

Procesamiento de mamografías para remoción del tejido muscular

Chiara V. Valenzuela L, Andrea C. Terraza S, Laura C. García

^a *Universidad Sergio Arboleda, Bogotá, Colombia*

Abstract

El cáncer de mama es la principal afecta a millones de mujeres en todo el mundo y la detección temprana es fundamental para mejorar las tasas de supervivencia y la eficacia del tratamiento. La mamografía toma un papel de suma importancia en el diagnóstico temprano, ya que permite a los especialistas hallar anomalías en el tejido mamario que pueden indicar neoplasias malignas. Sin embargo, la interpretación de las mamografías se enfrenta a retos como el ruido, el bajo contraste y la presencia de zonas irrelevantes, que complican la identificación precisa de las lesiones. Este proyecto tiene como finalidad mejorar las condiciones para el análisis de las mamografías mediante técnicas de procesamiento de imágenes que se centran en la eliminación del tejido muscular visible. Se espera que, luego de aplicar las técnicas de procesamiento, se tengan mamografías más claras al eliminar estructuras no deseadas, lo que ayudará a identificar masas de manera más eficaz.

Keywords: cáncer de mama, mamografía, procesamiento de imágenes, análisis

1. Introducción

El cáncer de mama es la principal causa de muerte por cáncer en mujeres a nivel mundial. Según la Organización Mundial de la Salud, cada año se diagnostican alrededor de 2.3 millones de casos, y aproximadamente 685,000 mujeres pierden la vida debido a esta enfermedad, (Organización Mundial de la Salud (OMS) (2024)). En este contexto, la detección temprana es importante para mejorar la tasa de supervivencia y la efectividad del tratamiento. Las mamografías son una herramienta fundamental, porque permiten a los especialistas

identificar anomalías en el tejido mamario que podrían indicar la presencia de tumores malignos.

No obstante, la interpretación de las mamografías enfrenta varios desafíos. Las imágenes a menudo presentan ruido, bajo contraste y la inclusión de áreas irrelevantes, como el músculo pectoral, lo que puede dificultar la identificación precisa de lesiones sospechosas. Este proyecto tiene como objetivo mejorar el análisis de las mamografías mediante técnicas de procesamiento de imágenes, enfocadas en la eliminación del tejido muscular visible. Al centrarse exclusivamente en las áreas que con-

tienen tejido mamario, estas técnicas facilitarán el análisis posterior y aumentarán la precisión del diagnóstico.

2. Contexto del problema

En las mamografías, especialmente en la proyección mediolateral oblicua (MLO), es común que aparezca el músculo pectoral, que puede ocupar una parte significativa de la imagen. Este tejido, al tener características visuales similares al tejido mamario, complica la segmentación y el análisis automatizado. Además, en otras proyecciones, como la cráneo-caudal (CC) y mediolateral (ML), las estructuras cercanas a los bordes como músculo y huesos, también pueden interferir en el análisis. (Avci and Karakaya (2023)). La presencia de estas áreas no deseadas puede generar diagnósticos menos precisos, aumentar el riesgo de falsos positivos o negativos, y complicar la tarea de los especialistas.

3. Problema

El principal desafío que trata este proyecto es implementar técnicas de procesamiento de imágenes que permitan la eliminación del músculo pectoral y otras estructuras cercanas a los bordes de las mamografías, sin recurrir a algoritmos de aprendizaje automático. La solución debe aplicar una serie de técnicas de pre-procesamiento y segmentación que actúen directamente sobre las áreas no deseadas, mejorando la claridad de las zonas donde se encuentra el tejido mamario. Estas técnicas incluirán la mejora del contraste, el uso de filtros morfológicos, la detección de bordes y la

segmentación basada en la intensidad de los píxeles.

El objetivo es obtener mamografías limpias, con el músculo pectoral eliminado de manera precisa, permitiendo un análisis más claro y eficiente de las áreas donde pueden aparecer lesiones, como masas o microcalcificaciones.

4. Descripción de las imágenes

En este proyecto se usarán imágenes provenientes del conjunto de datos **CBIS-DDSM (Curated Breast Imaging Subset of the Digital Database for Screening Mammography)**, una base de datos ampliamente reconocida y utilizada en el ámbito de la investigación de cáncer de mama. Este dataset contiene imágenes de mamografías digitales obtenidas a partir de exámenes de detección de cáncer de mama. Sawyer-Lee et al. (2016).

4.1. Modalidad y Formato:

- Las imágenes son mamografías digitales en formato DICOM, aunque se van a transformar a un formato JPG lo que asegura una alta calidad y mejor procesamiento de las imágenes.
- Incluyen vistas cráneo-caudal (CC), medio-lateral (ML) y medio-lateral oblicua (MLO) de cada seno.

4.2. Anotaciones y Etiquetado:

- Las imágenes están etiquetadas y anotadas para indicar la presencia de lesiones sospechosas, como masas y microcalcificaciones,
- Se proporciona información adicional sobre el tipo de lesión (benigna o maligna).

4.3. Resolución y Calidad:

- Las mamografías cuentan con una alta resolución, lo que permite un análisis detallado del tejido mamario.
- El dataset incluye imágenes de diferentes calidades, lo que permite evaluar la robustez de los algoritmos en situaciones de imágenes con diferentes niveles de ruido y contraste.

4.4. Diversidad de Casos:

El CBIS-DDSM incluye una amplia gama de casos, desde imágenes normales hasta aquellas con diversas anomalías.

5. Ejemplos de las imágenes a utilizar

En las mamografías seleccionadas del conjunto de datos CBIS-DDSM, se incluye la proyección mediolateral oblicua (MLO). A continuación, se relatan los aspectos que se pueden encontrar en cada una de las imágenes ejemplo:

- La Figura 1 es una mamografía normal tomada desde el lado izquierdo con una resolución de 3064x4728.
- La Figura 2 tiene una mamografía benigna capturada del lado derecho con una resolución de 3080x4568 píxeles.
- Finalmente, la figura 3 tiene una neoplasia maligna y una resolución de 3080x4554 píxeles

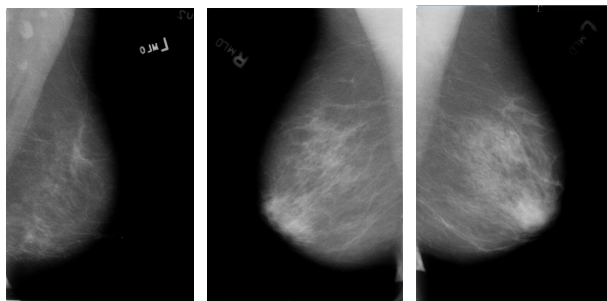


Figure 1: LMLO Figure 2: RMLO Figure 3: LMLO

Figure 4: Mamografías de proyección (Oblicua Latero-Medial) izquierda o derecha con resolución entre 2500 y 5800 píxeles de ancho y alto.

6. Posible resultado

La Figura 5 muestra una mamografía original con el músculo pectoral visible, y la Figura 6 presenta la misma mamografía después de aplicar las técnicas de procesamiento para eliminar el tejido muscular.

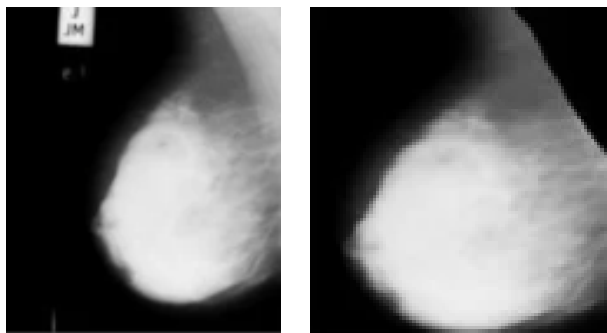


Figure 5: Mamografía original con el tejido muscular

Figure 6: Mamografía resultante con el tejido muscular removido

Figure 7: Mamografías antes y después de procesamiento.

Estas comparaciones destacan cómo la imagen procesada mejora la visibilidad del tejido

mamario, evitando la confusión entre tejidos y facilita la identificación de posibles lesiones, lo que contribuye a mejorar la precisión del diagnóstico asistido por computadora (CAD).

References

Avcı, H. and Karakaya, J. (2023). A novel medical image enhancement algorithm for breast cancer detection on mammography images using machine learning. *MDPI, Escuela de medicina de la Universidad Hacettepe*.

Organización Mundial de la Salud (OMS) (2024). Breast cancer. *OMS*.

Sawyer-Lee, R., Gimenez, F., Hoogi, A., and Rubin, D. (2016). Curated breast imaging subset of digital database for screening mammography (cbis-ddsm) [data set]. *The Cancer Imaging Archive*.