval o irregular; sus márgenes pueden ser circunscritos, microlobulados, mal definidos y espiculados; y la densidad puede ser alta, isodensa

(igual densidad), baja (baja atenuación, pero no contiene grasa), y radiolúcida (contiene grasa) (Bozek, et al., 2009; Velasquez, 1997).

Existe una gran complejidad en la detección y descripción de masas, debido principalmente a las características del tejido mamario, y a la calidad de las imágenes de mamografía. Muchas veces, las masas no pueden ser visualizadas debido a que la mama está constituida por tejidos muy similares entre sí, y porque, algunas veces las lesiones son muy pequeñas y se confunden con el tejido mamario normal (Bozek, et al., 2009). La interpretación correcta de mamografías depende en gran medida de la experiencia del radiólogo, pudiendo existir una alta variabilidad (entre el 65-75%) cuando es analizada por diversos radiólogos (Cheng, et al., 2006).

A fin de mejorar el desempeño en el análisis de mamografías, se han desarrollado sistemas computacionales que apoyan en el diagnóstico de lesiones (CADx-Computer-Aided Diagnosis). Una de las tareas de estos sistemas, es la clasificación de lesiones como benignas o malignas. En este trabajo, se presenta la investigación realizada sobre el uso de las Redes Bayesianas en la clasificación.

Las Redes Bayesianas son modelos gráficos probabilísticos que tienen la capacidad de trabajar en dominios inciertos y de capturar el conocimiento de expertos. Debido a estas ventajas han sido ampliamente utilizados en el área de la medicina, principalmente en problemas de diagnóstico (Antal, et al., 2004; Aussem, et al., 2012; Flores, et al., 2011; van der Gaag, et al., 2002; Wu, et al., 2001).

También se ha explorado la aplicación de las redes Bayesianas, en el diagnóstico del cáncer de mama a partir del análisis de imágenes de mamografía.

¹ Terminología establecida por Breast Imaging Reporting and Data System, la primera palabra del nombre indica la posición del tubo de rayos X, y la segunda indica la localización del receptor de imagen.

Existe el reporte de su buen desempeño, cuando son entrenadas a partir de datos clínicos del paciente y hallazgos mamográficos descritos por radiólogos especializados (Burnside, et al., 2009; Burnside, et al., 2000; Fischer, et al., 2004; Kahn Jr, et al., 1997). Sin embargo, poco se ha analizado sobre su desempeño cuando son entrenados a partir de características obtenidas de forma automática (Patrocínio, et al., 2004; Samulski, et al., 2007; Velikova, et al., 2007; Velikova, et al., 2009; Velikova, et al., 2012). Por esto último, y considerando que las redes Bayesianas son modelos gráficos sencillos de interpretar y validar por expertos, que permiten combinar datos con el conocimiento especializado, además de que son capaces de manipular información incompleta, este trabajo tiene como objetivo analizar su desempeño en la discriminación de masas benignas y malignas, utilizando características obtenidas de forma automática.

La organización de este artículo es como sigue. En la Sección 2 se describe el conjunto de datos utilizados, el conjunto de características extraídas y los tipos de Redes Bayesianas analizadas. Los experimentos realizados y los resultados obtenidos se explican en la Sección 3. Finalmente, en la Sección 4 se presentan las conclusiones y el trabajo a futuro.

Materiales y Métodos

Datos

Para el análisis de clasificación de este trabajo, se utilizó un subconjunto de 58 imágenes de la base de datos pública *Mammographic Image Analysis Society* (mini-MIAS) (Suckling, 1994). Mini-MIAS es una de las pocas bases de datos pública de imágenes de mamografía que ha sido utilizada en diversos proyectos de investigación. Esta base de datos contiene imágenes de proyecciones mamográficas del tipo medio-lateral-

oblicua, de la mama izquierda y derecha de 161 pacientes. También incluye anotaciones sobre el tipo, tamaño aproximado y ubicación de la lesión presente en cada imagen.

El conjunto de datos utilizado en esta investigación se conformó con imágenes que contienen únicamente la región de la masa (ROI). Estos ROIs rectangulares se extrajeron de las imágenes originales con la ayuda del programa ImageJ (Rasband, 2004) y considerando las anotaciones del MIAS. En nuestro conjunto de datos se incluyeron masas malignas y benignas, con bordes circunscritos, mal definidos y espiculados. En el Cuadro 1 se resume nuestro conjunto de datos y en la Fig. 2 se presentan algunos ejemplos de estas imágenes.

Cuadro 1. Descripción del conjunto de imágenes de prueba

| Categoría \ Forma | benignas | malignas |
|-------------------|----------|----------|
| circunscritas | 20 | 4 |
| mal definidas | 7 | 8 |
| espiculadas | 11 | 8 |
| TOTAL | 38 | 20 |

Extracción de características

De cada ROI se extrajeron 67 descriptores de ubicación, intensidad, forma, y textura. Este conjunto de características fueron seleccionadas considerando la literatura médica (Sickles, 1989) y trabajos relacionados con la clasificación de masas (Cheng, et al., 2006; Delogu, et al., 2008; Samulski., 2006).

Ubicación

La ubicación de la lesión es un parámetro importante debido a que aproximadamente el 45% de las masas malignas se desarrollan en el cuadrante superior

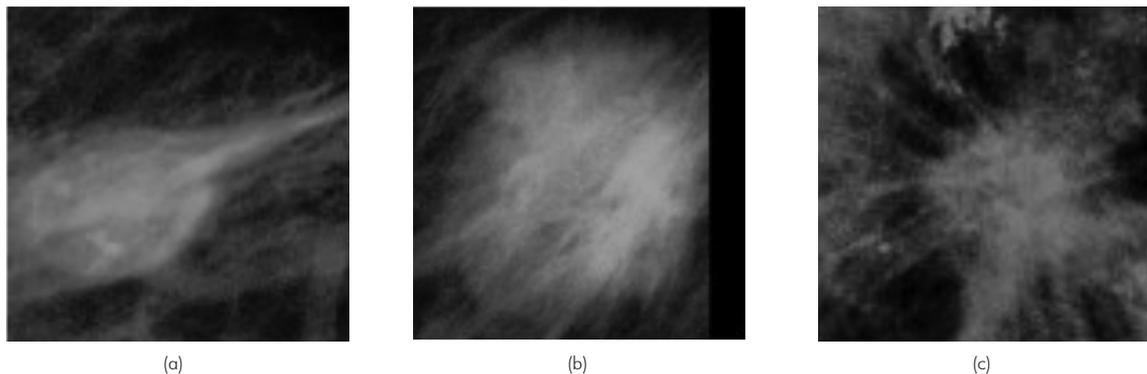


Fig. 2. Ejemplos de imágenes de masa del conjunto de prueba: (a) circunscrita, (b) mal definida y (c) espiculada.

externo de la mama. Para estimar la ubicación de la masa se aplicó el procedimiento descrito en (Samulski., 2006), el cual utiliza a toda la imagen de mamografía, y transforma el sistema de coordenadas, de forma tal, que el *eje y* coincide con el borde del músculo pectoral, y una perpendicular a este eje (dirigida hacia el pezón), se convierte en el *eje x*. En este nuevo sistema de coordenadas, se estimó la ubicación, (U_x, U_y) , como las coordenadas del centro de la región de la masa.

Descriptores de intensidad

Este tipo de mediciones fueron utilizados para describir la densidad de una masa, que es un parámetro importante en el diagnóstico. La densidad se refiere a la cantidad de tejido graso en la masa y a cómo se comporta el tejido de ésta con respecto al contenido en las estructuras densas de la mama (Jackson, et al., 1991).

A partir de los niveles de intensidad de los píxeles que pertenecen a la región de la masa, se calcularon el contraste y cinco estadísticos simples como descriptores de intensidad. Los estadísticos utilizados fueron media, varianza, mediana, asimetría y curtosis (Gonzalez & Woods, 2006).

Contraste: Es la diferencia relativa entre la intensidad promedio de los niveles de gris de los píxeles internos de la región de la masa $E(M)$, y la intensidad promedio de los niveles de gris de los píxeles que pertenecen a la banda que rodea a esta región $E(O)$ (Brake, 2000),

$$(1) \text{Contraste}(M) = E(M) - E(O)$$

Para construir la banda, se consideraron a todos los píxeles que se encontraban a una distancia de $0.6R$ del borde de la región de la masa. Donde R se refiere al radio aproximado de la masa el cual fue calculado de acuerdo a la siguiente ecuación,

$$(2) R = \sqrt{A(M)/\pi}$$

$A(M)$ representa el área de la masa.

Descriptores de forma

Los descriptores de forma fueron calculados para capturar la forma y los márgenes de la masa. Los

márgenes, que se refiere a los límites o contorno de la masa, son evaluados conjuntamente con la forma para determinar la posible malignidad (Sickles, 1989). A continuación se describen las mediciones de forma calculadas para cada ROI.

Área $A(M)$: Se define como el número de píxeles contenidos dentro de los límites de la región de la masa. Este descriptor proporciona información sobre el tamaño de la masa.

Perímetro $P(M)$: Es el número de píxeles que se encuentran en el contorno de la región de la masa.

Compacidad: Este descriptor indica qué tan compacta es la masa, mientras más compacta sea la región, su valor será cercano a uno. Esta característica relaciona el perímetro $P(M)$, y el área $A(M)$ de la región por medio de la siguiente ecuación (Samulski., 2006),

$$(3) C(M) = (P(M)^2) / 4\pi A(M)$$

Longitud radial normalizada (LRN): Se define como la distancia Euclidiana normalizada que existe entre el centro y un punto del contorno de la región de la masa. La distancia se normaliza al considerar la máxima distancia Euclidiana que existe hacia un punto del borde (Tsui, et al., 2010). Los descriptores de la LRN proporcionan información acerca del grado de irregularidad en los márgenes de la masa. Por lo que, a partir de la distancia radial normalizada $d(i)$ de todos los píxeles del contorno, se obtuvieron las siguientes mediciones: media (d_{AVG}), desviación estándar (d_{STD}), entropía (d_{ENT}), índice de área (d_{IA}), rugosidad del contorno (d_{IR}) y cruces por cero (d_{CC}) (Delogu, et al., 2008).

$$(4) d_{AVG} = \sum_{i=1}^{P(M)} d(i) / P(M)$$

$$(5) d_{STD} = \sqrt{\frac{1}{P(M)} \sum_{i=1}^{P(M)} (d(i) - d_{AVG})^2}$$

$$(6) d_{ENT} = - \sum_{i=1}^{Nbins} p_k \log(p_k)$$

Donde p_k es la probabilidad de que la LRN se encuentre entre $d(i)$ y $d(i+1) + 1/Nbins$, con $Nbins$ como el número de intervalos del histograma de la LRN.

$$(7) \quad d_{IA} = \frac{1}{d_{AVG} * P(M)} \sum_{i=1}^{P(M)} (d(i) - d_{AVG})$$

$$(8) \quad d_{IR} = \frac{1}{P(M)} \sum_{i=1}^{P(M)} |d(i) - d(i+1)|$$

$$(9) \quad d_{CC} = \sum_{i=1}^{P(M)} h(d(i), d_{AVG})$$

donde,

$$h(d(i), d_{AVG}) = \begin{cases} 1 & \text{si } d(i) > d_{AVG} \\ 0 & \text{en otro caso} \end{cases}$$

Momentos invariantes de Hu: Estos momentos son descriptores que son invariantes a traslación, rotación y cambio de escala. Para esta investigación, se calcularon los siete momentos de Hu para cada ROI (Pajares & Cruz., 2008).

Espiculación: El grado de espiculación de una masa proporciona información acerca de su posible malignidad. Las masas con bordes espiculados se caracterizan por tener un patrón estrellado el cual está constituido por un conjunto de líneas que apuntan hacia el centro de la lesión. Para medir el grado de espiculación de una región, se utilizaron los descriptores, *F1_estrellado*, *F2_estrellado*, *F1_estrellado_prom*, *F2_estrellado_prom* y *textura lineal*, propuestos en (Brake, 2000).

F1_estrellado: Un patrón estrellado puede identificarse al detectar un incremento en el número de píxeles que tienen una orientación lineal hacia un píxel en particular. El descriptor *F1_estrellado*, es una medida normalizada de la fracción de píxeles con orientación lineal que apunta hacia el centro de la masa (cx, cy) . Para estimar esta característica, se analizan todos aquellos píxeles que pertenecen a una vecindad circular S , la cual está constituida por píxeles que se localizan a una distancia entre $rmin$ y $rmax$ del centro de la masa (se consideró como $rmax$ a la distancia mínima que existe entre el centro y los límites de la imagen del ROI, y como $rmin=rmax/4$). Un píxel de S se cuenta como parte de *F1_estrellado*, si tiene una orientación lineal que apunta hacia un disco de radio $R = rmin/2$ y centro (cx, cy) .

F2_estrellado: Este descriptor ayuda a determinar si en la región existen espículas con diversas orientaciones. Mientras mayor sea el número de espículas con diferentes orientaciones, mayor será la proba-

bilidad de la presencia de un patrón estrellado en la región. Esta medición se obtiene al dividir la vecindad circular S en cuatro sectores circulares y estimando en qué proporción el descriptor *F1_estrellado* se distribuye uniformemente en cada uno de los sectores.

F1_estrellado_prom: Este descriptor se estima a partir del promedio de los valores de *F1_estrellado* correspondientes a cada uno de los píxeles de la masa. En este caso, para obtener el *F1_estrellado* de un píxel, cada píxel es considerado como el centro del patrón estrellado.

F2_estrellado_prom: Para estimar este descriptor, se calcula el promedio de los valores de *F2_estrellado* para cada uno de los píxeles que constituyen la masa.

Textura lineal: Este descriptor también ayuda a detectar la presencia de un patrón lineal en la región. Para estimar esta medición, se calcula la orientación lineal de todos los píxeles de la región. Posteriormente, utilizando la representación de doble ángulo, se suman todos los vectores resultantes de esta orientación lineal (Johansson, 2000), el descriptor se obtiene finalmente, al calcular la longitud de este vector de suma total, dividido entre la suma de la longitud de cada vector.

Descriptores de Textura

Los descriptores de las matrices de Co-ocurrencias de niveles de gris (GLCM), Diferencias de niveles de gris (GLDM) y Longitud de secuencias de niveles de gris (GLRLM) fueron utilizados para caracterizar a la densidad de la masa. A continuación se explican las mediciones obtenidas de cada ROI.

GLCM: Los valores de esta matriz representan el número de pares de píxeles con intensidades $I1$ y $I2$, que se encuentran separados por una distancia d y un ángulo de inclinación ϑ (Nixon & Aguado, 2012). Con este método, las mediciones que se pueden obtener son: momento angular de segundo orden, contraste, entropía, energía, correlación, homogeneidad, momento de la diferencia inversa, varianza, suma promedio, suma varianza y suma entropía. A partir del promedio y del rango de estas mediciones, para $d=1$ y $\vartheta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$, se obtuvieron 22 descriptores de textura.

GLDM: Los elementos de esta matriz representan la diferencia absoluta entre las intensidades de un par de píxeles que se encuentran separados por una distancia d y un ángulo de inclinación ϑ . A partir de

las matrices GLDM para $d=1$ y $\vartheta=\{0^\circ,45^\circ,90^\circ,135^\circ\}$, se calcularon las siguientes mediciones: media, entropía, contraste, y varianza. El valor promedio y el rango de cada medición fueron utilizados como descriptores.

GLRLM: Cada elemento de esta matriz representa el número de veces que un pixel con intensidad l , aparece en una secuencia de longitud j y con orientación ϑ (Tang, 1998). A partir de las matrices GLRLM para $\vartheta=\{0^\circ,45^\circ,90^\circ,135^\circ\}$, se calcularon las siguientes mediciones: énfasis de secuencia corta (SRE), énfasis de secuencia larga (LRE), niveles de gris no uniforme (GLNU) y longitud de secuencias no uniforme (RLNU). Con el promedio y el rango de estas mediciones se obtuvieron otros ocho descriptores para la textura.

Clasificadores

Una red Bayesiana es un tipo de modelo gráfico probabilístico en donde los nodos representan variables y los arcos dependencias entre las variables. Formalmente, una red Bayesiana es un par (D,P) , donde D es un grafo acíclico dirigido, $P=\{p(x_1|\pi_1), p(x_2|\pi_2), \dots, p(x_n|\pi_n)\}$ es un conjunto de n distribuciones de probabilidad condicional, uno para cada variable, y π_i es el conjunto de padres del nodo X_i en D (Castillo, et al., 1998).

Diversos modelos de redes Bayesianas (RB) han sido propuestos como clasificadores, algunos de ellos son: Naive-Bayes, Naive Bayes aumentado a árbol

(*Tree augmented Naive-Bayes – TAN*), Clasificador Bayesiano K-dependiente (*K-dependence Bayesian classifier-KDB*) y Naive Bayes aumentado a bosque (*Forest augmented Naive Bayes-FAN*).

Naive-Bayes

Es el modelo más simple de redes Bayesianas en el cual el nodo de clase es el padre de todos los nodos atributo, tal y como se muestra en la Fig. 3(a). Este tipo de red Bayesiana no requiere de un proceso de aprendizaje estructural, y además asume que todas las características son independientes entre sí. Debido a estas ventajas, es un modelo que ha sido utilizado en diversas tareas de clasificación (Broos, et al., 2011; Cruz-Ramírez, et al., 2007; Samulski, et al., 2007; Sharma & Mukherjee, 2012).

Naive Bayes aumentado a árbol (*Tree Augmented Naive Bayes -TAN*)

Este modelo es una extensión de una Naive-Bayes en el que se permite que los nodos atributo formen un árbol. La propiedad más importante que debe cumplir una TAN, es que cada nodo atributo debe tener a lo más dos padres: el nodo clase y algún otro nodo atributo (ver Fig. 3. Modelos de Redes Bayesianas para clasificación: (a) Naive-Bayes, (b) TAN, (c) KDB, y (d) FAN.3(b)).

Uno de los métodos más utilizados para aprender la estructura de una TAN es el propuesto por Friedman

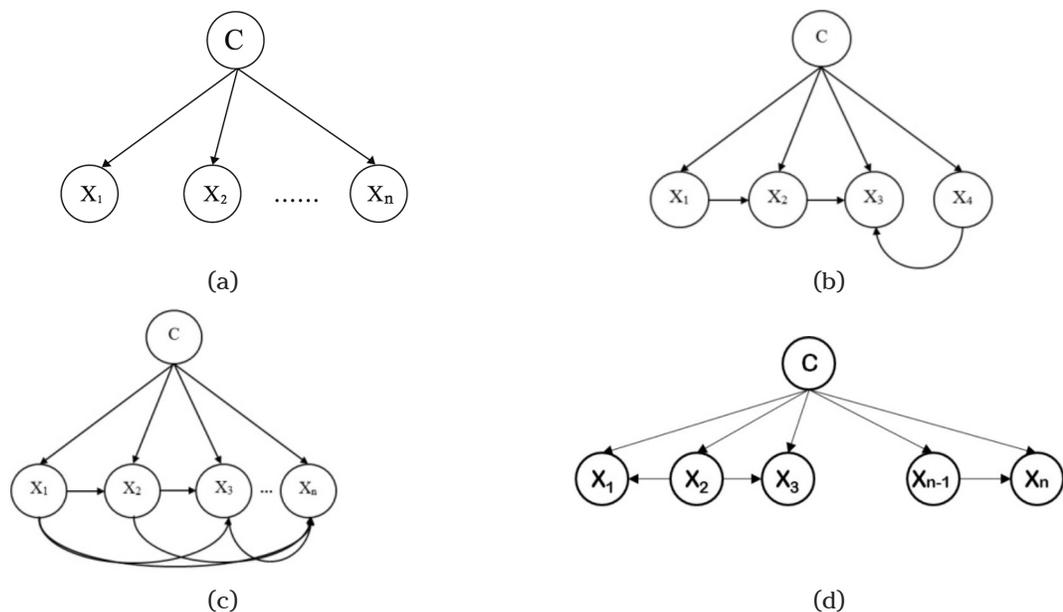


Fig. 3. Modelos de Redes Bayesianas para clasificación: (a) Naive-Bayes, (b) TAN, (c) KDB, y (d) FAN.

(1997). Este algoritmo es una variante del método de Chow-Liu (1968), que utiliza la información mutua condicional entre dos atributos dada la clase para encontrar la mejor estructura.

Clasificador Bayesiano K-dependiente (K-dependence Bayesian classifier-KDB)

Un clasificador KDB es una red Bayesiana que también es una extensión de una Naive-Bayes. En este modelo, cada atributo puede tener como padres a k - nodos atributos, además del nodo clase (Sahami, 1996). En la Fig. 3. Modelos de Redes Bayesianas para clasificación: (a) Naive-Bayes, (b) TAN, (c) KDB, y (d) FAN.3(c) se presenta un ejemplo de este tipo de clasificadores.

El algoritmo propuesto por Sahami (1996) permite aprender la estructura de este tipo de clasificadores. Este método utiliza la información mutua condicional entre dos atributos dada la clase y un valor de umbral (de información mutua) para encontrar las relaciones entre los atributos.

Naive Bayes aumentado a bosque (Forest augmented Naive Bayes-FAN)

Este modelo es una variante de una TAN que permite que los nodos atributos formen un grafo del tipo bosque. Una de las ventajas de la FAN es que elimina relaciones innecesarias entre los atributos que son consideradas, algunas veces, por una TAN. El algoritmo propuesto en Lucas (2002) puede ser utilizado para aprender la estructura de este tipo de clasificadores. Un ejemplo de este tipo de modelo se presenta en la Fig. 3. Modelos de Redes Bayesianas para clasificación: (a) Naive-Bayes, (b) TAN, (c) KDB, y (d) FAN.3(d).

Experimentos y Resultados

Los objetivos principales de nuestros experimentos son tres. Primero, analizar el desempeño de diversos modelos de redes Bayesianas en la discriminación de masas benignas y malignas. Segundo, determinar cuál es el modelo de red Bayesiana más adecuado para el diagnóstico. Finalmente, analizar la capacidad de las características para describir masas benignas y malignas.

Los tipos de modelos analizados fueron Naive-Bayes, TAN, KDB y FAN. Todas las redes Bayesianas se construyeron y evaluaron en Matlab® 7.10.0.499 (R2010a), con la ayuda de *Bayes Net toolbox* de Kevin

Murphy (Murphy, 2007) y *BNT Structure Learning Package* de Philippe Leray (Francois, 2007). Para construir los modelos, de nuestro conjunto de 58 ROIS con masas se extrajeron las 67 características de ubicación, intensidad, forma y textura que se describen en la Sección . Los valores de las características fueron estandarizados de tal forma que los valores de los datos tuvieran media cero y desviación estándar unitaria, y discretizados utilizando el método de Intervalos por igual amplitud (*Equal Width Discretization*).

El desempeño de los clasificadores fue estimado a partir del método *Leave-one-out cross validation*. Las métricas utilizadas para evaluar el desempeño fueron: exactitud, sensibilidad y especificidad. La exactitud se refiere a la proporción de masas correctamente clasificadas con respecto al total de masas. Sensibilidad es la proporción de verdaderos positivos con respecto al número de masas malignas; y especificidad es la proporción de verdaderos negativos con respecto al total de masas benignas. En este caso, un resultado se considera como verdadero positivo, cuando una lesión maligna es también catalogada como maligna por el clasificador; y un resultado es verdadero negativo cuando el clasificador determina correctamente que la lesión es benigna (Sampat, et al., 2005).

Evaluación de redes Bayesianas con el conjunto completo de características

Como primer experimento se construyeron redes Bayesianas del tipo Naive-Bayes, TAN, KDB ($k=2$) y FAN, utilizando el conjunto completo de 67 características que se describen en la Sección . Los resultados obtenidos para cada clasificador se muestran en el Cuadro 2. Como se puede observar en este cuadro, el desempeño de todos los clasificadores fue muy bajo, lo cual puede explicarse debido al número de características y a la poca cantidad de datos (58 muestras).

Cuadro 2. Desempeño obtenido por los clasificadores con el conjunto completo de características

| Modelo | Exactitud | Sensibilidad | Especificidad |
|-------------|-----------|--------------|---------------|
| Naive-Bayes | 0.43 | 0.2 | 0.55 |
| TAN | 0.27 | 0.2 | 0.31 |
| KDB | 0.32 | 0.2 | 0.39 |
| FAN | 0.27 | 0.2 | 0.31 |

Evaluación de redes Bayesianas con un subconjunto de características

En el segundo experimento, se realizó una selección de características, debido que era deseable obtener un modelo compacto que describiera fácilmente las relaciones entre las características y la clase. Para el proceso de selección de características se aplicó el método de Análisis de discriminantes de Fisher (Theodoridis & Koutroumbas, 2008).

El subconjunto de características se obtuvo al seleccionar primero veintidós características que se valoraron

como relevantes. Las primeras veinte de esta selección, correspondían a las más discriminativas en el espacio total de acuerdo a su factor de Fisher, y las últimas fueron las siguientes dos más importantes dentro del grupo de características de intensidad (ver Cuadro 3). Estas mismas características se muestran en el Cuadro 4 considerando que todos los tipos (ubicación, intensidad, forma y textura) son importantes para la clasificación de masas. El número de éstas se determinó de manera heurística, tratando de reducir a una tercera parte la dimensión del espacio total de características.

Cuadro 3. Primera selección de características ordenadas de acuerdo a su factor de discriminación de Fisher

| Num. | Característica | Factor de discriminación de Fisher | Trabajos relacionados |
|------|--|------------------------------------|---|
| 1 | compacidad ($C(M)$) | 0.2777 | Samulski., 2006 |
| 2 | perímetro ($P(M)$) | 0.1825 | Cheng, et al., 2006 |
| 3 | cruces por cero LRN (d_{cc}) | 0.1703 | Delogu, et al., 2008; Tsui, et al., 2010 |
| 4 | ubicación (U_y) | 0.1661 | Samulski, et al., 2007; Samulski., 2006 |
| 5 | contraste de intensidad ($Contraste(M)$) | 0.1365 | Brake, 2000; Samulski., 2006 |
| 6 | rango de suma promedio GLCM | 0.1349 | Cheng, et al., 2006 |
| 7 | rango de varianza GLCM | 0.0943 | Cheng, et al., 2006 |
| 8 | textura lineal | 0.089 | Brake, 2000; Samulski., 2006 |
| 9 | F1 promedio ($F1_estrellado_prom$) | 0.0848 | Brake, 2000; Samulski., 2006 |
| 10 | desviación estándar LRN (d_{std}) | 0.0833 | Delogu, et al., 2008; Tsui, et al., 2010 |
| 11 | media LRN (d_{avg}) | 0.0767 | Delogu, et al., 2008; Tsui, et al., 2010 |
| 12 | F2 estrellado ($F2_estrellado$) | 0.0697 | Brake, 2000; Samulski., 2006 |
| 13 | 1er. momento de Hu | 0.0624 | Cheng, et al., 2006 |
| 14 | rango de varianza GLCM | 0.0537 | Cheng, et al., 2006 |
| 15 | 3er. momento de Hu | 0.0531 | Cheng, et al., 2006 |
| 16 | F2 promedio ($F2_estrellado_prom$) | 0.048 | Brake, 2000; Samulski., 2006 |
| 17 | promedio de varianza GLDM | 0.0447 | Cheng, et al., 2006 |
| 18 | promedio de correlación GLCM | 0.0431 | Cheng, et al., 2006 |
| 19 | área ($A(M)$) | 0.0417 | Brake, 2000; Samulski., 2006 |
| 20 | 7o. momento de Hu | 0.0331 | Cheng, et al., 2006 |
| 21 | mediana de intensidad | 0.0077 | Cheng, et al., 2006; Delogu, et al., 2008 |
| 22 | media de intensidad | 0.0055 | Cheng, et al., 2006; Delogu, et al., 2008 |

Cuadro 4. Primera selección de características ordenadas de acuerdo a su factor de discriminación y tipo

| Tipo de característica | Característica | Factor de discriminación de Fisher | Trabajos relacionados |
|------------------------|--|------------------------------------|---|
| Ubicación | ubicación (U_j) | 0.1661 | Samulski, et al., 2007; Samulski., 2006 |
| Intensidad | contraste ($Contraste(M)$) | 0.1365 | Brake, 2000; Samulski., 2006 |
| | mediana | 0.0077 | Cheng, et al., 2006; Delogu, et al., 2008 |
| | media | 0.0055 | Cheng, et al., 2006; Delogu, et al., 2008 |
| Forma | compacidad ($C(M)$) | 0.2777 | Samulski., 2006 |
| | perímetro ($P(M)$) | 0.1825 | Cheng, et al., 2006 |
| | cruces por cero LRN (d_{CC}) | 0.1703 | Delogu, et al., 2008; Tsui, et al., 2010 |
| | textura lineal | 0.089 | Brake, 2000; Samulski., 2006 |
| | F1 promedio ($F1_estrellado_prom$) | 0.0848 | Brake, 2000; Samulski., 2006 |
| | desviación estándar LRN (d_{STD}) | 0.0833 | Delogu, et al., 2008; Tsui, et al., 2010 |
| | media LRN (d_{AVG}) | 0.0767 | Delogu, et al., 2008; Tsui, et al., 2010 |
| | F2 estrellado ($F2_estrellado$) | 0.0697 | Brake, 2000; Samulski., 2006 |
| | 1er. momento de Hu | 0.0624 | Cheng, et al., 2006 |
| | 3er. momento de Hu | 0.0531 | Cheng, et al., 2006 |
| | F2 promedio ($F2_estrellado_prom$) | 0.048 | Brake, 2000; Samulski., 2006 |
| | área ($A(M)$) | 0.0417 | Brake, 2000; Samulski., 2006 |
| 7o. momento de Hu | 0.0331 | Cheng, et al., 2006 | |
| Textura | rango de suma promedio GLCM | 0.1349 | Cheng, et al., 2006 |
| | rango de varianza GLCM | 0.0943 | Cheng, et al., 2006 |
| | rango de suma varianza GLCM | 0.0537 | Cheng, et al., 2006 |
| | promedio de varianza GLDM | 0.0447 | Cheng, et al., 2006 |
| | promedio de correlación GLCM | 0.0431 | Cheng, et al., 2006 |

Finalmente, con diversas combinaciones de este subconjunto de veintidós características, se construyeron clasificadores del tipo Naive-Bayes y TAN. Las combinaciones se obtuvieron al ir eliminando características en cada prueba, y la mejor que se obtuvo quedó constituida por las once características que se presentan en el Cuadro 5.

Cuadro 5. Subconjunto final de características

| | |
|-------------------|---|
| Intensidad | F1:mediana F2:media |
| Forma | F3:área ($A(M)$) F4:perímetro ($P(M)$) F5:compacidad ($C(M)$) F6:media LRN (d_{AVG}) F7:desviación estándar LRN (d_{STD}) F8:cruces por cero LRN (d_{CC}) F9:3er. momento de Hu F10:textura Lineal |
| Textura | F11:rango de varianza GLCM |

Haciendo un análisis del subconjunto final de características, se observa que éstas son parámetros importantes para el diagnóstico: tamaño, forma, márgenes y densidad de la masa. El tamaño se describe con las características F3 y F4; la forma, con F5 y F9; los márgenes, con F6, F7, F8 y F10; y la densidad, con F1, F2 y F11. De estas once características, nueve son consistentes con la literatura médica; sólo el tamaño no es considerado relevante para el diagnóstico, pero es un parámetro importante que ayuda al radiólogo a determinar el tratamiento adecuado (Sickles, 1989).

Los modelos obtenidos, del tipo Naive-Bayes, TAN, KDB (k=2) y FAN, con el conjunto de once características seleccionadas se muestran en la Fig. 4 y en el Cuadro 6. Desempeño obtenido por los clasificadores con el subconjunto seleccionado de once características⁶ los resultados correspondientes. En este caso, Naive-Bayes fue el modelo con el peor desempeño, y los mejores fueron TAN y FAN con un desempeño similar. El desempeño superior de los modelos TAN, FAN y KDB, demuestra que existen dependencias entre las características, las cuales, de acuerdo a los resultados, fueron capturadas mejor por la TAN y FAN. De hecho, como se muestra en la Fig. 4, estos modelos son casi equivalentes.

Cuadro 6. Desempeño obtenido por los clasificadores con el subconjunto seleccionado de once características

| Modelo | Exactitud | Sensibilidad | Especificidad |
|-------------|-----------|--------------|---------------|
| Naive-Bayes | 0.63 | 0.4 | 0.76 |
| TAN | 0.81 | 0.65 | 0.89 |
| KDB | 0.74 | 0.55 | 0.84 |
| FAN | 0.81 | 0.65 | 0.89 |

Los resultados obtenidos por los mejores modelos TAN y FAN, en sensibilidad y especificidad, explican que estos modelos tienen una capacidad aceptable para clasificar masas benignas, y regular, para masas malignas. Esto último indica que las características contempladas por estos modelos, describen adecuadamente las masas benignas, y que es necesario considerar otro tipo de características que ayuden a describir mejor las propiedades de las masas malignas.

Una interpretación subjetiva para la topología de los modelos TAN y FAN obtenidos, tomando en cuenta que existe una relación causal entre las características, sería la siguiente. El tamaño de la masa tiene influencia en la forma, factor que a su vez afecta a los márgenes. También, en estos modelos se puede identificar, que los márgenes repercuten en el tamaño y en

la densidad de la masa; y que la densidad influye tanto en los márgenes como en la forma. Aunque es difícil determinar la validez de estas relaciones sin la ayuda de un experto, considerando la literatura médica, se puede decir que estas relaciones reflejan la manera en la que el experto efectúa el análisis de una masa para determinar su diagnóstico. No es común que el radiólogo realice la evaluación de características de forma independiente, regularmente considera a más de una: márgenes y tamaño son analizados conjuntamente; al analizar la forma, también observan los márgenes; y la densidad se evalúa con la descripción de los márgenes, y algunas veces, con la de forma (Jackson, et al., 1991; Sickles, 1989).

Comparación de los modelos de Redes Bayesianas con otro tipo de clasificadores

En el Cuadro 7. Comparación del desempeño de nuestros mejores modelos de Redes Bayesianas con el obtenido por otro tipo de clasificadores⁷ se presenta una comparación del desempeño de nuestros mejores modelos de redes Bayesianas con el obtenido por una Máquina de Soporte Vectorial (SVM). La SVM también fue implementada en Matlab® con la ayuda de la biblioteca LIBSVM (Chang & Lin, 2011). En este cuadro se puede observar que los resultados de las redes Bayesianas son mejores a los del clasificador SVM, lo que indica que el conocimiento sobre la clasificación de las masas, implícito en los datos, en este caso es mejor descrito por las redes Bayesianas.

Cuadro 7. Comparación del desempeño de nuestros mejores modelos de Redes Bayesianas con el obtenido por otro tipo de clasificadores

| Modelo | Exactitud | Sensibilidad | Especificidad |
|-----------------------------------|-----------|--------------|---------------|
| TAN | 0.81 | 0.65 | 0.89 |
| FAN | 0.81 | 0.65 | 0.89 |
| SVM | 0.76 | 0.60 | 0.84 |
| Bayesiano de Rojas & Nandi (2009) | 0.73 | 0.45 | 0.9 |
| SVM de Rojas & Nandi (2009) | 0.75 | 0.45 | 0.95 |

Por otra parte, también en el Cuadro 7. Comparación del desempeño de nuestros mejores modelos

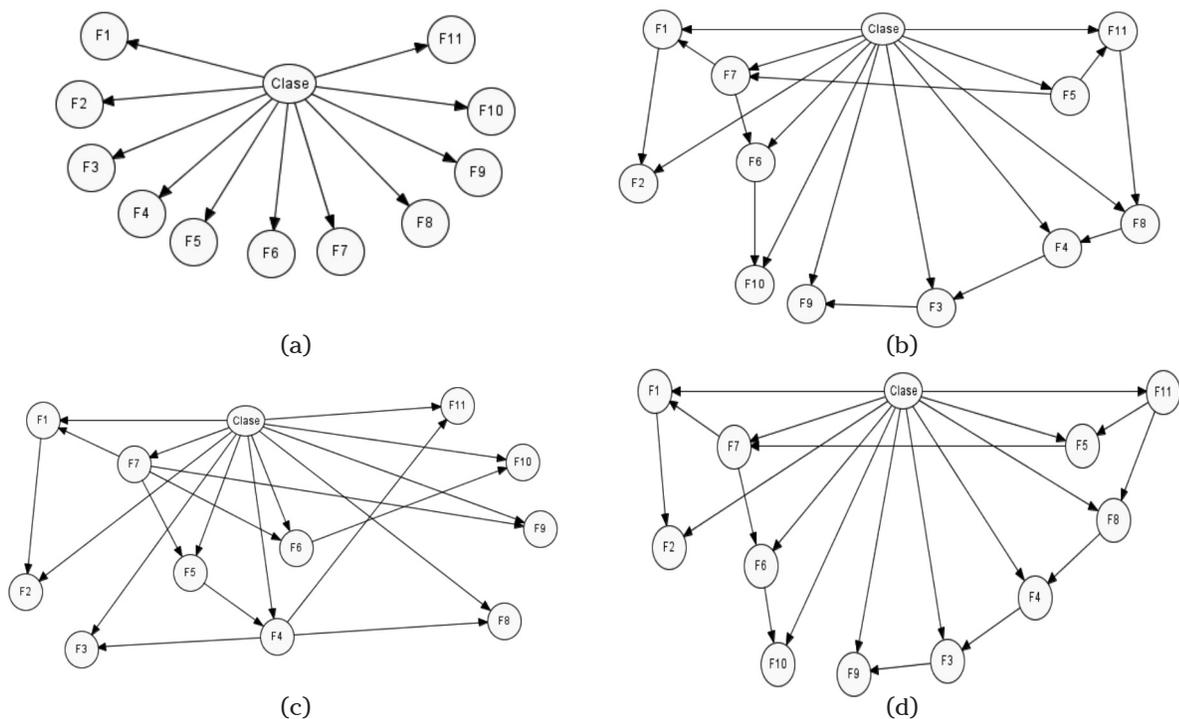


Fig. 4. Topología de las Redes Bayesianas obtenidas con el subconjunto de once características: (a) Naive-Bayes, (b) TAN, (c) KDB y (d) FAN.

de Redes Bayesianas con el obtenido por otro tipo de clasificadores⁷ se muestra una comparación del desempeño de nuestros modelos, con los obtenidos por los clasificadores propuestos por Rojas & Nandi (2009) quienes también utilizan la base de datos mini-MIAS para sus experimentos. Es importante mencionar que es difícil realizar una comparación efectiva, debido a que en la evaluación de estos clasificadores, aunque se utiliza la misma base de datos, varían las condiciones experimentales. Se puede apreciar en este cuadro, el desempeño superior de nuestros modelos en exactitud y sensibilidad, y que la especificidad es comparable y aceptable. También se puede observar, que tanto nuestros modelos como los de Rojas & Nandi (2009), tienen un mejor desempeño en la clasificación de masas benignas. El desempeño superior en sensibilidad de nuestros modelos explica que nuestras características son más apropiadas para la descripción de masas malignas.

Los resultados de esta comparación muestran que nuestros modelos de redes Bayesianas son prometedores para la clasificación de masas. Los modelos obtenidos son compactos, sencillos de validar e interpretar por expertos, se les pueden incorporar

fácilmente datos clínicos del paciente y son capaces de trabajar con información incompleta. Estas últimas ventajas son importantes en problemas de diagnóstico médico y muy difícilmente se pueden obtener con otro tipo de clasificadores.

Conclusiones

La mamografía es considerada como una herramienta fundamental para la detección oportuna del cáncer de mama. Debido a la complejidad de su interpretación y a la sobrecarga de radiólogos especializados, existe la necesidad de proporcionar herramientas computacionales que apoyen en su análisis. En este trabajo se exploró el uso de las redes Bayesianas en la discriminación de masas benignas y malignas, utilizando características extraídas de forma automática. Nuestros resultados experimentales han mostrado que los modelos de redes Bayesianas TAN y FAN, con una sensibilidad de 0.65 y especificidad de 0.89, comparados con los reportados en la literatura (Rojas & Nandi, 2009), son adecuados para esta tarea. Estos modelos, además de tener una estructura compacta, logran capturar parte del conocimiento especializado utilizado por los radiólogos en el diagnóstico. Por esto último, y considerando que la exactitud del 81% obte-

nida por estos modelos es comparable con el desempeño promedio (75%) de un radiólogo (Cheng, et al., 2006) y con el 82% de sistemas CADx experimentales (Sampat, et al., 2005), se puede concluir que las redes Bayesianas construidas con características obtenidas de forma automática, son modelos prometedores para la clasificación de masas.

Como trabajo a futuro se planea realizar la interpretación y validación de nuestros modelos de redes Bayesianas por un experto, así como, comparar su desempeño con el de otro tipo de clasificadores. También, se pretende construir redes Bayesianas que consideren, además de los descriptores de masa, datos clínicos del paciente 

Referencias

- Antal, P., Fannes, G., Timmerman, D., Moreau, Y., & De Moor, B. (2004). Using literature and data to learn Bayesian networks as clinical models of ovarian tumors. *Artificial Intelligence in medicine*, 30 (3), 257-281.
- Aussem, A., De Morais, S. R., & Corbex, M. (2012). Analysis of nasopharyngeal carcinoma risk factors with Bayesian networks. *Artificial Intelligence in Medicine*, 54 (1), 53-62.
- Bozek, J., Mustra, M., Delac, K., & Grgic, M. (2009). A survey of image processing algorithms in digital mammography. *Recent Advances in Multimedia Signal Processing and Communications*, 231, 631-657.
- Brake, G. t. (2000). Computer Aided Detection of Masses in Digital Mammograms. *PhD in Medical Science*. Netherlands: Radboud University Nijmegen.
- Broos, P. S., Getman, K. V., Povich, M. S., Townsley, L. K., Feigelson, E. D., & Garmire, G. P. (2011). A Naive Bayes Source Classifier for X-ray Sources. *The Astrophysical Journal Supplement Series* 194 (1), 4.
- Burnside, E. S., Davis, J., Chhatwal, J., Alagoz, O., Lindstrom, M. J., Geller, B. M., y otros. (2009). Probabilistic Computer Model Developed from Clinical Data in National Mammography Database Format to Classify Mammographic Findings. *Radiology*, 251 (3), 663-672.
- Burnside, E., Rubin, D., & Shachter, R. (2000). A Bayesian network for mammography. *Proceedings of the AMIA Symposium*, (pág. 106).
- Cárdenas Sánchez, J., Bargalló Rocha, E., Erazo Valle, A., Maafs Molina, E., & Poitevin Chacón, A. (2013). Quinta revisión del Consenso Mexicano sobre Diagnóstico y Tratamiento del Cáncer Mamario. *Folleto*. Masson Doyma México.
- Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1998). *Sistemas Expertos y Modelos de Redes Probabilísticas*. Academia Española de Ingeniería.
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2 (3), 1-27.
- Cheng, H. D., Shi, X., Min, R., Hu, L., Cai, X., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition*, 39 (4), 646-668.
- Chow, C. K., & Liu, C. N. (1968). Approximating discrete probability distributions with dependence trees. *IEEE Transactions on Information Theory*, 14, 462-467.
- Cruz-Ramírez, N., Acosta-Mesa, H. G., Carrillo-Calvet, H., Nava-Fernández, L. A., & Barrientos-Martínez, R. E. (2007). Diagnosis of breast cancer using Bayesian networks: A case study. *Computers in Biology and Medicine*, 37 (11), 1553-1564.
- Delogu, P., Fantacci, M. E., Kasae, P., & Retico, A. (2008). Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier. *Computers in Biology and Medicine*, 1479-1491.
- Fischer, E. A., Lo, J. Y., & Markey, M. K. (2004). Bayesian networks of BI-RADS™ descriptors for breast lesion classification. *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, (págs. 3031-3034).
- Flores, M. J., Nicholson, A. E., Brunskill, A., Korb, K. B., & Mascaro, S. (2011). Incorporating expert knowledge when learning Bayesian network structure: A medical case study. *Artificial intelligence in medicine*, 53 (3), 181-204.
- Francois, O. (2007). *Structure Learning Package for Bayes Net Toolbox*. Obtenido de

- <http://www.mathworks.com/matlabcentral/fileexchange/13562-structure-learning-package-for-bayes-net-toolbox>
- Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian Network Classifiers. *Machine Learning* (29), 131-163.
- Gonzalez, R. C., & Woods, R. E. (2006). *Digital Image Processing* (3rd. ed.). USA: Prentice-Hall, Inc.
- Instituto Nacional del Cáncer de Argentina. (2012). Manual operativo para el uso de Mamografía en tamizaje. Ministerio de Salud de la Nación.
- Instituto Nacional del Cáncer. (2013). *Información general sobre el cáncer del seno (mama) (en español)*. Recuperado el 13 de noviembre de 2013, de <http://www.cancer.gov/espanol/pdq/tratamiento/seno/Patient>
- Jackson, V. P., Dines, K. A., Bassett, L. W., Gold, R. H., & Reynolds, H. E. (1991). Diagnostic importance of the radiographic density of noncalcified breast masses: analysis of 91 lesions. *AJR. American journal of roentgenology*, 157 (1), 25-28.
- Johansson, B. (2000). Backprojection of Some Image Symmetries Based on a Local Orientation Description. Technical Report LiTH-ISY-R-2311, Dept. EE, Linköping University.
- Kahn Jr, C. E., Roberts, L. M., Shaffer, K. A., & Haddawy, P. (1997). Construction of a Bayesian network for mammographic diagnosis of breast cancer. *Computers in biology and medicine*, 27 (1), 19-29.
- Leod, P. M., & Verma, B. (2013). Variable Hidden Neuron Ensemble for Mass Classification in Digital Mammograms [Application Notes]. *Computational Intelligence Magazine, IEEE*, 8 (1), 68-76.
- Lucas, P. (2002). Restricted Bayesian network structure learning. *Proceedings of the First European Workshop on Probabilistic Graphical Models*, (págs. 117-126).
- Murphy, K. (octubre de 2007). *The bayes net toolbox for Matlab*. Obtenido de <https://code.google.com/p/bnt/>
- Nixon, M. S., & Aguado, A. S. (2012). *Feature Extraction & Image Processing for Computer Vision*. Academic Press.
- Pajares, M. G., & Cruz., G. J. (2008). *Visión por computador: imágenes digitales y aplicaciones* (2a. ed.). México: Alfaomega.
- Pant, C. (2011). *Atlas of Breast Imaging with Mammography, Ultrasound and MRI correlation*. Jaypee Brothers Medical Publishers.
- Patrocínio, A. C., Schiabel, H., & Roseli Ap F, R. (2004). Evaluation of Bayesian Network to classify clustered microcalcifications. *Proc. of SPIE Vol*, (pág. 1027).
- Rasband, W. (2004). ImageJ. *Image Processing and Analysis in Java*. National Institutes of Health, USA.
- Rojas, D. A., & Nandi, A. K. (2009). Toward breast cancer diagnosis based on automated segmentation of masses. *Pattern Recognition*, 42 (6), 1138-1148.
- Sahami, M. (1996). Learning limited dependence Bayesian classifiers. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, (págs. 335-338).
- Sampat, M., Markey, M., & Bovik, A. (2005). Computer-aided detection and diagnosis in mammography. En A. Bovik, *Handbook of Image and Video Processing* (págs. 1195-1217). Elsevier Academic Press.
- Samulski, M., Karssemeijer, N., Lucas, P., & Groot, P. (2007). Classification of mammographic masses using support vector machines and Bayesian networks. *Medical Imaging*, (págs. 65141J-65141J).
- Samulski, M. R. (2006). Classification of Breast Lesions in Digital Mammograms. *Master's thesis*. Netherlands: University Medical Center Nijmegen.
- Sharma, N., & Mukherjee, S. (2012). Layered approach for intrusion detection using naive Bayes classifier. *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, (págs. 639-644).
- Sickles, E. A. (1989). Breast masses: mammographic evaluation. *Radiology*, 173 (2), 297-303.
- Suckling, J. (1994). The mammographic image analysis society digital mammogram database. *International Congress Series* (1069), 375-378.
- Tang, X. (1998). Texture Information in Run-Length Matrices. *IEEE Transactions on Image Processing*, 7 (11), 1602-1609.

- Theodoridis, S., & Koutroumbas, K. (2008). *Pattern Recognition*. Academic Press.
- Tsui, P.-H., Liao, Y.-Y., Chang, C.-C., Kuo, W.-H., Chang, K.-J., & Yeh, C.-K. (2010). Classification of benign and malignant breast tumors by 2-d analysis based on contour description and scatterer characterization. *IEEE Transactions on Medical Imaging*, 29 (2), 513-522.
- van der Gaag, L. C., Renooij, S., Witteman, C. L., Aleman, B. M., & Taal, B. G. (2002). Probabilities for a probabilistic network: a case study in oesophageal cancer. *Artificial Intelligence in medicine*, 25 (2), 123-148.
- Velasquez, M. V. (1997). Lenguaje en imágenes del seno. *Revista Peruana de Radiología*, 3 (7).
- Velikova, M., de Carvalho Ferreira, N., & Lucas, P. (2007). Bayesian network decomposition for modeling breast cancer detection. En *In Artificial Intelligence in Medicine* (págs. 346-350). Springer Berlin Heidelberg.
- Velikova, M., Lucas, P. J., Samulski, M., & Karssemeijer, N. (2012). A probabilistic framework for image information fusion with an application to mammographic analysis. *Medical Image Analysis*, 16 (4), 865-875.
- Velikova, M., Samulski, M., Lucas, P. J., & Karssemeijer, N. (2009). Improved mammographic CAD performance using multi-view information: A Bayesian network framework. *Physics in Medicine and Biology*, 54 (5), 1131.
- Wang, X. H., Zheng, B., Good, W. F., King, J. L., & Chang, Y. H. (1999). Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network. *International Journal of Medical Informatics*, 54 (2), 115-126.
- Wu, X., Lucas, P., Kerr, S., & Dijkhuizen, R. (2001). Learning Bayesian-network topologies in realistic medical domains.