

Desarrollo de una Metodología de Procesamiento de Imágenes para la Detección Temprana del Cáncer de Mama

Miguel Angel Thomas Gonzalez
miguel.thomas01@usa.edu.co
¹ and Carlos Mario Bermudez Galvan
carlos.bermudez01@usa.edu.co²

¹Escuela de Ciencias Exactas e Ingeniería

¹Universidad Sergio Arboleda

¹Bogotá, Colombia 2024

1. Abstract

Early detection of anomalies in mammography images is crucial for the effective diagnosis and treatment of breast cancer. However, low contrast visibility and tissue overlap make precise segmentation of the area of interest challenging. This study aims to develop an image processing method that improves segmentation of areas of interest in mammograms using advanced image processing techniques. To achieve this, we apply a series of contrast expansion filters, edge detection techniques, intensity transformations, morphological operations, and thresholding. The results demonstrate a significant improvement in the segmentation of the area of interest, facilitating the precise identification of anomalous regions in the images.

2. Resumen

La detección temprana de anomalías en imágenes de mamografía es crucial para el diagnóstico y tratamiento efectivo del cáncer de mama, pero la baja visibilidad de contrastes y la superposición de tejidos dificultan la segmentación precisa del área de interés. Este estudio tiene como objetivo principal desarrollar un método de procesamiento de imágenes que mejore la segmentación de áreas de interés en mamografías mediante técnicas avanzadas de procesamiento de imágenes. Para lograr esto, aplicamos una serie de filtros de expansión de contraste, técnicas de detección de bordes, transformaciones de intensidad, operaciones morfológicas y umbralización. Los resultados demostraron una mejora significativa en la segmentación del área de interés, facilitando la identificación precisa de regiones anómalas en las imágenes.

Palabras clave: morfología, mamografía, segmentación, bordes, umbralización, contraste.

3. Introducción

El cáncer de mama es una enfermedad que afecta a millones de mujeres en todo el mundo, siendo una de las principales causas de morbilidad y mortalidad en el ámbito de la salud femenina. Se origina en las mamas, glándulas esenciales compuestas por lobulillos, conductos y tejido conectivo, responsables de la producción y transporte de leche, así como de brindar soporte estructural. La multiplicación descontrolada de células mamarias puede derivar en un cáncer que, de no ser detectado y tratado a tiempo, puede propagarse a otras partes del cuerpo a través del sistema circulatorio y linfático, dando lugar a metástasis [CDC, 2023].

La detección temprana de este tipo de cáncer es crucial para el éxito del tratamiento y la mejora de los resultados de salud de las pacientes. La mamografía se posiciona como la herramienta más efectiva para la detección temprana, permitiendo identificar anomalías en las imágenes de las mamas. No obstante, para lograr una detección más precisa y eficiente, se requiere el empleo de técnicas avanzadas de procesamiento de imágenes y análisis computacional [ACS, 2022].

En la literatura, diversos estudios han abordado estas dificultades mediante el empleo de técnicas avanzadas de procesamiento de imágenes. Márquez Sosa y María Mercedes destacaron la importancia del procesamiento digital en la caracterización de lesiones mamarias, utilizando métodos como la segmentación y el análisis morfológico para mejorar la interpretación de las imágenes mamográficas [Márquez Sosa, 2021]. Por otro lado, Díaz Sotolongo y Costa Alonso combinaron técnicas de procesamiento de imágenes digitales e inteligencia artificial para clasificar lesiones patológicas, utilizando algoritmos que incorporan parámetros de forma y textura en la formación de vectores de características [Alejandro Díaz Sotolongo, 2013].

A pesar de los avances logrados en el campo del procesamiento de imágenes para la detección temprana del cáncer de mama, aún persisten desafíos significativos que requieren atención continua. Uno de estos desafíos es la mejora de las técnicas de segmentación para garantizar una detección más precisa y temprana de posibles cánceres. En este contexto, el objetivo principal del proyecto fue desarrollar y validar un algoritmo de segmentación altamente eficiente y preciso, que combinara una variedad de técnicas avanzadas.

Se implementó un enfoque multidisciplinario que integró transformaciones de intensidad, detección de bordes, operaciones morfológicas y umbralización para optimizar la segmentación del área de interés en las imágenes de mamografía. Este enfoque se basó en la aplicación sistemática de filtros y técnicas de procesamiento de imágenes, adaptadas específicamente para abordar los desafíos de visibilidad de contrastes y superposición de tejidos en las mamografías. Además, se realizó una exhaustiva validación del algoritmo utilizando conjuntos de datos clínicos reales, lo que permitió evaluar su rendimiento en términos de precisión, sensibilidad y especificidad.

El trabajo se centró en aumentar la precisión de la segmentación y en reducir la tasa de falsos positivos y negativos. Esto fue fundamental para garantizar una detección más efectiva y eficiente de las lesiones mamarias, lo que a su vez puede contribuir significativamente a mejorar los resultados clínicos y la calidad de vida de las pacientes.

4. Metodología

4.1. Descripción de imágenes

La base de datos MIAS MiniMammographic proporciona una colección de imágenes mamográficas digitales que serán utilizadas en este proyecto para el análisis de anomalías mamarias. Estas imágenes están en formato .pgm (Portable Graymap), digitalizadas a una resolución de 200 micrones por píxel y tienen un tamaño de 1024 píxeles por 1024 píxeles.

Cada imagen está etiquetada con información detallada que incluye:

- Número de referencia en la base de datos MIAS.
- Carácter del tejido de fondo: Fatty (graso), Fatty-glandular (graso-glandular), Dense-glandular (denso-glandular).
- Clase de anomalía presente: Calcificación, Circumscribed Mass (masa circunscrita), Spiculated Mass (masa espiculada), Architectural Distortion (distorsión arquitectural), Asymmetry (asimetría), Normal (sin anomalía).

- Gravedad de la anomalía: Benigna o Maligna.
- Coordenadas x e y del centro de la anomalía en la imagen.
- Radio aproximado de la anomalía en píxeles.

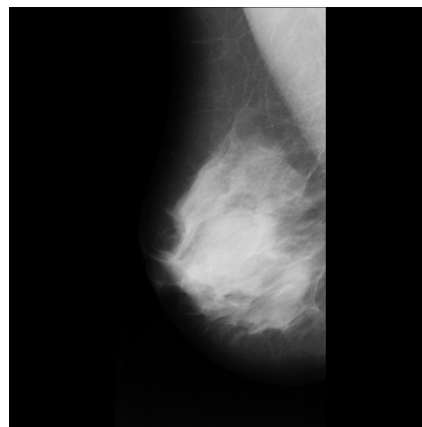


Figura 1: Ejemplo de Mamografía presente en las imágenes escogidas.

Las imágenes en la base de datos están identificadas por un número de referencia único, seguido de la información sobre el tipo de tejido, la clase de anomalía y su gravedad. Además, se proporcionan las coordenadas del centro de la anomalía en la imagen y su radio aproximado en píxeles.

Por ejemplo, la imagen mdb028 tiene la siguiente información:

- Número de referencia: mdb028
- Carácter del tejido de fondo: Fatty (graso)
- Clase de anomalía presente: Circumscribed Mass (masa circunscrita)
- Gravedad de la anomalía: Maligna
- Coordenadas del centro de la anomalía: (338, 314)
- Radio aproximado de la anomalía: 56 píxeles

Estos datos detallados serán fundamentales para el desarrollo y la evaluación de algoritmos de procesamiento de imágenes destinados a la detección temprana y caracterización de anomalías mamarias, contribuyendo así a mejorar el diagnóstico del cáncer de mama y el bienestar de las pacientes.

4.2. Diagrama del Pipeline de Procesamiento

A continuación se visualiza el diagrama del pipeline desarrollado.

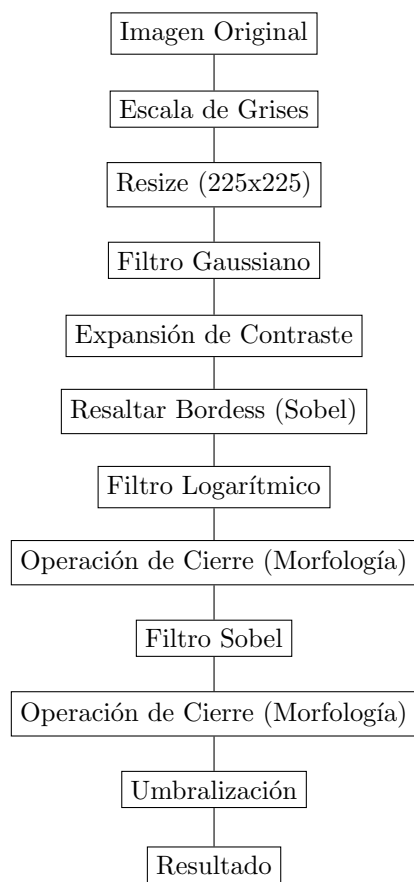


Figura 2: Diagrama del Pipeline de Procesamiento de Imágenes

4.3. Descripción del Pipeline de Procesamiento

El pipeline de procesamiento de imágenes desarrollado para este proyecto tiene como objetivo mejorar la segmentación de áreas de interés en mamografías, permitiendo una detección más precisa de anomalías. A continuación, se describen los pasos del pipeline:

1. Conversión a Escala de Grises:

Las imágenes originales en formato RGB se convierten a escala de grises para simplificar el procesamiento y enfocarnos en la información de intensidad, crucial para la detección de anomalías en mamografías. Este paso reduce la complejidad computacional y elimina la influencia del color.

2. Redimensionamiento de la Imagen:

Las imágenes se redimensionan a 225x225 píxeles. Este tamaño estándar facilita el procesamiento uniforme y reduce la carga computacional sin comprometer la resolución necesaria para identificar características importantes en las mamografías. Además de conservar la relación de aspecto original (1024x1024, cuadrada).

3. Filtro Gaussiano:

Se aplica un filtro gaussiano para suavizar la imagen y reducir el ruido. Esto es especialmente importante en mamografías donde el ruido puede oscurecer detalles críticos. El filtro gaussiano atenúa las variaciones de intensidad no deseadas, permitiendo un análisis más claro de las estructuras internas.

4. Expansión de Contraste:

La expansión de contraste se utiliza para mejorar la visibilidad de las estructuras en la imagen. Este paso ajusta los niveles de intensidad de los píxeles, maximizando el rango de valores de intensidad. Esto es crucial en mamografías, donde las diferencias sutiles en contraste pueden indicar la presencia de anomalías.

5. Detección de Bordess con Sobel:

El operador Sobel se utiliza para resaltar los bordes en la imagen, calculando las derivadas de intensidad en las direcciones horizontal y vertical. Este paso es vital para identificar los límites de las regiones de interés, ayudando a delinear áreas potencialmente anómalas con mayor precisión.

6. Filtro Logarítmico:

Se aplica un filtro logarítmico para comprimir el rango dinámico de la imagen, lo que puede ayudar a resaltar características importantes en presencia de variaciones de iluminación. Esta transformación mejora la diferenciación de estructuras que podrían no ser visibles con una escala lineal de intensidad.

7. Operación de Cierre Morfológico:

La operación de cierre morfológico se utiliza para cerrar pequeños agujeros dentro de las regiones de interés y conectar elementos disjuntos. Este paso es esencial para crear contornos más definidos y eliminar pequeñas imperfecciones que podrían interferir con la segmentación.

8. Sobel:

Se aplica nuevamente el operador Sobel para una detección de bordes refinada después de la operación de cierre morfológico. Esto asegura que los bordes de las regiones de interés estén claramente definidos, mejorando la precisión en la segmentación.

9. Operación de Cierre Morfológico:

Se repite la operación de cierre morfológico para eliminar cualquier pequeño artefacto introducido por la detección de bordes y para suavizar aún más las regiones segmentadas. Este paso garantiza que las regiones de interés sean continuas y bien definidas.

10. Umbralización:

Finalmente, se aplica una técnica de umbralización para segmentar las regiones de interés. Este paso convierte la imagen en una binaria, destacando las regiones

anómalas con mayor claridad. La umbralización facilita la identificación de áreas sospechosas al simplificar la imagen a una representación de alto contraste.

Cada uno de estos pasos contribuye significativamente a la mejora de la segmentación en las mamografías, permitiendo una detección más precisa y fiable de anomalías. Las técnicas aplicadas no solo mejoran la visibilidad de las estructuras de interés, sino que también minimizan las interferencias y artefactos que podrían complicar el análisis. A pesar de las ventajas, es importante considerar las limitaciones inherentes a cada técnica y cómo estas pueden afectar los resultados finales.

5. Resultados

5.1. Pipeline de Procesamiento

1. Escala de Grises

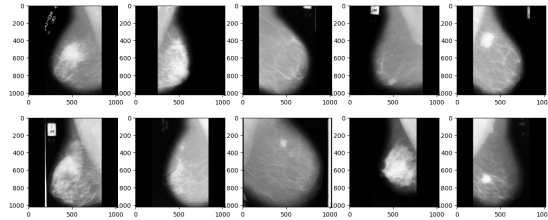


Figura 3: Resultado Escala de Grises

2. Resize de Imagen a 225x225

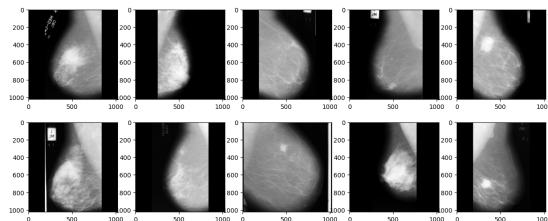


Figura 4: Resultado Resize

3. Filtro Gaussiano

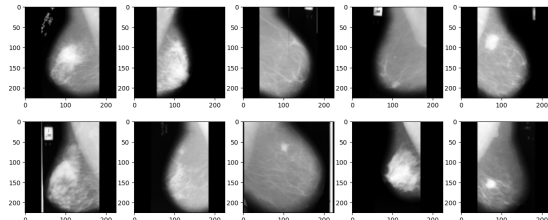


Figura 5: Resultado Filtro Gaussiano

4. Expansión de Contraste

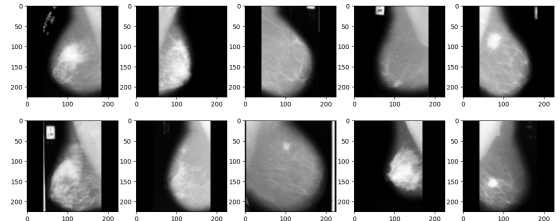


Figura 6: Resultado Expansión de Contraste

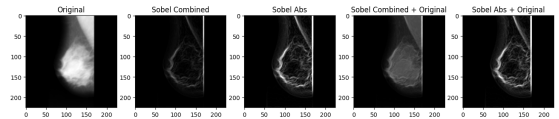


Figura 7: Resultado Resaltar Bordes con Sobel

5. Resaltar Bordes con Sobel

6. Filtro Logarítmico

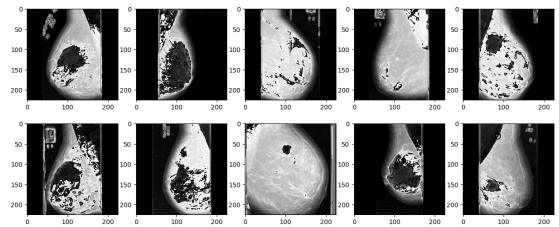


Figura 8: Resultado Filtro Logarítmico

7. Operación de Cierre (Morfología)

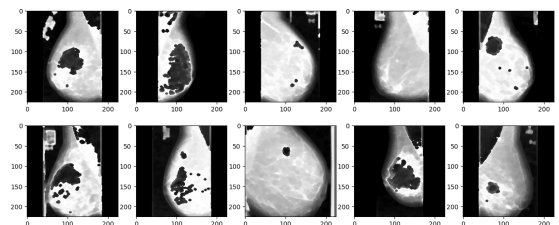


Figura 9: Resultado Operación de Cierre (Morfología)

8. Segundo Resaltado de Bordes con Sobel

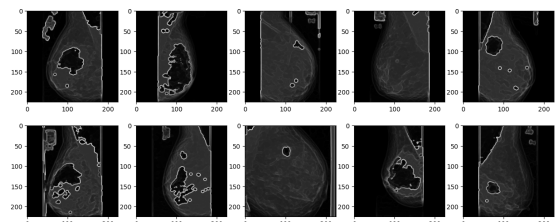


Figura 10: Resultado Segundo Resaltado de Bordes con Sobel

9. Segunda Operación de Cierre (Morfología)

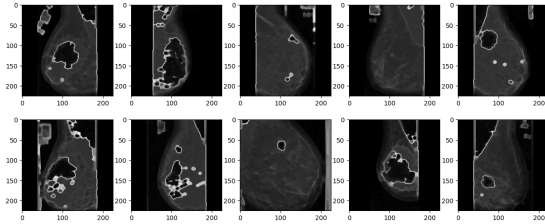


Figura 11: Resultado Segunda Operación de Cierre (Morfología)

10. Umbralización

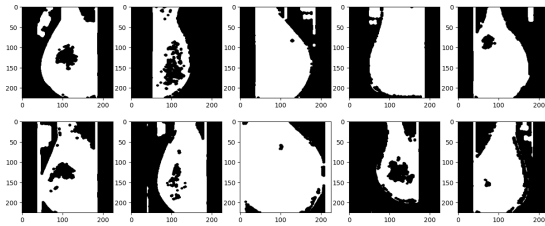


Figura 12: Resultado Umbralización

5.2. Resultados Finales

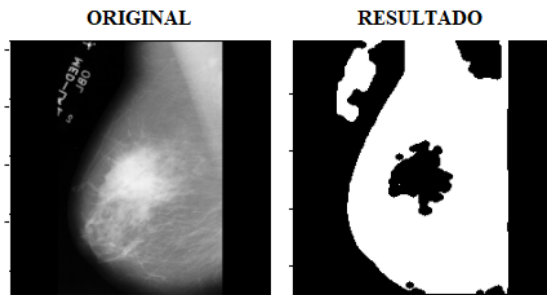


Figura 13: Resultado Final con la Primera Imagen de Entrada

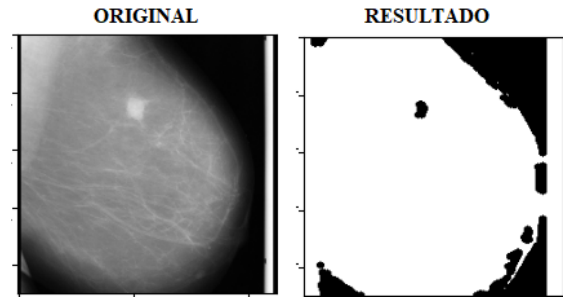


Figura 14: Resultado Final con la Segunda Imagen de Entrada

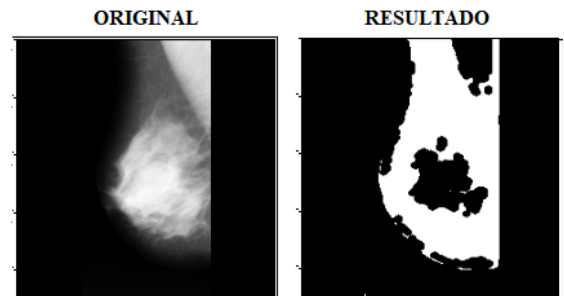


Figura 15: Resultado Final con la Tercera Imagen de Entrada

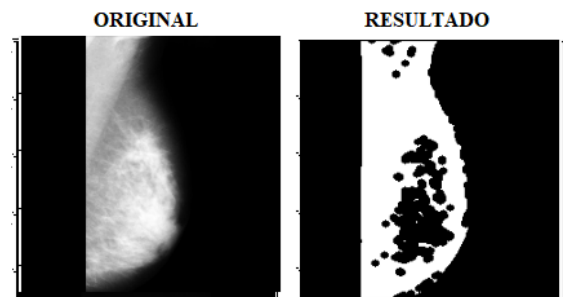


Figura 16: Resultado Final con la Cuarta Imagen de Entrada

6. Discusión

Durante el proceso de segmentación del área de interés en mamografías para la detección de tumores, se ha observado un rendimiento generalmente satisfactorio del algoritmo. Sin embargo, se han identificado áreas de mejora que son importantes de discutir.

Una de las limitaciones encontradas es la omisión o pérdida del rastro de la región de interés en imágenes de tamaño reducido. Esto sugiere la necesidad de investigar técnicas específicas para mejorar la detección en este tipo de imágenes, manteniendo la precisión del algoritmo.

Otro aspecto relevante es la segmentación incorrecta de elementos anatómicos como nervios, los cuales pueden ser malinterpretados como la región de interés. Este desafío resalta la importancia de incorporar análisis detallados de texturas y características específicas de los tumores para mejorar la especificidad del algoritmo.

7. Conclusiones

* El algoritmo desarrollado ha demostrado ser efectivo para la segmentación de áreas de interés en mamografías, utilizando técnicas como la escala de grises, resaltado de bordes con Sobel y operaciones morfológicas.

* La incorporación de filtros como el logarítmico y el gaussiano ha contribuido significativamente a mejorar la calidad de las imágenes procesadas, facilitando la detección de estructuras relevantes.

* Aunque se logró una segmentación precisa en la mayoría de los casos, persisten desafíos en imágenes de baja resolución y en presencia de estructuras anatómicas complejas, como nervios.

Estas conclusiones destacan la importancia del trabajo realizado y señalan direcciones prometedoras para futuras investigaciones en el campo del procesamiento de imágenes médicas para la detección temprana de enfermedades.

/

8. Referencias

- [ACS, 2022] ACS (2022). Recomendaciones de la american cancer society para la detección temprana del cáncer de seno.
- [Alejandro Díaz Sotolongo, 2013] Alejandro Díaz Sotolongo, D. B. C. A. (2013). Combinación de técnicas de procesamiento digital de imágenes e inteligencia artificial para el análisis de imágenes de mamografías.
- [CDC, 2023] CDC (2023). ¿qué es el cáncer de mama? Technical report, Centros para el Control y la Prevención de Enfermedades.
- [Márquez Sosa, 2021] Márquez Sosa, M. M. (2021). Procesamiento de imágenes de mamografía digital para la caracterización de lesiones mamarias.