

Análisis de imágenes de mamografía para la detección de cáncer de mama

1. Introducción

El cáncer de mama es una enfermedad en la que se forman células malignas en los tejidos de la mama (Instituto Nacional del Cáncer, 2012). Es considerado actualmente el cáncer más frecuente y el de mayor mortalidad entre las mujeres del mundo. En México, de acuerdo al INEGI, a partir del 2006 este tipo de cáncer ocupa el primer lugar de causa de muerte por cáncer entre las mujeres (Knaul, Arreola, Lozano, & Gómez, 2011). Hasta el momento, para este tipo de cáncer, se desconocen las causas y/o los medios de prevención, el único método que ha demostrado ser efectivo en la detección oportuna es la mamografía.

La mamografía es un examen radiográfico de la glándula mamaria que puede ser dividido en dos categorías, escrutinio y diagnóstico. La mamografía de escrutinio, es un estudio que se realiza en mujeres sanas para detectar tempranamente signos de cáncer de mama. La mamografía de diagnóstico se usa para buscar cáncer después de haberse encontrado un abultamiento u otro signo o síntoma de dicho cáncer. El examen de diagnóstico arroja información acerca de la localización, el número y las características de una lesión palpable. Un examen mamográfico típico consiste en la obtención de dos proyecciones radiográficas 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 loca-

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.

lización 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).

Algunos de los signos de la presencia de anomalías, que un radiólogo busca al analizar una mamografía son calcificaciones, nódulos o masas, asimetría en la densidad, y/o distorsión de la arquitectura. Dependiendo de las características de las anomalías encontradas, el radiólogo formula su interpretación y una serie de recomendaciones.

La interpretación de una mamografía, usualmente, se elabora utilizando el sistema BI-RADS. El BI-RADS (*Breast Imaging Reporting and Data System*) es un sistema de datos y reportes, elaborado por el Colegio

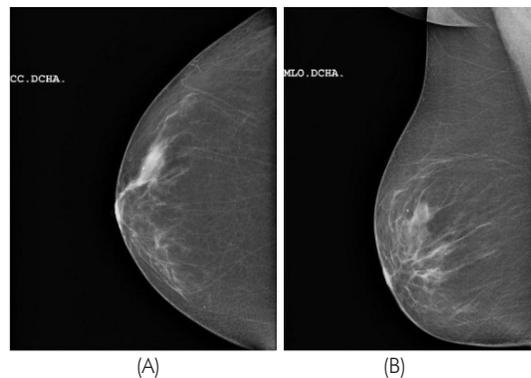


FIG. 1 MAMOGRAFÍA DE LA MAMA DERECHA: (A) PROYECCIÓN CRÁNEO-CAUDAL QUE PERMITE OBSERVAR CON MAYOR PRECISIÓN EL TEJIDO MEDIAL Y (B) PROYECCIÓN MEDIO-LATERAL-OBLICUA EN DONDE SE PUEDE APRECIAR LOS TEJIDOS DESDE LA PARTE ALTA DE LA AXILA HASTA LA ZONA POR DEBAJO DEL PLIEGUE INFRAMAMARIO.

Estadounidense de Radiología, que ayuda al radiólogo a elaborar un reporte estandarizado y reduce la posible confusión en la interpretación de la imagen mamográfica. En México, la Norma Oficial NOM-041-SSA2-2002 (Secretaría de Salud, 2002) establece que los resultados de la mamografía deben reportarse de acuerdo con la clasificación del BI-RADS.

Existe una gran complejidad en la interpretación de mamografías, debido principalmente a la variabilidad en la apariencia de las anomalías, a las características del tejido mamario, y a la calidad de las imágenes de mamografía. Muchas veces, las anomalías 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. La variabilidad de las características de las lesiones se refiere principalmente, a que un grupo de características que puede ayudar a identificar malignidad, también puede ayudar a determinar que la lesión es benigna (Bozek, Mustra, Delac, & Grgic, 2009).

Por otra parte, la interpretación humana de una mamografía es subjetiva y cualitativa, además de que requiere de mucho tiempo del radiólogo. La determinación de la presencia de lesiones se realiza a través de un examen visual de las imágenes de mamografía. Por lo que la interpretación correcta de estos estudios, depende en gran medida de la experiencia del radiólogo, y aún entre radiólogos experimentados puede existir una alta variabilidad, entre el 65-75% (Cheng, Shi, Min, Hu, Cai, & Du, 2006). La interpretación incorrecta de estudios trae como consecuencia que algunas lesiones (posiblemente malignas) pasen desapercibidas, o en otros casos, la recomendación innecesaria de biopsias.

A fin de mejorar el desempeño en el análisis de mamografías, se han desarrollado sistemas computacionales que apoyan en la detección (*CAD-Computer Aided Detection*) o en el diagnóstico de lesiones (*CADx-Computer-Aided Diagnosis*). Estos sistemas, cuando son usados en combinación con la opinión de radiólogos especializados, han demostrado que pueden ayudar a incrementar el desempeño en el análisis de mamografías (Bozek, Mustra, Delac, & Grgic, 2009; Cheng, Shi, Min, Hu, Cai, & Du, 2006).

En la Universidad Tecnológica de la Mixteca, el Cuerpo Académico de Reconocimiento de Patrones

está desarrollando un proyecto que tiene como objetivo la implementación de un sistema CADx que apoye a hospitales de la región en la detección y diagnóstico de masas en mamografías. En este trabajo, se presenta un breve resumen sobre la investigación realizada hasta el momento. En las primeras secciones se describe el tipo de lesiones que pueden encontrarse en una mamografía, enfatizando en las características de una de las más importantes, las masas. Posteriormente, se presentan las tareas que deben realizar los sistemas CAD, así como, una concisa descripción de los tipos de algoritmos utilizados para la detección de masas. Finalmente, en la última sección se exponen los alcances y los objetivos de la presente investigación.

2. Hallazgos mamográficos

Existe un gran número de lesiones o anomalías que pueden estar presentes en la mama. Algunos de estas anomalías, que pueden ser indicadores de cáncer de mama y que un radiólogo busca al analizar una mamografía, son asimetría, distorsión de la arquitectura, calcificaciones y masas (Bozek, Mustra, Delac, & Grgic, 2009).

Al analizar una mamografía, una de las primeras tareas que realiza un radiólogo, es comparar las mamografías de ambas mamas para detectar si existe una **asimetría** entre éstas. Las **asimetrías** son áreas de tejido de densidad fibroglandular que son más extensas en una mama que en la otra (Sickles, 2007). La asimetría puede ser el indicador de la presencia de una masa o de una variación en la distribución del tejido normal (Malagelada, 2007). Un ejemplo de asimetría se muestra en la Fig. 2(a).

La **distorsión de la arquitectura** es definida como una distorsión focal de la arquitectura normal de la mama sin la presencia de masas visibles. Se puede visualizar como espículas que irradian desde un punto común, formando un patrón similar al de una estrella (De Paredes, 2007). La distorsión de la arquitectura puede ser un indicador de malignidad cuando es complementado con la identificación de masas, calcificaciones o asimetrías (Bozek, Mustra, Delac, & Grgic, 2009). Un ejemplo de distorsión de la arquitectura se muestra en la Fig. 2(b).

Las **calcificaciones** son acumulaciones cristalinas de calcio en el tejido mamario. Regularmente tienen

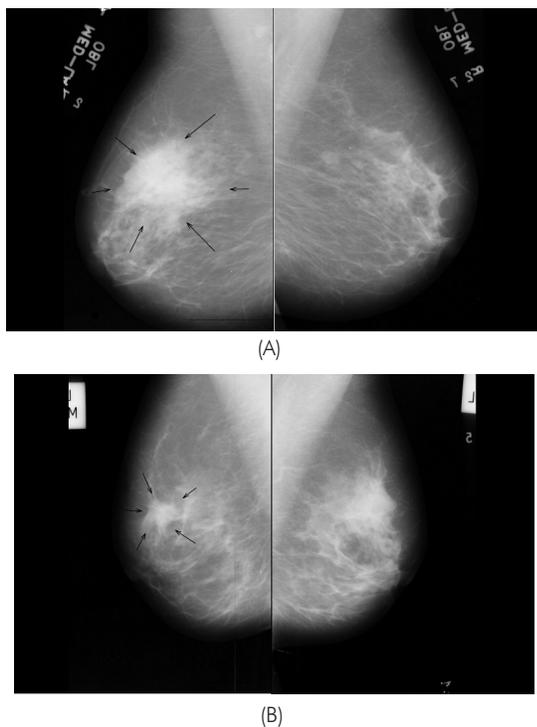


FIG. 2 EJEMPLOS DE (A) ASIMETRÍA QUE SE PUEDE APRECIAR EN LA PROYECCIÓN MLO IZQUIERDA Y (B) DISTORSIÓN DE LA ARQUITECTURA QUE SE VISUALIZA COMO UN ÁREA ESPICULADA EN LA PROYECCIÓN MLO IZQUIERDA .

un tamaño de micras, su patrón de agrupación y morfología son indicadores de malignidad. En la Fig. 3(a) se presenta un ejemplo de una agrupación de microcalcificaciones. Por regla general, las calcificaciones más grandes con forma oval o redonda y tamaño uniforme (entre calcificaciones), están asociadas a procesos benignos. Mientras que las calcificaciones asociadas a procesos malignos son más pequeñas, irregulares, polimórficas y con variaciones en su tamaño (Bozek, Mustra, Delac, & Grgic, 2009; Shinde, 2003).

Una **masa** es una estructura tridimensional que tiene, total o parcialmente, bordes externos convexos y es evidente en al menos dos proyecciones de la mama (Pant, 2011). Cuando una masa es sólo vista en una de las proyecciones se le denomina asimetría hasta que un análisis confirme su tridimensionalidad. Regularmente, los radiólogos caracterizan una masa por su densidad, sus márgenes y su forma. Un ejemplo de una masa se muestra en la Fig. 3(b).

La **densidad** de una masa se utiliza para definir la atenuación de los rayos X por la lesión, considerando la atenuación esperada para un volumen de tejido glandular similar. Esta característica es importante, ya que la mayoría de los cánceres de mama que for-

man una masa visible son de igual o mayor densidad que un volumen similar de tejido glandular normal (Shinde, 2003). Es raro (aunque no imposible) que un cáncer de mama sea de más baja densidad. El cáncer de mama nunca contiene grasa (radiolúcido) aunque si puede atrapar grasa. La densidad de una masa puede ser alta, isodensa (igual densidad), baja (baja atenuación, pero no contiene grasa), radiolúcido (contiene grasa) (Velasquez, 1997).

Los **márgenes** se refieren a los bordes de una masa, los cuales deben ser cuidadosamente caracterizados debido a que es uno de los criterios más importantes en la determinación de la malignidad o benignidad de una masa. Los radiólogos clasifican los márgenes de una masa en cinco clases (Malagelada, 2007):

- **Circunscrita:** Una masa con márgenes bien definidos y demarcados con una abrupta transición entre la lesión y el tejido que la rodea.
- **Microlóbulados:** Cuando los márgenes son ondulados con ciclos cortos produciendo pequeñas ondulaciones.
- **Mal definidos:** En este caso los márgenes están pobremente definidos y son dispersos.
- **Espiculados:** La lesión se caracteriza por líneas que se irradian desde los márgenes de la masa.

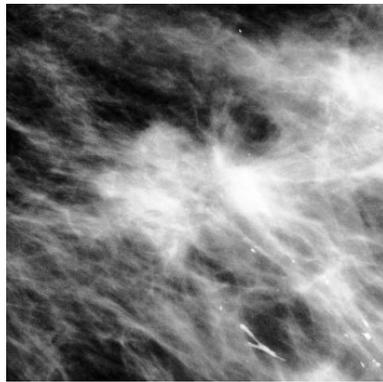
Las **formas** que puede tener una masa son (Velasquez, 1997):

- **Redonda:** Una masa que es esférica, circular o globular.
- **Oval:** Una masa que es elíptica o en forma de huevo.
- **Lobular:** Una masa que tiene contornos ondulados.
- **Irregular:** La forma de la lesión no puede caracterizarse por ninguna de las terminologías descritas anteriormente.

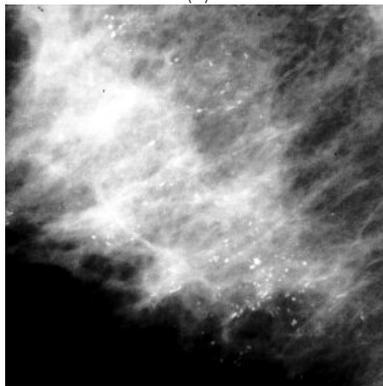
Los diferentes tipos de márgenes y formas de las masas se presentan en la Fig. 4.

Masas redondas con márgenes suaves y circunscritos usualmente indican cambios benignos. Por otra parte, las masas malignas, frecuentemente, tienen márgenes espiculados y borrosos (ver Fig. 3(b)). Sin embargo, existen casos atípicos de masas benignas con márgenes macrolóbulados o espiculados; así como también, masas malignas con márgenes mi-

crolobulados y circunscritos (Bozek, Mustra, Delac, & Grgic, 2009).



(A)



(B)

FIG. 3 EJEMPLOS DE (A) UN GRUPO DE MICROCALCIFICACIONES Y (B) DE UNA MASA ESPICULADA².

3. Sistemas de detección asistida por computadores

Un sistema de detección asistida por computadora CAD, es utilizado para ayudar a los radiólogos en la detección de anomalías en mamografías. Estas herramientas se utilizan como un segundo lector, el radiólogo primero tiene que formular su interpretación de la mamografía antes de usar los sistemas CAD. Básicamente, la tarea de estos sistemas es sólo indicar las áreas a donde posiblemente puede existir una lesión, usando marcas especiales. Por lo que, el trabajo de delimitar el área de la lesión para que pueda ser caracterizada y diagnosticada, queda a cargo del médico radiólogo. Las tareas que deben realizar estos sistemas sobre las imágenes de mamografía son (Bozek, Mustra, Delac, & Grgic, 2009; Cheng, Shi, Min, Hu, Cai, & Du, 2006):

1. Pre-procesamiento: Etapa en la cual se debe reducir el ruido y mejorar la calidad de la
2. Imágenes obtenidas de *Mini-MIAS database* (Suckling, 1994).

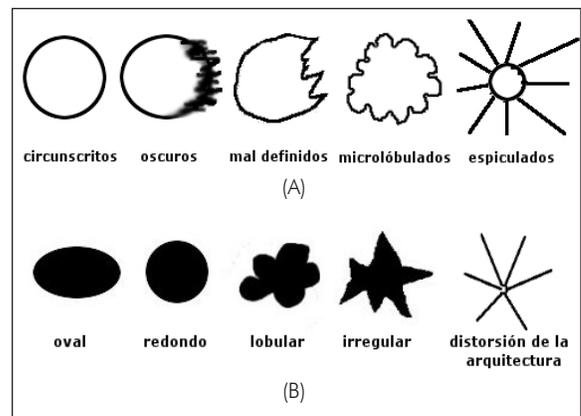


FIG. 4 POSIBLES (A) MÁRGENES Y (B) FORMAS DE UNA MASA.

imagen. En algunos sistemas, en esta etapa es a donde se elimina el fondo de la imagen y el músculo pectoral.

2. Segmentación: Etapa en la cual se localizan y se busca aislar regiones sospechosas de pertenecer a una anomalía.
3. Extracción de características: Etapa en la que se obtienen las características de las regiones sospechosas obtenidas en la etapa anterior.
4. Clasificación: Etapa en la que se clasifican las regiones sospechosas utilizando sus características. Esta etapa la aplican algunos sistemas CAD a fin de validar que la región detectada corresponde a una lesión y no a tejido mamario normal.

No existe un consenso en la forma en la cual se deba reportar el desempeño de los algoritmos de detección. Algunos autores reportan el desempeño, simplemente, en función del número de verdaderos positivos (TP) y de falsos positivos (FP). Un TP es una marca hecha por el sistema CAD que corresponde a la localización de una lesión. Un FP es una marca hecha por el sistema CAD que no corresponde a la localización de una lesión. Algunas otras métricas usadas para reportar el desempeño son matrices de confusión (1), exactitud (2), precisión (3), sensibilidad (4), y el índice de FP (5) (Bozek, Mustra, Delac, & Grgic, 2009). Una matriz de confusión consiste de verdaderos negativos (TN), verdaderos positivos (TP), falsos positivos (FP) y falsos negativos (FN).

Existen diversos sistemas comerciales tipo CAD que han sido aprobados por la FDA (Food and Drug Administration, USA) para auxiliar en la detección de anomalías en mamografías. Algunos de estos sistemas son (Sampat, Markey, & Bovik, 2005):

$$(1) C = \begin{pmatrix} TN & FP \\ FN & TP \end{pmatrix}$$

$$(2) \text{ exactitud} = \frac{TP+TN}{TN+FP+FN+TP}$$

$$(3) \text{ precision} = \frac{TP}{FP+TP}$$

$$(4) \text{ sensibilidad} = \frac{TP}{FN+TP}$$

$$(5) FPI = \frac{FP}{TN+FP}$$

- *R2 Technology's Image Checker*. Fue el primer sistema aprobado por la FDA y es usado para ayudar a detectar masas y calcificaciones.
- *MammoReader*. Desarrollado por Intelligent System's Software, Inc, puede detectar masas, calcificaciones, distorsión de la arquitectura y asimetría.
- *Second Look*. Desarrollado por CADx Medical Systems, puede detectar masas y calcificaciones.

Por otra parte, los sistemas CADx (Computer-Aided Diagnosis) son herramientas empleadas para auxiliar en el diagnóstico de lesiones. Las tareas de los sistemas CADx son mucho más complejas que la de los CAD's. Estos sistemas, además de localizar lesiones, deben de caracterizarlas y determinar su grado de malignidad o benignidad. Debido a esta complejidad en sus funciones, aún no existen sistemas CADx comerciales, sólo existe el reporte de algunos desarrollados con fines de investigación (Bozek, Mustra, Delac, & Grgic, 2009; Sampat, Markey, & Bovik, 2005).

4. Detección de masas

La gran mayoría de los algoritmos de detección de masas consideran dos etapas: detección de regiones sospechosas y clasificación de las regiones sospechosas como masa o tejido normal. Estas etapas se muestran en la Fig. 5. Los algoritmos para realizar la

primera etapa son generalmente de dos tipos, basados en píxeles y basados en regiones (Bozek, Mustra, Delac, & Grgic, 2009; Sampat, Markey, & Bovik, 2005).

Los métodos basados en píxeles, realizan el siguiente proceso para realizar la detección. Primero, extraen las características de cada píxel a partir de su vecindad local. Después, estas características son utilizadas en el proceso de clasificación del píxel (como normal o sospechoso). Finalmente, los píxeles sospechosos son agrupados en regiones, al recolectar píxeles conectados. Una de las ventajas de este tipo de métodos, es que se cuenta con un gran número de muestras para entrenar al clasificador. Sin embargo, estas técnicas no consideran la información espacial de los píxeles, el cual es un factor importante para discriminar masas de tejido normal.

Por otra parte, los métodos basados en regiones, primero extraen las regiones de interés al aplicar técnicas de segmentación o de filtrado de imágenes. Posteriormente, extraen las características de cada región y cada región es clasificada como sospechosa o normal. Una de las ventajas de estos métodos es que consideran la información espacial. También, las características obtenidas están directamente correlacionadas con información diagnóstica, como lo son la forma y los márgenes de las regiones extraídas. Sin embargo, su desventaja principal es que no se cuenta con un suficiente número de muestras para entrenar al clasificador.

La etapa de clasificación, en los algoritmos de detección, se usa para reducir el número de falsos positivos que pudieran haberse generado en la primera etapa. Usualmente, la clasificación se realiza a partir de las características de textura y utilizando una Red Neuronal. Otros métodos para realizar esta discriminación comparan las regiones sospechosas (obtenidas en la primera etapa) con imágenes representativas de masas (Cheng, Shi, Min, Hu, Cai, & Du, 2006).

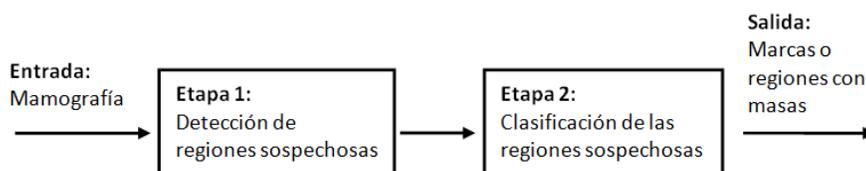


FIG. 5 DIAGRAMA DE FLUJO QUE MUESTRA LAS ETAPAS INVOLUCRADAS EN LA DETECCIÓN DE MASAS.

5. Alcance y objetivos de la investigación

En la Universidad Tecnológica de la Mixteca, el Cuerpo Académico de Reconocimiento de Patrones está desarrollando un proyecto de investigación relacionado con la construcción de un sistema CADx. Este proyecto pretende aportar un sistema que ayude al médico radiólogo en la detección temprana y en el diagnóstico de masas en mamografías digitales; y que al mismo tiempo, pueda servir como un auxiliar para el entrenamiento de médicos radiólogos que deseen especializarse en radiografía mamográfica.

Como se mencionó anteriormente, el cáncer de mama ocupa el primer lugar por neoplasia maligna entre las mujeres mexicanas. Dada esta problemática, y las relacionadas con la interpretación de mamografías, existe una gran necesidad de desarrollar sistemas computacionales que puedan ayudar a los radiólogos a realizar análisis más precisos y objetivos de las mamografías. Y considerando que aún no se tiene el reporte del desarrollo de un sistema CADx de forma comercial y de que los hospitales de la Región no cuentan con suficientes radiólogos especializados, la implementación de este tipo de herramientas es primordial.

A fin de alcanzar los objetivos del proyecto, su desarrollo está dividido en dos fases: detección y diagnóstico de masas. En la primera fase titulada, “**Detección, segmentación e identificación de masas en imágenes de mamografía**”³, se plantea la investigación de técnicas de Reconocimiento de Patrones, principalmente, que permitan desarrollar una herramienta computacional que apoye a los médicos radiólogos en el desempeño de las siguientes tareas:

- Localización de masas, es decir, en el proceso de detectar en qué partes de la imagen posiblemente se encuentra una masa.
- Segmentación de masas, entendiéndolo como el proceso de aislar las regiones que posiblemente contengan masas.
- Identificación de masas, este proceso decidirá cuáles de las regiones obtenidas por el segmentador son realmente masas.

Por otra parte, la segunda fase está integrada al proyecto titulado “**Sistema de análisis computacional de características benignas y malignas en**

³ Proyecto PROMEP-2012 para el fortalecimiento del Cuerpo Académico de Reconocimiento de Patrones.

mamografías digitales mediante herramientas de procesamiento de imágenes y aprendizaje automático”⁴. Este proyecto tiene como objetivos principales:

- La creación de un conjunto de imágenes de mamografías reales con características específicas de masas de acuerdo al BI-RADS.
- La obtención y selección de las características de las regiones de la imagen identificadas como masas.
- Implementación de diversos clasificadores que determinen si una masa es benigna o maligna.

Esta segunda fase del proyecto está siendo desarrollado en forma conjunta con personal capacitado y experimentado del Hospital Regional de Alta Especialidad de Oaxaca (HRAEO). Se espera que el HRAEO nos proporcione el conocimiento en el área para integrarlo en la construcción del sistema CADx. Así como también, nos facilite imágenes de mamografías y nos apoye en el proceso de validación del sistema.

6. Conclusiones

La detección oportuna del cáncer de mama puede ayudar a reducir los índices de mortalidad y a incrementar las posibilidades de recuperación de la paciente. Hasta el momento, la mamografía es la única herramienta que ha demostrado ser efectiva en la detección oportuna. Dada la complejidad de su interpretación, existe la necesidad de contar con herramientas computacionales que apoyen en esta tarea. En este trabajo se presentaron las funciones de los sistemas CAD que apoyan a los radiólogos en el proceso de detección de anomalías. Se hizo énfasis en la descripción de la detección de las masas ya que pueden ser un indicador muy importante de la presencia de cáncer. Aunque en la actualidad existen un gran número de sistemas CAD comerciales, su estudio es importante debido a que sus tareas son parte de otras herramientas computacionales, los sistemas CADx, cuyo objetivo es apoyar en el diagnóstico del cáncer de mama.

El trabajo aquí presentado forma parte de dos proyectos de investigación que están siendo desarrollados por integrantes del Cuerpo Académico de Reconocimiento de Patrones con los cuales se pretende la implementación de un sistema CADx para el diagnóstico de masas 

⁴ Proyecto UTM-HRAEO.

7. Bibliografía

- 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.
- 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.
- De Paredes, E. S. (2007). *Atlas of Mammography* (3^a ed.). Lippincott Williams & Wilkins.
- 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. (2012). *Información general sobre el cáncer del seno (mama) (en español)*. Disponible en: <http://www.cancer.gov/espanol/pdq/tratamiento/seno/Patient> [2012, 24 de agosto]
- Knaul, F. M., Arreola, O. H., y Lozano, R. G. (2011). *Numeralia de cáncer de mama*. México: Fundación Mexicana para la Salud.
- Malagelada, A. O. (2007). *Automatic mass segmentation in mammographic images*. España: PhD thesis, Universitat de Girona.
- Pant, C. (2011). *Atlas of Breast Imaging with Mammography, Ultrasound and MRI correlation*. Jaypee Brothers Medical Publishers.
- Sampat, M., Markey, M., & Bovik, A. (2005). *Computer-aided detection and diagnosis in mammography*. A. Bovik, *Handbook of Image and Video Processing Elsevier Academic Press* (1195-1217).
- Secretaría de Salud. (2002). *Norma oficial mexicana nom-041-ssa2-2002, para la prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama*. México.
- Shinde, M. (2003). *Computer aided diagnosis in digital mammography: Classification of mass and normal tissue*. USA: Master's thesis, University of South Florida.
- Sickles, E. A. (2007). *El espectro de las asimetrías en la mama: características radiológicas, estudio diagnóstico y tratamiento*. *Clínicas Radiológicas de Norteamérica*, 45, 765-771.
- Suckling, J. (1994). *The mammographic image analysis society digital mammogram database*. *International Congress Series* (1069), 375-378.
- Velasquez, M. V. (1997). *Lenguaje en imágenes del seno*. *Revista Peruana de Radiología*, 3 (7).

Verónica Rodríguez López
Universidad Tecnológica de la Mixteca