

# Método para la clasificación de lesiones patológicas en imágenes de mamografías utilizando características de textura.

Ing. Alejandro Díaz Sotolongo<sup>1</sup>, Ing. Danays B. Costa Alonso<sup>1</sup>

<sup>1</sup>Departamento de Computación, Facultad de Ingeniería, Universidad de Ciego de Ávila, Cuba  
email: alejandro@informatica.unica.cu, danays@ingenieria.unica.cu

## Resumen

El cáncer de mama constituye una de las principales causas de muerte entre mujeres, según estadísticas publicadas por la Organización Mundial de la Salud, en las últimas dos décadas se ha reportado un considerable aumento de la mortalidad a escala mundial. La mamografía ha demostrado ser el método más eficaz para la de detección precoz del cáncer de mama, ya que permite diagnosticar la enfermedad en su etapa inicial. Por esta razón los Sistemas de Diagnóstico Asistidos por Computadora (CAD) han desempeñado un papel clave en la detección precoz de dicha enfermedad, y de este modo ayudan a reducir la tasa de mortalidad entre las mujeres. En este trabajo se desarrolla un método para el apoyo al especialista médico en el diagnóstico de lesiones patológicas presentes en imágenes de mamografías. Este diseño, consistirá en un algoritmo elaborado a base de un lenguaje computacional, fundamentado en técnicas empleadas en la clasificación y detección de patrones en imágenes de mamografías.



Figura 1. Etapas que integran el método a desarrollar

## **Primera etapa: Preprocesamiento**

En esta etapa se inicia con el ingreso de la imagen original tomada de la base de datos del repositorio Digital Database for Screening Mammography (DDSM), el cual ha servido de referencia para la mayoría de los trabajos realizados con este propósito. Son imágenes de alta resolución en escala de grises. La etapa de pre-procesamiento, tiene como finalidad eliminar aquellos ruidos o elementos en la imagen que pueden distorsionar el proceso de identificación de microcalcificaciones, así como reducir el área de trabajo sólo a la región de la mama, para ello se aplican dos procedimientos: el filtro mediana y el corte automático. En este procedimiento se obtiene como resultado la imagen pre-procesada y una imagen espejo binarizada.

El pre-procesamiento, es fundamental para eliminar aspectos de distorsión de las imágenes, causadas principalmente por la colposcopia, en cuya técnica, el experto humano dirige la cámara manualmente y por consiguiente a este procedimiento, se generan algunas alteraciones como calidad de la resolución, brillo excesivo, pérdida de la profundidad, distorsión por movimientos del paciente, reactivos duplicados en la zona, ruido en la imagen debido a la presencia de instrumentos médicos utilizados en el procedimiento.

Con la aplicación de filtros, se procede a la disminución de ruidos ocasionados por diferentes factores: errores de impresión de la mamografía, rayones en la misma y algunas incrustaciones de partículas de polvo. La aplicación de filtros depende en gran medida del tipo de ruido que se desea eliminar. En esta parte del sistema se utilizó el técnica denominada filtros gaussianos.

### **Filtro Gaussiano**

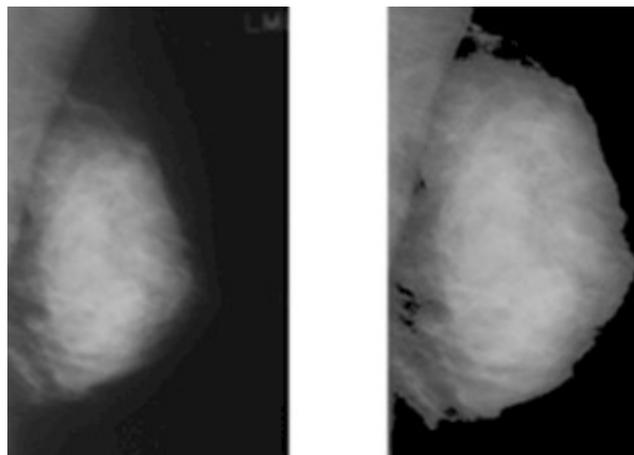
El filtro gaussiano, es considerado uno de los mejores métodos para la eliminación de ruidos de las imágenes. Para el sistema automatizado de apoyo al diagnóstico de lesiones en imágenes de mamografía, se utilizó la umbralización global, técnica clásica más sencilla en la que su funcionamiento se basa en el pico generado en el histograma por las regiones que contienen anomalías, es distinto del pico generado por el tejido sano. Esta técnica, sin embargo, no tiene gran precisión en la detección de ROIs, por lo que, la salida

generada por esta umbralización se utiliza como entrada de algún otro proceso posterior. Una técnica algo más precisa consiste en la umbralización local. En este caso, el valor de umbral se determina localmente para ventanas de menor tamaño en función de los niveles de gris de los píxeles contenidos en la ventana.

En el Preprocesado, se va a aislar la región de interés (mama) de las demás áreas, con esto se pretende reducir el área de procesamiento de la imagen, eliminando parte del fondo y etiqueta de la mamografía (en la mayoría de los casos presentes) dejando centrado únicamente la mama.

Después de analizar 100 mamografías se determinó que la mama presenta niveles de intensidad de gris comprendidos entre 130 y 250, por lo tanto estos son los valores de umbral que se van a utilizar en este trabajo.

Una vez determinado el intervalo de umbral se procede a reemplazar las intensidades de nivel de gris inferiores a 130 y superiores a 250 por el color negro total dejando la región de la mama (tejido fibroglandular, conductos galactóforos, ganglios linfáticos y músculo pectoral) sin cambios tal como se muestra en la Figura 2.



**Figura 2.** Imagen mostrando el efecto de reducir el área de procesamiento, (a) Imagen original, (b) imagen reducida.

En la figura anterior, puede apreciarse también como el filtro Gaussiano elimina mejor el ruido y enborrona menos los bordes.

## **Segunda etapa: Selección y extracción de características**

La extracción de características es, básicamente, un proceso de extracción de la información útil contenida en la señal a tratar. Los algoritmos de extracción de características, por tanto, crean nuevas características a partir de transformaciones y/o combinaciones de las características del conjunto original. La elección de un determinado método de extracción de características (o la combinación de varios de estos métodos) es una decisión que tendrá gran repercusión en las prestaciones del sistema.

En el caso de trabajar con imágenes mamográficas, y más concretamente con detección de lesiones, lo habitual será tratar de diferenciar únicamente entre dos clases: lesión y tejido normal o sano. Aunque tener que discriminar únicamente entre dos clases pueda parecer una reducción de la complejidad del problema, lo cierto es que éstas, especialmente la correspondiente a tejido sano, pueden tener apariencias muy dispares, por lo que lograr una buena representación de ambas puede llegar a ser una tarea de una complejidad considerable.

En un sistema de reconocimiento de formas, la etapa de extracción de características tendrá que aplicarse, utilizando exactamente los mismos criterios y algoritmos, en dos fases distintas. En primer lugar, dado un conjunto de muestras (imágenes) de entrenamiento, tras el preproceso de las mismas, el algoritmo de extracción de características obtendrá un conjunto de prototipos para cada una de las clases que se desean modelar. Posteriormente, ante una muestra de test (imagen) no vista anteriormente, se obtendrán las características del mismo modo a como se hizo con las muestras de entrenamiento. Aplicando entonces alguna regla de decisión se estimará la clase a la que, con mayor probabilidad, pertenece la nueva observación

### **Niveles de gris**

Para caracterizar los distintos píxeles de la imagen a analizar, lo más sencillo es utilizar directamente los valores de nivel de gris de dichos píxeles. La caracterización de una imagen mediante los niveles de gris de sus píxeles, aun siendo un método simple, ofrece buenos resultados cuando se combina con PCA.

## Filtro de Sobel

Una forma habitual de mejorar o transformar una imagen digital es mediante la aplicación de operadores espaciales o filtros. Para calcular el nuevo valor de cada píxel de la imagen, dichos operadores tienen en cuenta tanto el valor del propio píxel, como el de sus vecinos. Para ello, la imagen es transformada mediante convoluciones realizadas con unas matrices denominadas máscaras espaciales o kernels, de modo que cada píxel de la imagen resultante se obtiene mediante una combinación lineal de sus píxeles vecinos en la imagen original. Es precisamente la máscara la que contiene los coeficientes de la combinación lineal.

## Matrices de Coocurrencia

En el análisis estadístico de texturas, la extracción de características se realiza a partir de la distribución estadística con la que se observan combinaciones de determinadas intensidades en posiciones relativas de la imagen. Según el número de puntos (píxeles) considerados en cada combinación, se obtienen estadísticos de primer orden, de segundo orden o superiores.

Las matrices de coocurrencia, también llamadas Matrices de Dependencia de Niveles de Gris o Space Gray Level Dependence Matrices (SGLDM) son un método para obtener características de texturas de segundo orden. Esta aproximación ha sido utilizada en muy diversas aplicaciones.

Básicamente, una matriz de coocurrencia  $P$  es una matriz cuadrada, en la que el número de filas y columnas coincide con el número de niveles de gris ( $G$ ) en la imagen a analizar y donde cada elemento de la matriz  $P(i; j)$  contiene la frecuencia relativa con la que dos píxeles de la imagen  $I$ , con intensidades  $i$  y  $j$  respectivamente y separados por una distancia  $d$  y ángulo  $\theta$ , ocurren en una determinada vecindad.

Cuando se trabaja con matrices de coocurrencia, si se utiliza una cantidad elevada de niveles de gris en la imagen, pueden producirse varias situaciones no deseables: por un lado podría darse la situación paradójica de que las matrices de coocurrencia fueran de mayor tamaño que la propia imagen

original. Por otro lado, las matrices de coocurrencia pueden ser muy sensibles al tamaño de la ventana de vecindad cuando se trabaja con un valor elevado de niveles de gris. Finalmente, es recomendable que las matrices de coocurrencia tengan un grado alto de ocupación o, dicho de otro modo, que no sean excesivamente dispersas. Es evidente que cuanto mayor sea la cantidad de niveles de gris considerados, más dispersa será la matriz de coocurrencia.

#### Reducción de la dimensionalidad

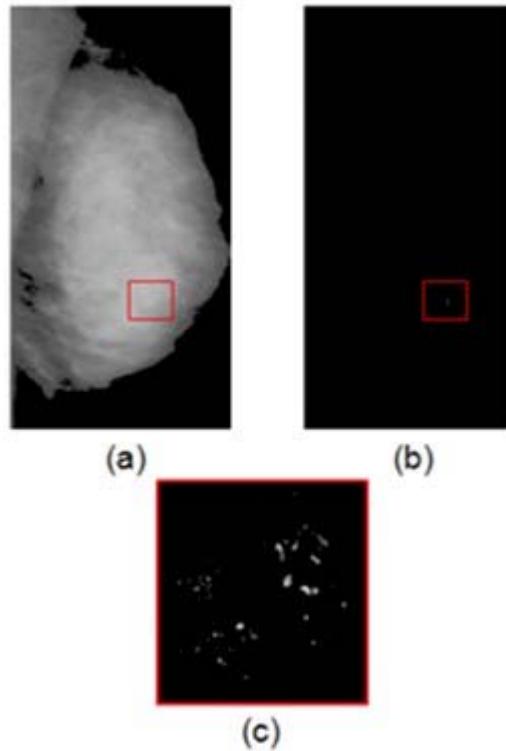
En algunas ocasiones puede ocurrir que el número de características obtenido para representar los objetos de nuestro problema sea excesivo de cara a una posterior etapa de clasificación. Básicamente existen dos motivos principales por los que podría ser interesante trabajar con dimensionalidades más bajas (menor número de características): reducir el coste computacional (tanto espacial como temporal) y mejorar la tasa de aciertos. Cuanto menor sea la dimensionalidad de los puntos a clasificar, más rápido será el clasificador y menos memoria requerirá.

#### **Extracción de características**

En esta etapa se busca resaltar las características de interés en la imagen de entrada, tales como: bordes, formas, regiones, color, tamaño, textura con la finalidad de obtener una imagen de salida filtrada (eliminación de ruidos). La finalidad es extraer las regiones potenciales (microcalcificaciones y/o masas), así como las características desde cada señal identificada en la etapa anterior, para luego clasificarlas como calcificación o no calcificación y/o masas.

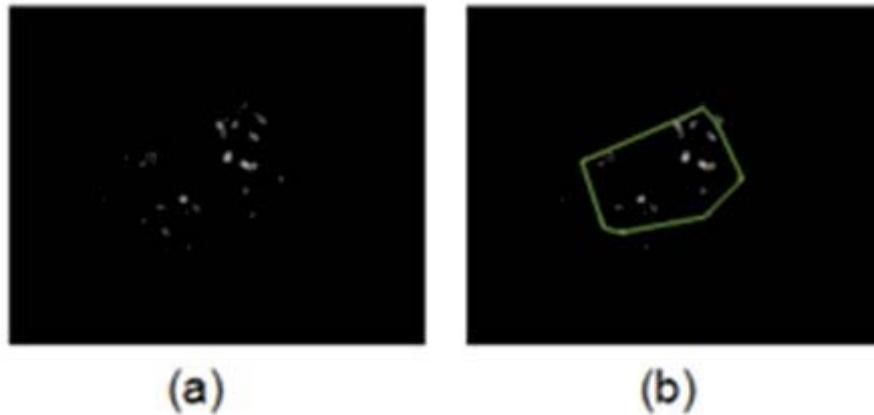
En el Análisis Morfológico, se trabaja sobre la región de la mama con el fin de detectar y aislar microcalcificaciones. Con este fin, en esta etapa se detectan regiones (grupos de tejidos) que comparten el mismo nivel de intensidad en escala de grises. Lo que se busca es detectar regiones sospechosas de contener microcalcificaciones, las cuales se determinó de forma experimental poseen tonos blancos con un nivel de intensidad mayor a 170, por lo tanto se realiza una umbralización en la región de interés con este valor. Una vez hecho esto se aíslan ya sea grupos de microcalcificaciones o cuerpos redondos de gran tamaño y varas tal como se ilustra en la Figura 3. Una vez que se tienen a

las regiones de interés aisladas se busca la ubicación de agrupaciones (microcalcificaciones) para esto se va aplicar la técnica de Convex Hull.



**Figura 3.** Ilustración del proceso para detectar microcalcificaciones, (a) muestra la imagen después del paso 1, (b) se aplica el umbral a la imagen previa y (c) se muestra un acercamiento a la región segmentada.

La técnica de Convex Hull, es una metodología también conocida como “Envoltura Convexa”, se encarga de encontrar los bordes de un conjunto de puntos, en este caso, una agrupación de microcalcificaciones. Este método consiste en encontrar los valores fronterizos de las posiciones de un conjunto de puntos dispersos y unirlos mediante una línea tomando en cuenta a los más cercanos. El objetivo en este paso es ubicar en la mamografía las regiones sospechosas de contener la lesión de microcalcificaciones. El resultado de este paso se observa en la Figura 4.



**Figura 4.** En esta imagen se ilustra un agrupamiento de microcalcificaciones, en (a) se determina la región de interés y en (b) la región es encerrada aplicando el método Convex Hull.

## **SEGMENTACIÓN**

La fase de segmentación trata de aislar las regiones sospechosas (ROIs) del resto de la imagen. La segmentación, es la parte del sistema en la que delimita la región de interés con la que se desea trabajar. Este proceso determinará la sensibilidad del sistema, esto es, su capacidad para detectar correctamente el tejido canceroso. La segmentación debería aislar el mayor número posible de anomalías, aunque entre las ROIs se encuentren regiones correspondientes a tejido sano (falsos positivos). Los falsos positivos deberán ser eliminados en una fase posterior. Atendiendo a su naturaleza, las aproximaciones a la segmentación pueden dividirse en tres grupos: técnicas clásicas, substracción de imágenes bilaterales y técnicas multiescala.

La técnica clásica más sencilla es la umbralización global, su funcionamiento se basa en que el pico generado en el histograma por las regiones que contienen anomalías, es distinto del pico generado por el tejido sano. Esta técnica sin embargo, no tiene gran precisión en la detección de ROIs, por lo que la salida generada por esta umbralización se utiliza normalmente como entrada de algún otro proceso posterior.

La umbralización local, es la técnica algo más precisa, en este caso el valor de umbral se determina localmente para ventanas de menor tamaño, en función de los niveles de gris de los píxeles contenidos en la ventana.

Una de las técnicas más populares para la segmentación de algunos tipos de tumores, caracterizados por presentar un nivel de brillo superior al tejido circundante, es la conocida como regiongrowing. La idea básica de este algoritmo, es determinar un conjunto de semillas en la imagen y a continuación hacer crecer de manera iterativa dichas semillas mediante la adición de píxeles vecinos, cuyas características sean similares a la de la semilla. Cuando la región deja de crecer, se compara el nivel de intensidad medio de la misma con la intensidad de las regiones de alrededor para determinar si se trata de una región sospechosa o no. Las claves para el correcto funcionamiento de este algoritmo radican en el criterio de selección de semillas, así como el criterio utilizado para determinar si un píxel debe o no añadirse a la región.

La técnica región clustering, similar a la anterior, las regiones se buscan directamente, sin necesidad de establecer semillas iniciales. El algoritmo k-medias, es una técnica de agrupamiento o clustering ampliamente utilizada.

En la substracción de imágenes, se pueden localizar diferencias entre dos imágenes médicas de un mismo paciente tomadas en distintos momentos. Para que esta aproximación genere resultados aceptables, es necesario realizar un alineamiento elástico entre ambas imágenes. Las diferencias encontradas entre una y otra imagen serán regiones sospechosas, aunque entre estas regiones puede haber un número importante de falsos positivos. Es necesario por tanto, hacer un análisis posterior basado en algún tipo de características que ayude a reducir los falsos positivos. En los casos de órganos simétricos, como por ejemplo la mama, es posible también buscar diferencias entre las imágenes de cada órgano, por ejemplo mamografías del pecho izquierdo y derecho tomadas en la misma sesión.

La técnica multiescala, puede facilitar la detección de los tumores, ya que en ocasiones el tamaño de los mismos puede variar considerablemente. Para ello,

el método de detección escogido se aplica a diferentes escalas y se combina de algún modo el nivel de sospecha detectado en cada una de las escalas. Por otro lado, mediante la transformación discreta de wavelets (DWT), también es posible hacer un análisis multiescala de la imagen.

Asimismo, la técnica ampliamente utilizada emparejamiento de patrones o template matching, es una aproximación que trata de segmentar las anomalías, buscando en la imagen patrones que guarden cierta similitud con un conjunto de prototipos obtenidos previamente de una muestra de entrenamiento. Cuando el tamaño de los patrones (anomalías) a segmentar se desconoce, es necesario disponer de un conjunto de prototipos que abarque todas las escalas posibles, o analizar la imagen de test a distintas escalas.

Otros algoritmos clásicos utilizados habitualmente para la segmentación de tumores son aquellos basados en la detección de bordes como los filtros de Sobel, que mejoran y transforman las imágenes médicas digitales.

### **Clasificación**

En esta etapa, se realiza la clasificación de imágenes sospechosas a través de un proceso de aprendizaje automatizado que simula el trabajo del experto humano, en el que se toma el vector que representa a la imagen para obtener la lesión más probable.

### **Métodos de selección de características**

Tal y como se ha explicado, usar un número excesivo de características puede degradar las prestaciones del clasificador debido al fenómeno de la maldición de la dimensionalidad. Por otro lado, es evidente que un número escaso de características puede producir una pérdida del poder de discriminación y, consecuentemente, reducir igualmente el rendimiento del clasificador.

Existe la necesidad, por tanto, de obtener un conjunto de características cuanto más reducido mejor, pero que al mismo tiempo recoja la información suficiente para caracterizar las distintas clases de nuestro problema. Dado un conjunto de características extraído de una muestra de test, los métodos de selección de

características tratan de detectar y rechazar aquellas características que presentan un bajo o nulo poder discriminatorio. Con ello, además de mejorar las prestaciones del clasificador, se consigue reducir costes computacionales.

### Matriz de covarianzas

La covarianza entre dos variables (o entre dos componentes de una misma variable) es un estimador de la dependencia lineal entre dichas variables. Dado que la covarianza se puede calcular entre cualesquiera dos dimensiones de un conjunto de datos, esta técnica se utiliza habitualmente para encontrar relaciones entre las distintas dimensiones de un conjunto de datos multidimensional. La matriz de covarianzas almacena todos los valores posibles de covarianza entre todas las dimensiones de un conjunto de datos.

### Vectores y valores propios

Los vectores propios sólo pueden encontrarse en matrices cuadradas, aunque no todas las matrices cuadradas tienen vectores propios (sí los tienen las matrices simétricas). Si una matriz de  $n \times n$  tiene vectores propios, entonces habrá  $n$  de ellos (sin contar, obviamente, distintos escalados de un mismo vector) que además serán ortogonales entre sí. Puesto que si escalamos un vector propio por cualquier valor real, el resultado sigue siendo un mismo vector propio, habitualmente se sigue el estándar de utilizar vectores propios cuyo módulo sea 1. En relación a la transformación PCA, la característica más interesante de los vectores propios es que, cuando se calculan sobre la matriz de covarianzas de un conjunto de datos, dichos vectores representan las direcciones del espacio que más cantidad de la varianza de los datos explica. Concretamente, cuanto mayor sea el valor propio asociado, el vector propio define una dirección del espacio con mayor varianza en los datos.

### Evaluación del modelo

La evaluación experimental, consiste en comprobar el modelo o diseño del sistema automatizado, así como el grado de cumplimiento de los objetivos general y específicos planteados al inicio de este trabajo de investigación. Al realizar pruebas con los datos obtenidos se encontraron aspectos que hay que tomar en cuenta, tales como los siguientes: Existe una gran variabilidad en los

niveles de intensidad presente en una mamografía digital sin embargo con los valores encontrados, se lograron disminuir casi en su totalidad la presencia de falsos positivos y aumentar considerablemente la detección de las microcalcificaciones.



**Figura 5.** Microcalcificaciones con intensidades similares al tejido mamario.

Debido a que se consideró solamente la intensidad como base para detectar la lesión es posible confundir en ocasiones la lesión con el tejido mamario, sin embargo hay que tomar en cuenta que este sistema es un apoyo al experto humano y es el quien va a realizar el diagnóstico final.

El modelo se va evaluar utilizando las siguientes medidas de desempeño:

- ✓ Exactitud = elementos correctamente clasificados / total de elementos
- ✓ Precisión = Verdaderos Positivos / (Verdaderos Positivos + Falsos Positivos)

Por ejemplo en la evaluación de 100 imágenes el sistema determinará cuantas son sospechosas.

Pruebas de funcionalidad del método

De acuerdo con Pressman, las pruebas de caja negra, llamadas también de comportamiento, se encuentran enfocadas a los requisitos funcionales de software, y permite al desarrollador centrarse en la coherencia de las entradas y salidas al sistema sin preocuparse de la estructura interna de la aplicación o sistema automatizado.

Las pruebas realizadas con el diseño, se enfocaron en primer término a la familiaridad con el usuario (facilidad de uso), manejo del sistema, acceso a la información (imágenes), almacenamiento, rapidéz y capacidad de respuesta. Cada uno de los aspectos examinados fue realizado con diferentes imágenes mamográficas, con la finalidad de determinar que los resultados obtenidos fueran consistentes bajo cualquier situación presentada.

#### Análisis de resultados mediante curvas ROC y FROC

Las curvas ROC (Receiver Operating Characteristic) aportan notables ventajas frente al error de clasificación, por lo que su uso está ampliamente extendido en problemas de clasificación binarios. Un caso particular son las aplicaciones médicas.

En dichas aplicaciones es habitual que, a partir del resultado de cierta prueba, se deba emitir un diagnóstico que será, de forma genérica, paciente sano o paciente enfermo.

#### Conceptos previos

- Verdadero Positivo (V P): Paciente enfermo diagnosticado como enfermo.
- Falso Positivo (FP): Paciente sano diagnosticado como enfermo.
- Verdadero Negativo (V N): Paciente sano diagnosticado como sano.
- Falso Negativo (FN): Paciente enfermo diagnosticado como sano.

Dado un clasificador (o un test clínico) y una muestra (paciente), definimos la siguiente notación:

T+: resultado del test positivo

T-: resultado del test negativo

E+: paciente enfermo. Nótese que el número total de pacientes enfermos será  $V P + FN$

E-: paciente no enfermo (sano). Nótese que el número total de pacientes sanos será  $V N + FP$

o demos definir:

Fracción de Verdaderos Positivos (FVP):

$$FVP = P (T+ E+) = VP / (VP+FN)$$

Fracción de Verdaderos Negativos (FVN):

$$FVN = P (T- E-) = VN / (VN+FP)$$

Fracción de Falsos Positivos (FFP):

$$FFP = P (T+ E-) = FP / (VN+FP)$$

Fracción de Falsos Negativos (FFN):

$$FFN = P (T- E+) = FN / (VP+FN)$$

Es importante observar que la fracción de falsos negativos es 1 menos la fracción de verdaderos positivos:

$$FFN = 1 - FVP \text{ y de modo similar } FFP = 1 - FVN$$

Valor Predictivo Positivo (VPP): es la probabilidad de que el paciente padezca la enfermedad cuando el test ha dado positivo.

$$VPP = P (E+ T+) = VP / (VP+FP)$$

Valor Predictivo Negativo (VPN): es la probabilidad de que el paciente no padezca la enfermedad cuando el test ha dado negativo.

$$VPN = P (E- T-) = VN / (VN+FN)$$

Los valores de VPP y VPN no deberían usarse en muestras en las que la probabilidad a priori de padecer la enfermedad haya sido controlada de forma artificial. Dichos valores no tendrían validez si para realizar el estudio se hubiesen utilizado, por ejemplo, pacientes sanos y enfermos en proporción uno a uno.

Habitualmente, en el campo de la medicina, se utilizan los términos:

Sensibilidad para referirse a la Fracción de Verdaderos Positivos (FVP), esto es, la probabilidad de clasificar correctamente a un individuo enfermo. La sensibilidad, por tanto, mide la capacidad del test para detectar la enfermedad.

Especificidad para referirse a la Fracción de Verdaderos Negativos (FVN), esto es, la probabilidad de clasificar correctamente a un individuo sano. La especificidad, por tanto, mide la capacidad del test para detectar los pacientes sanos

Un test muy sensible será aquel que detecta con una alta probabilidad a todos los pacientes enfermos, aunque sea a costa de cometer más errores de falsos positivos.

Por contra, un test será más específico cuantos más pacientes sanos detecte, aunque sea a costa de cometer más errores de falsos negativos.

Una aproximación para decidir el umbral, consiste en evaluar la utilidad (U) del test en función de dónde se sitúe dicho umbral. Para ello se define:

BVP: beneficio de verdadero positivo. Si la enfermedad tuviese un tratamiento eficaz y seguro, el beneficio de un verdadero positivo sería 1.

CFN: coste de falso negativo.

BVN: beneficio de verdadero negativo.

CFP: coste de falso positivo.

### **Referencias Bibliográficas**

1. Avalo, E. (2003). Información para el público en general. [www.samas.org.ar](http://www.samas.org.ar). Página vigente al 17/08/2003.
2. Bueno Gloria, 2008. "Detección automática de lesiones en la mama", Universidad de Castilla - La Mancha. Ciudad Real, 10 de septiembre de 2008.
3. Brandan y Villaseñor, Cancerología 1 (2006) Detección del Cáncer de Mama: Estado de la Mamografía en México. Instituto Nacional de Cancerología.
4. Cancer Mondial, International Agency for Research on Cancer, IARC, <http://www-dep.iarc.fr>
5. Calderón-Garcidueñas AL. et al., Risk factors of breast cancer in Mexican women. *Salud Publica Mex.* 2000; 42:26-33.
6. Castro C.A, Pérez T.S., González Barcena J. M., and Santiago J.R., 1999. Interferón y cáncer de mama avanzado. *Revista Cubana de Oncología*, 15(2):8994.

8. Crouse M. S., Nowak R. D. and Baraniuk R.G., 2008. "Wavelet-Based Statistical Signal Processing Using Hidden Markov Models", IEEE Trans. Signal Proc. 46, PP. 886-902, april 2008.
9. D. Laya, E. Urra, J. Rodríguez, P. Linares, F. Saldivia y A. Reigosa. Diseño de Software para la Estimación de la Relación de Volumen Tumor/Mama a Partir de Mamografía. En Salus, revista de la Facultad de Ciencias de la Salud, Vol. 12 Núm. 3. Universidad de Caraboo, Diciembre de 2008.
10. DDSM: Digital Database for Screening Mammography <http://marathon.csee.usf.edu/Mammography/Database.html>
11. Egan R, 1988. Breast Imaging: Diagnosis and Morphology of Breast Diseases. Philadelphia. W.B. Saunders Company.
12. Ferrero G., Britos P., García-Martínez R.. Detection of Breast Lesions in Medical Digital Imaging Using Neural Networks. In IFIP International Federation for Information Processing, Volume 218, Professional Practice in Artificial Intelligence, eds. J. Debenham, (Boston: Springer), pp. 1-10. 2006.
13. González Rafael y Woods Richard, 1992. Digital Image Processing. Ed. Addison-Wesley Publishing Company, USA.
14. Hayken S, 1994. Neural networks: A comprehensive foundation. Nueva York: Macmillan.
15. ICPR 2006. 18th International Conference on Volume 4, Issue, 20-24 Aug. 2006 Page(s): 958 – 958 Digital Object Identifier 10.1109/ICPR.2006.749.
16. I. Anttinen, M. Pamilo, M. Soiva, and M. Roiha. Double reading of mammography screening films: one radiologist or two ClinRadiol, (48):414{421, 1993.
17. INEGI, Anuario Estadístico de los Estados Unidos Mexicanos, eds. 1999-2005.
18. INEGI, 2008. Instituto Nacional de Estadística, Geografía e Informática. Proyecciones de la población con base en el Censo 2005. México.
19. Lauria A., Palmiero R., Forni G., Cerello P., Golosio B., Fauci F., Magro R., Raso G., Tangaro S., Indovina P. The CALMA System: an artificial neural network for detecting masses and microcalcifications in digitalized mammograms.
20. Levi F, Bosetti C, Lucchini F, Negri E, and La Vecchia C, Monitoring the decrease in breast cancer mortality in Europe. Eur. J. Cancer Prevention 2005, 14(6): 497-502.
21. López Rios O, Lazcano-Ponce E, La epidemia de cáncer de mama en México. ¿Consecuencia de la transición demográfica? Salud Pública Mex. 1997;39(4): 259-265.
22. McPherson K, Steel CM and Dixon JM, Breast cancer – epidemiology, risk factors and genetics. BMJ 2000; 321:624-628.
23. [3] Narváez Espinoza Fabián Rodrigo (Ing), 2010. "Sistema de anotación para el apoyo en el seguimiento y diagnóstico de cáncer de seno", Bogotá D.C., 26 de marzo de 2010.

24. NOM-041-SSA2-2002, Para la prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama. Diario Oficial, 17 septiembre 2003.
25. Oporto A, 2004. Detección Automática de Agrupamientos de Microcalcificaciones en Mamografías Digitalizadas. TESIS. Maestría en Ciencias en Sistemas Inteligentes. Instituto Tecnológico y de Estudios Superiores de Monterrey. Por Samuel Alonso Oporto Díaz. Diciembre 2004.
26. Oporto Díaz Samuel A., Detección de Microcalcificaciones en Imágenes de Mamografías usando Diferencia de Filtros Gaussianos Optimizados. Universidad Nacional de Ingeniería.
27. Oporto Díaz Samuel Alonso, 2007. "Detección automática de agrupamientos de microcalcificaciones en mamografías digitalizadas". Diciembre 2007.
28. Oporto S. Comparación de Técnicas para el Diagnóstico Automático de Cáncer de Mama. Filtros Gaussianos Optimizados y Wavelets. Cursos y recursos de Inteligencia Artificial, Propuesta de Tema de Investigación. [www.wiphala.net/research/proposal/cancer\\_mama\\_diagnostico.html](http://www.wiphala.net/research/proposal/cancer_mama_diagnostico.html). Abril 2008.
29. OMS, 2008. Organización Mundial de la Salud. Cáncer.
30. Peto R, Borrrham J, Clarke M, Davies C, and Beral V, UK and USA breas cancer deaths down 25% in year 2000 at ages 20-69 years. Lancet 2000, 355:1822.
31. Rangayyan R. M, El-Faramawy N. M., Leo Desautels J. E. and Alim O. A., 1997. Measures of acurate and shape for classification of breast tumors. IEEE Transactions on Medical Imaging, 16(6):799-810.
32. R. Campanini, A. Bazzani, A. Bevilacqua, D. Bollini, D.N. Dongiovanni, E. Iampieri, N.Lanconelli, A. Riccardi, M. Rolli, and R. Tazzoli. A novel approach to mass detection in digital mammography based on support vector machines (svm). In Proc. of IWDM2002, pages 399{401, Bremen, Germany, June 22
33. Ristori Bogajo Esther, 2007. "Aplicaciones de la segmentación interactiva y el análisis morfométrico de nódulos en mamografía al diagnóstico asistido por ordenador". Málaga, España. Septiembre de 2007.
34. Rodríguez-Cuevas S, Macías CG, Franceschi D, and Labastida S, Breast carcinoma presents a decade earlier in Mexican Women than in Women in the United Status or European countries. Cáncer 2001; 91(4): 863-868.
35. Rodríguez J. M. Algoritmo de Segmentación de la Imagen para la Detección de Cáncer de Mama. Reporte técnico. Escuela Universitaria Politécnica de Mataró, España, 2004.
36. Sandoval N. Fernando. (2002). Enfoques epistemológicos en la elaboración de diseños de investigación, pp. 90-91. Chilpancingo, Gro. México.
37. Sucar, L. E., Martínez-Arroyo M., Acosta G., Cruz Nicandro. Bayesian Model Combination and Its Application to Cervical Cancer Detection. Book Series Lecture Notes in Computer Science. Publisher Springer Berlin / Heidelberg, ISSN 0302-9743, Subject Computer Science, Volume 4140/2006.

38. Sucar L.E., Montero J.A., 2006. A decision-theoretic assistant based on gesture recognition, 5th Mexican International Conference on Artificial Intelligence, IEEE Computer Society, pp. 78-87.
39. Sucar L.E., Martínez M., 1998. "Interactive Structural Learning of Bayesian Networks", Expert Systems and Applications, Vol. 15, No. 3/4, pp., Pergamon Press.
40. Sucar L.E., Martínez M., 2006. "Learning an optimal naive bayes classifier", International Conference on Pattern Recognition (ICPR), Agosto 2006, Vol. 3, pp. 1236-1239.
41. Sucar L.E., Martínez M, "Aprendizaje de clasificadores bayesianos dinámicos" Open Discusión Track Proceedings. IBERAMIA 2002, VIII Iberoamerican Conference on Artificial Intelligence. Sevilla España. Nov. 12-15 2002 Ed. F. Garijo, J.C. Riquelme, M. Toro.
42. Sucar L.E., Martínez M. Learning an Optimal Naive Bayes Classifier. 36° Congreso de Investigación y Desarrollo del Tecnológico de Monterrey, congreso CIDTEC" 2006. Enero de 2006.
43. Velásquez, M. G. Business Ethics Concepts and Cases. (1992) (3rd. ed). New Jersey: Prentice Hall.
44. Zhang W., Doi K., Giger M. L., Nishikawa R. M., and Schmidt R. A., An improved Shift-Invariant Artificial Neural Network for Computerized Detection of Clustered Microcalcifications in Digital Mammograms Med. Phys., vol. 23, pp. 595-601, 1996.