



LA FORMACIÓN DE INGENIEROS:
UN COMPROMISO PARA EL
DESARROLLO Y LA SOSTENIBILIDAD

15 al 18
DE SEPTIEMBRE

20
20

www.acofi.edu.co/eiei2020

ALGORITMO DE DETECCIÓN Y CARACTERIZACIÓN DE TUMORES MALIGNOS O BENIGNOS EN MAMOGRAFÍAS

Bohórquez Casallas Alejandra, Salamanca Parra María Valentina, Albarracín Rojas Julián David, Cancino Suárez Sandra Liliana

**Escuela Colombiana de Ingeniería Julio Garavito, Universidad del Rosario
Bogotá, Colombia**

Resumen

El presente trabajo describe una herramienta basada en técnicas de procesamiento de imágenes mamográficas para la clasificación de tumores de mama malignos o benignos. Se utilizaron imágenes de la base de datos MIAS (Mammographic Image Analysis Society) que incluyen anotaciones del tipo de tumor que se visualiza en cada una de ellas. La metodología incluye la implementación de una herramienta que aplica inicialmente técnicas de preprocesamiento o mejoramiento de la calidad de la imagen como el filtrado espacial, las cuales permiten realizar una adecuada extracción del músculo pectoral, para poder minimizar los posibles errores que este puede causar dentro de la detección de los tumores. A continuación, se utiliza una combinación entre técnicas de umbralización y crecimiento de regiones para lograr la segmentación del área perteneciente al tumor. Una vez segmentado y localizado el tumor dentro de la imagen, se procede a su caracterización, por medio de la cuantificación de su área, junto con la proporción en porcentaje con respecto al área de toda la mama. Asimismo, se sabe que el 90% de los tumores malignos, crecen cerca a los conductos mamarios o emigran a los tejidos adiposos. Por esta razón, la metodología propuesta incluye la cuantificación de la distancia del tumor a las estructuras descritas, con el fin de determinar si el crecimiento anormal de células es benigno o maligno. Para la evaluación del desempeño de la herramienta propuesta, se utilizaron 121 imágenes que contenían 67 tumores benignos y 54 imágenes mamográficas de tumores malignos, obteniendo como resultados un error promedio del 20%, una sensibilidad de 77% y una especificidad de 73%.

Palabras clave: procesamiento de imágenes; segmentación; tumor

Abstract

The present work describes a tool based on mammographic image processing techniques for the classification of malignant or benign breast tumors. Images from the MIAS (Mammographic Image Analysis Society) database were used, which included annotations of the type of tumor displayed in each one. The methodology includes the implementation of a tool that initially applies preprocessing techniques for image quality improvement such as spatial filtering, which allow for adequate extraction of the pectoral muscle, in order to minimize the possible errors of tumor detection. Next, a combination of thresholding techniques and region grow is used to achieve segmentation of the tumor area. Once the tumor is segmented and located within the image, it is characterized, by quantifying its area and the proportion in percentage related to the breast total area. Likewise, it is known that 90% of malignant tumors grow close to the mammary ducts or migrate to adipose tissues. For this reason, the proposed method includes the quantification of the distance from the tumor to the described structures, in order to determine if the abnormal cell growth is benign or malignant. For performance evaluation of the proposed tool, 121 images containing 67 benign tumors and 54 mammographic images with malignant tumors were used, obtaining as results an average error of 20%, a sensitivity of 77% and a specificity of 73%.

Keywords: *image processing; segmentation; tumor*

1. Introducción

Actualmente el 31% de los casos de cáncer de mama en el mundo se encuentran en América Latina. El cáncer cervicouterino junto con el cáncer de mama es la principal causa de muerte en mujeres entre los 35 y 64 años en esta región (Maita, et al., 2012). La falta de una detección temprana de cáncer de mama se ha asociado directamente con una menor supervivencia de las personas que han dado positivo, es decir, con un tiempo superior a los tres meses entre la aparición de los síntomas y el inicio de su tratamiento, la supervivencia global se disminuye en un total de 12% (Piñeros, et al., 2011).

Debido a lo anterior es que en Colombia y en otros países de la región el cáncer de mama se ha comenzado a perfilar como una problemática de salud pública, haciéndose necesario buscar alternativas de detección temprana de este tipo de enfermedad. Sin embargo, la detección temprana debe complementarse con un sistema de salud bien estructurado e incluyente para poder contrarrestar los índices de mortalidad (Llanos, et al., 2012).

Las técnicas de procesamiento de imágenes pueden ser usadas como herramienta para la detección temprana del cáncer de mama. Para desarrollar algoritmo de detección de tumores benignos y malignos en mamografías basados en procesamiento de imágenes, se emplean bases de datos de imágenes médicas en donde se suministre además la información necesaria para utilizar como patrón de comparación o de referencia, por ejemplo, el tipo de tumor y su tamaño.

2. Marco Teórico

Cada mama tiene entre 15 y 20 secciones denominadas lóbulos, distribuidos como los pétalos de una margarita. Cada lóbulo tiene muchos lobulillos más pequeños que terminan en bulbos diminutos que producen leche. Los lóbulos, lobulillos y bulbos se unen mediante conductos finos denominados ductos. Estos ductos conducen al pezón en el centro de un área oscura de la piel denominada areola (ver Fig.1) (Allen, 2016).

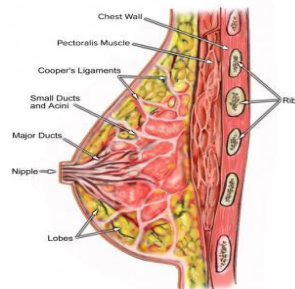


Figura 1. Anatomía del seno (Allen, 2016).

El cáncer de mama es un proceso oncológico en el cual un grupo de células sanas ubicadas en la glándula mamaria se degeneran, lo que produce una proliferación de estas células que dan origen a la conformación de un tumor. Por otro lado, se debe tener en cuenta que no todos los tumores son cáncer, ya que, un tumor se define como un conjunto de células que se han multiplicado de manera descontrolada y han dado lugar a una masa anormal. Los tumores benignos se pueden diferenciar de los malignos, en que las células de los primeros son más parecidas a las del tejido sano (Biolatto, 2019) (Sacada, 2020).

Con respecto a los tumores malignos (ver Fig.2A), estos se pueden clasificar en: Carcinoma ductal in situ que se origina en las células de las paredes de los conductos mamarios. Carcinoma ductal invasivo se origina en el conducto mamario. Carcinoma lobular in situ, se origina en las glándulas mamarias. Carcinoma lobular infiltrante se inicia en las glándulas mamarias, pero es capaz de extenderse y destruir otros tejidos. Carcinoma Inflamatorio, es un cáncer poco común, representa un 1% del total de los tumores cancerosos de la mama (Solves, 2009). El diagnóstico del cáncer se puede realizar con una autoexploración de mamas que consiste en detectar anomalías en las mamas por medio del tacto, sin embargo, una técnica más sofisticada y precisa es la mamografía (ver Fig. 2B), que emplea radiación ionizante de baja potencia aplicada directamente sobre las mamas y en la imagen resultado se pueden detectar tumores de tamaño reducido (Defilippo, 2017).

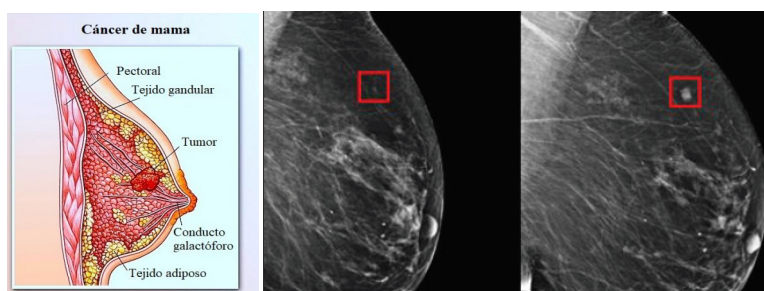


Figura 2. A. Cáncer de mama. B. Mamografía con presencia de cáncer de mama (Raya, 2019).

Algunos de los criterios para determinar si los descubrimientos de la mamografía son benignos o malignos se pueden evidenciar en la figura 3:

CARACTERÍSTICAS	SOSPECHA DE TU BENIGNO	SOSPECHA DE TU MALIGNO
INICIO	Precoz	Tardío
NÚMERO	Generalmente única	Masa solitaria
LOCALIZACIÓN	---	41% CSE, 34% Central, 14% CSI, 6% CIE, 5% CII
TAMAÑO	0,5- 5cm	Inicialmente 0,5- 1cm
FORMA	Redondeada o Lobulado	Irregular
DELIMITACIÓN	Bien Delimitados	Límites difícil de definir
CONSISTENCIA	Elástica	Pétrea o Leñosa
SUPERFICIE	Lisa	Irregular

Figura 3. Clasificación de los tumores de mama (Quijada, 2011).

3. Método

La metodología propuesta se subdivide en una etapa de preprocesamiento, para el realce de la región del tumor dentro de la mamografía; seguida de una fase de segmentación o extracción de características de la región del tumor, una etapa de toma de decisión o identificación de tipo de tumor (benigno o maligno), y finalmente, la validación correspondiente de los resultados obtenidos.

3.1. Datos Experimentales

Se usa la base de datos de mamografías MIAS. La Mammographic Image Analysis Society (MIAS) es una organización de grupos de investigación del Reino Unido dirigidos a la comprensión de las mamografías, y por esto han generado una base de datos de mamografías digitales. Las películas tomadas se han digitalizado a un borde de píxeles de 50 micras. La base de datos contiene 322 películas digitalizadas, incluye anotaciones de radiólogos en las ubicaciones de cualquier anomalía que pueda estar presente (2011, April). Brinda el carácter del tejido, la clase de anomalía presente, así como su gravedad (benigno o maligno) y las coordenadas "x,y" del centro de la anomalía, además del radio aproximado (en píxeles) de un círculo que encierra la anomalía. El tamaño de todas las imágenes es de 1024x1024 píxeles, en formato pgm (2012, December). Para el desarrollo del presente trabajo, se realiza una selección de 121 imágenes de MIAS que contienen tumores, 54 de tumores malignos y 67 de tumores benignos

3.2. Preprocesamiento

Para el preprocesamiento de las imágenes de mamografía, inicialmente se recorta cada una de ellas con el fin eliminar información irrelevante. Las mamografías tienen una orientación definida, sin embargo, esta orientación también se modificó por facilidad en la implementación de ciertos procesos descritos más adelante (ver figura 4A).

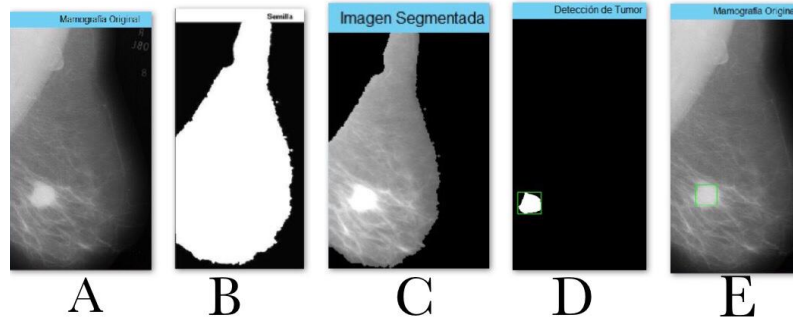


Figura 4. A. Imagen recortada y orientada a la derecha. B. Resultado de crecimiento de regiones. C. Imagen original solo con la región del seno D. Segmentación de tumor. E. Identificación del tumor sobre la imagen original.

Debido a que algunas imágenes presentan ciertas líneas originadas por el ruido propio del proceso de digitalización, se aplicó un filtro de mediana. Para eliminar las anotaciones y el músculo pectoral presente en la mamografía, se toma la imagen en escala de grises y se le aplican operadores morfológicos para eliminar pixeles aislados y rellenar las regiones de interés, para continuar así a la etapa de segmentación.

3.3. Segmentación

Se utiliza la técnica de crecimiento de regiones. La semilla inicial es ubicada en la zona del músculo pectoral de tal forma que el algoritmo extraiga la región que lo contiene (Fig. 4B). Se realiza una multiplicación de la imagen binaria resultado del crecimiento de regiones con la imagen original en escala de grises, para obtener la imagen original de mamografía únicamente con la región del seno (Fig. 4C).

Una vez obtenida esta imagen, se inicia la etapa de identificación del tumor, el cual se logra aplicando un filtro no lineal que realiza operaciones en la vecindad de los pixeles de forma deslizante. Luego se procede a ubicar los bordes del tumor y a dilatarlos, a rellenar los espacios pequeños entre ellos. Por último, se eliminan las estructuras conectadas al tumor, dando como resultado una imagen con el tumor segmentado (Fig. 4D). Cabe resaltar que para cada imagen los tumores son de diferente tamaño, por eso, se implementó un "slider" que varía el valor de referencia en pixeles que se deben eliminar, para así, tener un acercamiento del tumor a cualquier mamografía dispuesta por la base de datos.

3.4. Extracción de características

Para la etapa de extracción de características se realizó un reconocimiento de las regiones en la imagen (tumor), usando funciones existentes en Matlab para calcular el centro de dicha región, así como su área en pixeles. Es importante mencionar que en algunos casos pese a todos los pasos realizados quedan aún secciones que no son de interés, razón por la cual se eliminan si tienen un área inferior a un valor predefinido. Debido a que se conoce el tamaño de cada píxel en unidades del sistema internacional se puede hacer un cálculo del porcentaje equivalente de la malformación respecto a la totalidad del seno, sin tener en cuenta el músculo pectoral, lo cual brinda una relación de tamaño que puede ser de interés.

3.5. Identificación de tipo de tumor

Para la identificación y clasificación del tumor se recurre a ubicar un recuadro luminoso en la imagen original (Fig. 4E), esto muestra la ubicación del tumor en el seno, seguido a esto, el especialista define si el tumor está ubicado cerca a los conductos o glándulas mamarias, si esto es afirmativo, el algoritmo arroja que el tipo de tumor es maligno, de lo contrario, será benigno.

4. Resultados

A continuación, se mostrarán los resultados obtenidos una vez se ha implementado una interfaz para facilitar el uso del algoritmo por parte de un usuario. La figura 5 muestra la interfaz resultante. Inicialmente se selecciona la mamografía en la que se va a realizar el proceso de detección del tumor. En cada etapa del proceso, la interfaz va guiando al usuario y dando ciertas recomendaciones con el fin de que este obtenga el resultado esperado.

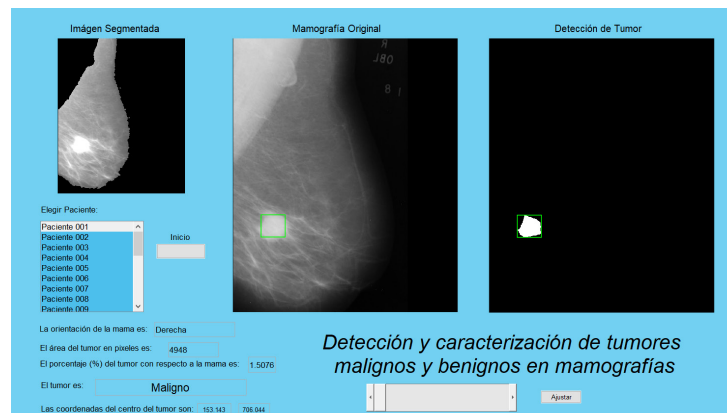


Figura 5. Mamografía con tumor maligno identificado en la interfaz.

Asimismo, el usuario puede apreciar la imagen original de la mamografía acompañada de un recuadro en un color llamativo que encierra la sección en donde el algoritmo ubica la anomalía. Esto es importante ya que podría servir de apoyo durante la revisión por parte de un radiólogo para que identifique con mayor rapidez y claridad la zona en donde se ubica la anomalía y así determinar la naturaleza del tumor. Lo anterior está basado en el hecho que dependiendo de la zona como los conductos mamaros o los lóbulos de la mama se puede considerar la anomalía como maligna o al menos potencialmente maligna, a diferencia de otros lugares que estén distanciados de estas zonas en donde la anomalía ya puede ser caracterizada como benigna.

Como información, se indica la orientación original de la mama, seguida de información como el área del tumor, el porcentaje de este comparado con la totalidad del seno, la clase asignada al tumor (benigno o maligno), y las coordenadas del centroide de la región del tumor dentro de la imagen (Fig.6).

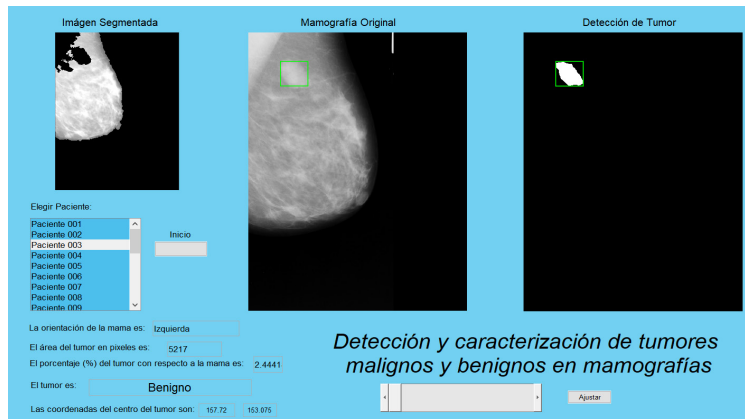


Figura 6. Mamografía con tumor benigno identificado en la interfaz.

En la tabla 1, se puede observar el promedio del porcentaje de error de las coordenadas (x, y) del centroide del tumor detectado por la interfaz, y las coordenadas del centroide del tumor proporcionadas en las anotaciones de las 121 imágenes utilizadas.

Tabla 1. Error entre los centroides del tumor de las anotaciones de la base de datos y la interfaz implementada.

	Porcentaje de error en x(%)	Porcentaje de error en y (%)
PROMEDIO	22.700	14.066

Posterior a esto se realizó una tabla de sensibilidad y especificidad, la primera se refiere a la proporción de personas que tienen un tumor maligno que posean una prueba positiva y la segunda se refiere a la proporción de personas que tengan un tumor benigno que posean una prueba negativa o normal (Vizcaíno, 2017). Este resultado se puede observar en la tabla 2.

Tabla 2. Prueba diagnóstica y resultados probables de la misma

	TUMOR MALIGNO	TUMOR BENIGNO
PRUEBA POSITIVA	51	15
PRUEBA NEGATIVA	15	40

	VERDADEROS POSITIVOS (VP)
	FALSOS POSITIVOS (FP)
	FALSOS NEGATIVOS (FN)
	VERDADEROS NEGATIVOS (VN)

Para encontrar la sensibilidad, la especificidad y la exactitud se deben de seguir las siguientes ecuaciones:

$$Sensibilidad = \frac{VP}{(VP + FN)} = \frac{51}{51 + 15} = 0.77$$

$$Especificidad = \frac{VN}{(FP + VN)} = \frac{40}{15 + 40} = 0.73$$

$$Exactitud = \frac{(VP + VN)}{n} = \frac{51 + 40}{121} = 0.75$$

Obteniendo una sensibilidad del 77%, una especificidad del 73% y una exactitud del 75%.

5. Análisis de Resultados

La interfaz tiene un error aceptable con respecto a la detección de tumores (aproximadamente 75%), esto quiere decir que, junto al conocimiento del usuario, es posible utilizar la interfaz implementada en un ambiente clínico para ayudar al diagnóstico de cáncer de mama. Por otro lado, la sensibilidad de 77% y una especificidad de 73%.

Se evidencia que la no estandarización de las características de las mamografías brindadas por la base de datos MIAS, lo cual impacta el desarrollo de un algoritmo de detección completamente automático, esto debido a factores como la orientación en la mamografía y la presencia de bandas negras a sus costados que conllevan a errores en la etapa de segmentación de la región del tumor.

Cabe resaltar que, en algunos casos de detección, además de la ubicación típica de los tumores se deben tener conocimientos del comportamiento de estos, así como su fisiología. Por ejemplo, en la figura 4A, se observa que el tumor maligno no se encuentra en las zonas típicas de crecimiento, sino que se encuentra en otra fase, en la que el tumor maligno emigra hasta el tejido adiposo. En este caso, la característica descrita sumada al hecho de que tiene una forma más irregular que el tumor benigno, ayuda a determinar su malignidad.

6. Conclusiones y trabajo futuro

Esta investigación se realizó con el fin de detectar tumores mamarios y poder de esta forma clasificarlos, haciendo uso de técnicas de segmentación, basadas en métodos de segmentación de umbrales y operadores morfológicos. Las mamografías fueron tomadas de la base de datos MIAS, las técnicas implementadas dan resultados eficientes, muy cercanos en comparación con investigaciones previas. Nuestra propuesta de algoritmo presentada es fácil de ejecutar e intuitiva, razón por la cual se puede gestionar fácilmente sin tener una capacitación previa. Como trabajo futuro, se propone implementar una interfaz de detección de tumores en imágenes de mamografía que sea 100% automática, facilitando su manipulación por cualquier tipo de usuario y disminuyendo los errores que puedan ocurrir por la selección de parámetros de forma manual y subjetiva.

7. Referencias

- Maita, F; Llanos, J. L; Panozo, S. V; Muñoz, L; Gutiérrez, C and Zegarra, W (2012). Diagnostic value of ultrasound and mammography in patients with breast neoplasms of the Hospital Obrero N ° 2 of the Caja Nacional de Salud. Gaceta Médica Boliviana, Vol. 35, No. 2.

- Piñeros, M; Sánchez, R; Perry, F; García, O .A; Ocampo, R and Cendales, R (2011). Demoras en el diagnóstico y tratamiento de mujeres con cáncer de mama en Bogotá, Colombia. Salud Pública de México, Vol. 53, No. 6.
- (2016, June). Breast Anatomy. Consultado el 4 de marzo de 2020 en <https://reference.medscape.com/article/1273133-overview>
- (2019, July). Tumor maligno y benigno, ¿cuáles son las diferencias? Consultado el 4 de marzo de 2020 en <https://mejorconsalud.com/tumor-maligno-y-benigno-cuales-son-las-diferencias/>
- (2020, May). Patologías benignas de mama. Consultado el 4 de marzo de 2020 en <https://www.webconsultas.com/patologias-benignas-de-mama/tumores-benignos-de-la-mama-2142>
- Solves, J. A (2009). Análisis de la densidad de mama asistido por ordenador. Universidad Politécnica de Valencia
- (2017, September). Carcinoma de mama: exámenes, tratamiento y pronóstico. Consultado el 4 de marzo de 2020 en <https://www.fisioterapiaparatodos.com/carcinoma-de-mama/>
- (2019, May). Predecir el cáncer de mama sólo con una mamografía es posible gracias a una nueva IA. Consultado el 4 de marzo de 2020 en https://www.lespanol.com/omicro/tecnologia/20190508/predecir-cancer-mamografia-posible-gracias-nueva-ia/396961683_0.html
- (2011, October). Conducta ante una tumoración en mama. Consultado el 4 de marzo de 2020 en <https://es.slideshare.net/LeopoldoJQuijadaB/conducta-ante-una-tumoracin-de-mama>
- (2011, April). Mammographic Image Analysis Homepage. Consultado el 4 de marzo de 2020 en <https://www.mammoimage.org/databases/>
- (2012, December). The mini-MIAS database of mammograms. Consultado el 4 de marzo de <http://peipa.essex.ac.uk/info/mias.html>
- Vizcaíno, G.J (2017). Importancia del cálculo de la sensibilidad, la especificidad y otros parámetros estadísticos en el uso de las pruebas de diagnóstico clínico y de laboratorio. Editora Médica Colombiana S.A, Vol. 23, No. 7-8.

Sobre los autores

- **Alejandra Bohórquez Casallas:** Estudiante de Ingeniería Biomédica. alejandra.bohorquez@mail.escuelaing.edu.co
- **María Valentina Salamanca Parra:** Estudiante de Ingeniería Biomédica. maria.salamanca-p@mail.escuelaing.edu.co
- **Julián David Albarracín Rojas:** estudiante de Ingeniería Biomédica. julian.albarracin@mail.escuelaing.edu.co
- **Sandra Liliana Cancino Suárez:** Ingeniera electrónica. Profesor Asociado. Programa de Ingeniería Biomédica. sandra.cancino@escuelaing.edu.co

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2020 Asociación Colombiana de Facultades de Ingeniería (ACOFI)