



BENEMERITA UNIVERSIDAD AUTÓNOMA DE  
PUEBLA

---

---

FACULTAD DE CIENCIAS FISICO  
MATEMATICAS

VARIEDADES DIFERENCIABLES EN  
IMAGENES MAMOGRAFICAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIATURA EN FISICA APLICADA

P R E S E N T A :

DAVID BERNARDO REMIGIO

TUTOR

DR. CRISTIAN HEBER ZEPEDA FERNANDEZ  
DR. JAVIER MIGUEL HERNANDEZ LOPEZ



CIUDAD UNIVERSITARIA, PUEBLA, MÉX., JUNIO 2024



*«Solo hazlo, es mas facil hacerlo que explicar como hacerlo.»*



# Agradecimientos

Estos años han sido particularmente excepcionales, quiero agradecer a todos aquellos que me acompañaron en el extenso camino que significó atravesar la universidad, que bien dicho, no fue fácil en absoluto.

Agradezco a mis padres por haber vuelto posible este logro que hoy puedo agradecer por mi mismo, por los constantes ánimos que me daban cada vez que las cosas parecían volverse algo más difíciles de lo común y por los consejos que desde la experiencia me dejaron.

A mis hermanas, con las que compartí mucho en este trayecto que visto de lejos suena simple, pero del que fueron siempre parte fundamental y por el cual estaré agradecido toda la vida de forma incondicional.

A mis amigos y amigas, que con mucho o poco y aun pese a la adversidad de sus propias vidas, me han regalado un poquito de su tiempo a consolarme en mis desvaríos. A Xochitl, Maritza, Grissel, Maria Fernanda, José Ángel, Leonardo, Brayan, Israel, Raymundo y los otros que quedo por mencionar, particularmente por su excepcional compañía en los tiempos turbulentos que vivimos y que cada uno de nosotros conocemos a la perfección.

Sin ninguno de ustedes nada de esto hubiera sido posible y no me queda nada más que agradecer infinitamente y para toda la vida este apoyo incondicional que me ha servido de tanto.

Finalmente, quiero expresar mis buenos deseos a todos y cada uno, que la vida nos trate bien y que al mirar atrás, desde la nostalgia, encontremos buenas razones para seguir viviendo nuestras vidas felices y en paz.



# Índice general

<b>Agradecimientos</b>	<b>v</b>
<b>1. Motivación</b>	<b>5</b>
<b>2. Objetivos</b>	<b>7</b>
2.1. Objetivo General . . . . .	7
2.2. Objetivos Específicos . . . . .	8
<b>3. Introduccion</b>	<b>9</b>
<b>4. Anormalidades mamarias</b>	<b>13</b>
4.1. Tumores y métodos de detección . . . . .	13
4.1.1. Tumores benignos . . . . .	14
4.1.2. Tumores malignos . . . . .	15
4.2. Metodos de deteccion . . . . .	17
4.2.1. Ultrasonido . . . . .	18
4.2.2. Mamografia . . . . .	20
4.2.3. Resonancia magnetica . . . . .	22
<b>5. Segmentacion de imagen</b>	<b>25</b>
5.0.1. Segmentacion de imagen . . . . .	25
5.0.2. Estado del arte . . . . .	26
<b>6. Variedades diferenciables</b>	<b>33</b>
6.1. Variedades diferenciables . . . . .	33
6.1.1. Metricas . . . . .	35
6.1.2. variedades diferenciables y metricas . . . . .	38
<b>7. Metodologia</b>	<b>41</b>
<b>8. Resultados y analisis</b>	<b>49</b>
8.1. Resultados . . . . .	49
<b>9. Conclusiones</b>	<b>57</b>
9.0.1. Retos y Metas a futuro . . . . .	58

Bibliografía	58
Apéndice A. Tablas Completas	61

# Lista de Figuras

3.1. Primer Radiografia de la Historia . . . . .	10
4.1. Esquema de tumor benigno . . . . .	14
4.2. Estructura de tumor maligno . . . . .	15
4.3. Ecografia mamaria . . . . .	18
4.4. Mamografia mamaria . . . . .	21
4.5. Imagen obtenida mediante resonancia magnetica . . . . .	22
5.1. Clasificacion de segmentacion de imagen . . . . .	26
5.2. Segmentacion por region . . . . .	28
5.3. Segmentacion basada en thresholding . . . . .	29
6.1. Bolas disjuntas . . . . .	34
6.2. Metricas en el el espacio . . . . .	36
6.3. Proceso de analisis de imagenes mamograficas . . . . .	38
7.1. Pagina del dataset . . . . .	41
7.2. Datos de las imágenes . . . . .	42
7.3. Mamografía . . . . .	43
7.4. zona analizada del tumor . . . . .	44
7.5. Escala de tonos . . . . .	44
7.6. De imagen a matriz de datos . . . . .	45
7.7. Zona analizada del tejido sano . . . . .	46
8.1. Histograma del tejido del tumor . . . . .	49
8.2. Histograma del tejido sano . . . . .	50
8.3. comparacion de metricas . . . . .	51
8.4. Metricas coorelacionadas . . . . .	52
8.5. Ejemplo de ajuste de histograma . . . . .	53
8.6. Medias de los tejidos . . . . .	54



# Lista de Tablas

7.1. Datos del dataset . . . . .	43
7.2. Datos del tejido sano . . . . .	46
8.1. Metricas de los tejidos . . . . .	50
8.2. Medias y sigma del tejido sano . . . . .	53
8.3. Medias y sigma de los tumores . . . . .	54
8.4. Promedios de las medias y las sigmas de los datos . . . . .	54
A.1. Datos del tejido del tumor . . . . .	61
A.2. Datos del tejido sano . . . . .	62
A.3. Metricas de los tejidos . . . . .	63
A.4. Medias y sigma del tejido sano . . . . .	64
A.5. Medias y sigma de los tumores . . . . .	65



# 1 Motivación

La física es una ciencia que puede ser bastante flexible en cuanto a los ámbitos en los que puede ser aplicados, debido a la variedad de herramientas que presenta y a la forma tan practica de desarrollar los problemas que surgen de vez en cuando en la vida cotidiana, además de que la comprensión de los fenómenos naturales hacen posible que exista una estrecha relación con el campo de la medicina.

Es común escuchar términos médicos como el de radiografía", "tomografía", "ultrasonido", inclusive en herramientas que pareciera que no tienen mucho que ver con la física como algunos aparatos ortopédicos que usan los principios de torque y palanca, entre muchos otros , sin embargo es útil reconocer que estos avances y otros mas se han dado gracias a que existe una muy buena comprensión de los fenómenos físicos y el estudio matemático a cerca de estas cuestiones.

En los ultimo años, la revolución tecnológica que ha traído consigo el desarrollo de herramientas como el de la inteligencia artificial, el de métodos de aprendizaje automático han vuelto a la física aun mas relevante que nunca, puesto que las potenciales aplicaciones de estas herramientas son variadas y prometen ser una excelente ayuda para que muchos de los problemas actuales sean un problema mucho menos preocupante de lo que es hoy en día.

En general se estima que enfermedades como el cáncer en cualquiera de sus presentaciones suele ser un gran problema en todo tipo de ámbitos, desde personales hasta económicos tanto para las familias como para los gobiernos del mundo. Desde hace unas décadas, la búsqueda por la cura contra esta mortífera enfermedad ha aumentado de manera impresionante, sin embargo, siendo muy poco optimista, la solución no se ve factible en un futuro próximo y aunque los indicios de una cura se presentaran tan pronto como sea posible, existe aun riesgo de que la investigación farmacéutica tome aun varios años para hacerla disponible a todo el mundo y aun con esas, existe la posibilidad de que una gran parte de la población no sea capaz de tener acceso a esta cura, y mientras mucha gente seguirá sufriendo las consecuencias del cáncer.

Por lo tanto, al día de hoy, el método mas factible que encontramos para combatir a esta patología es el de la prevención, pero nuestros métodos de prevención suelen acarrear problemas significativos, en primer lugar por que se requiere un alto grado

de especialización médica para reconocer los signos una vez que los estudios han sido realizados y por otro lado, que existen errores humanos que pueden tener un impacto significativo y que en algunos casos marca la diferencia entre que una persona pueda ser tratada a tiempo y sanar, o dejar evolucionar la enfermedad y tener la posibilidad de complicaciones a futuro.

En este sentido, el propósito de la tesis que se presenta a continuación es la de explorar el comportamiento de los tumores benignos para ofrecer algún factor de caracterización para que este pueda ser considerado y usado en futuras investigaciones, ya sean propias o para futuros investigadores.

Además se pretende dar un repaso a cerca de algunas consideraciones matemáticas sobre el análisis de estructuras para posteriormente presentar un resultado en el análisis de imágenes mamográficas como estudio particular de anomalías benignas, que posteriormente podría ser extendido para analizar afecciones malignas.

## 2 Objetivos

### 2.1. Objetivo General

Estudiar el comportamiento de la métrica Euclideana en tumores benignos detectados en imágenes mamográficas para conocer el comportamiento de estas, para realizar una revisión a cerca de si pueden ser usadas como un método de caracterización.

## 2.2. Objetivos Específicos

Con la investigación que se presenta a continuación, se pretende poner a prueba el análisis de la métrica euclídea como un factor de decisión para la detección de anomalías mamarias, centradas en las mamografías en el campo de la física médica, que puede servir de forma alternativa o complementaria a otros métodos que existen en la actualidad para poder ayudar a clasificar las imágenes médicas de una forma más eficiente y de esta forma ayudar a los profesionales de la salud a realizar diagnósticos más fiables, rápidos y seguros.

El objetivo primordial es el de lograr una fiabilidad alta al momento de realizar la detección para que posteriormente, con ayuda de herramientas más sofisticadas lograr incluso predecir la aparición de alguna clase de malformación; por el momento el estudio se centra en las mamografías por ser un área que contiene en su mayoría tejidos blandos y puede servir de ventana de prueba, pero que podría también extenderse a el análisis en las imágenes de otras zonas del cuerpo.

Así mismo, se pretende realizar una revisión de algunos conceptos matemáticos y físicos que nos puedan ayudar a comprender de una forma sencilla y metodológica el proceso por el cual se realiza el método y su funcionamiento en general para detectar problemas al llevarlo a cabo y poder mejorarlo sistemáticamente una vez completado.

## 3 Introducción

Hasta el siglo XIX existía un gran desconocimiento a cerca del mundo físico que nos rodeaba, la falta de una teoría que pusiera en orden la observación empírica de los fenómenos naturales volvía difícil que se produjeran avances significativos en campos que hoy presentan una gran importancia como todas aquellas en las que esta involucrado el electromagnetismo.

La unificación de las leyes electromagnéticas propuestas por James Clerk Maxwell en 1856 condujeron a que se diera un salto tecnológico en estas cuestiones, produciendo además de una gran revolución en el el campo de la industria eléctrica, un salto tecnológico e importante el el campo de la óptica, la interferometría y mas importante, en el campo de la medicina con la diversa cantidad de invenciones tecnológicas basadas en dicho marco teórico.

Hasta 1895 se desconocía gran parte de lo que hoy consideramos el espectro electromagnético, tanto la radiación 'infrarroja'(ondas de radio, microondas,etc.) usadas principalmente en las telecomunicaciones, como la radiación 'ultravioleta'(rayos x, rayos gamma,etc.) usada ampliamente en el campo de la medicina, la industria nuclear,entre otros, era muy poco estudiada, sin embargo, con el descubrimiento de los rayos x, se encontraron una buena cantidad de aplicaciones, como lo pueden ser las radiografías, ya que podríamos a partir de ese momento, permitirnos explorar el interior del cuerpo humano sin necesidad de intervenciones riesgosas, así pues, los 'rayos x', probaron ser tan efectivos que aun hoy en día y a pesar de los avances tecnológicos, las imágenes obtenidas de esta manera siguen siendo de gran utilidad especialmente por ser relativamente accesibles en cualquier parte del mundo y por que existe un entendimiento bastante solido a cerca de su comportamiento cuando se habla de interacciones con materiales de diferentes densidades.

En la figura **Figura 3.1** Se muestra la primer radiografía de la historia, a partir de entonces, el uso de los rayos X se extendió al rededor del mundo y se estableció como un método confiable para la obtención de imágenes médicas.

Se ha estudiado la interacción de este tipo de radiación con diferentes tipos de tejido, de donde inmediatamente se pueden obtener algunas conclusiones, y es que este tipo de radiación atraviesa con mucha facilidad los tejidos blandos pero muy poco los tejidos que contienen estructuras solidas, además de que es fácil distinguir estructuras con una gran cantidad de detalles, por lo tanto es perfecta para que se



Figura 3.1: Primer Radiografía de la Historia  
*En ella podemos notar la interaccion  
de los los rayos x, con distintos tipos de tejido*

puedan detectar malformaciones especialmente en tejidos blandos (como en las mamas) en donde dichas estructura son fácilmente reconocibles.

Se estima que cerca de 1 de cada 12 mujeres en el mundo, sufrirá a lo largo de su vida, de algún tipo de afección causada por el cáncer de Mama, según datos de la Organización Mundial de la Salud (OMS) [1] por lo cual es de suma importancia, por lo menos en el campo de la prevención, tener a la mano las herramientas necesarias para que el diagnostico de esta enfermedad pueda ser llevado a cabo de forma sencilla, eficaz y con un buen margen de confiabilidad.

Las mamografías son una herramienta clave en la detección temprana y el diagnóstico de enfermedades mamarias, en particular, el cáncer de mama. La detección temprana de esta enfermedad es fundamental para aumentar las tasas de supervivencia y mejorar el pronóstico de las pacientes.

Sin embargo, el análisis puede resultar un desafío debido a la complejidad de analizar las imágenes y la necesidad de identificar características sutiles así como potenciales lesiones para poder estandarizar indicadores que puedan ser usados en el campo de la medicina, para realizar diagnósticos mas confiables y precisos.

Debido a esto, en los últimos años y con el avance de la tecnología se a decidido implementar ciertas técnicas con este objetivo, el de detectar de forma temprana posibles afecciones que puedan existir y que de alguna u otra forma, en comparación con un médico especializado, suelen tener mejor efectividad, aunque se debe reconocer que estas cuestiones necesitan seguir siendo mejoradas.

En esta tesis revisaremos el comportamiento de algún indicador para observar la forma en la que nos pueda ayudar a facilitar la detección y caracterización de las anomalías mamarias, nos enfocaremos primordialmente en el estudio de los tumores benignos dado que son un excelente método de prueba dado que son fácilmente reconocibles y que pueden darnos pista de como continuar mas adelante o darle las herramientas a futuros investigadores que puedan completar las ideas que se presentan a continuación.

Desde hace algunos los investigadores han notado que el uso de la tecnología pue-

de ser beneficioso para este tipo de estudios en el existen patrones que pueden ser explotados, por lo que es importante definir características, en este caso, dentro de las mamografías para que en combinación con técnicas mas avanzadas, se puedan ofrecer resultados mejores y mas precisos.



# 4 Anormalidades mamarias

## 4.1. Tumores y métodos de detección

La salud mamaria es un tema de gran importancia para las mujeres y, en ocasiones, para los hombres.

Los tumores en el tejido mamario, en particular, son motivo de preocupación y un tema de gran relevancia en el campo de la medicina.

En esta exploración, abordaremos algunas características generales sobre los padecimientos mamarios.

Los tumores en el tejido mamario pueden variar ampliamente en su naturaleza, desde benignos hasta malignos.

Segun el Instituto Mexicano del Seguro Social, tan solo en 2020 el cancer de mama fue la enfermedad mas frecuente (15.3%), catalogada como la principal causa de muerte en México, ademas de que la cifras al rededor del mundo incluso en paises desarrollados suelen ser elevadas, lo cual significa que millones de personas en todo el mundo se enfrentan a la posibilidad de desarrollar un tumor en la mama en algun momento de su vida. [2]. Cuyas afecciones en muchos casos representan una severa afección para la calidad de vida de las personas que llegan a padecerlas, causando desde molestias leves, mal funcionamiento de algunos organos, hasta consecuencias mas severas como la muerte.

Exploraremos algunas características de los diferentes tipos de tumores mamarios, los métodos de diagnóstico y algunos inconvenientes que estos presentan al momento de ser llevados a cabo y por que es importante proponer nuevos metodos alternativos que puedan ser mas eficientes y asequibles para mas personas alrededor del mundo.

Es fundamental comprender la diversidad y complejidad de los tumores en el tejido mamario para brindar una atención médica efectiva y mejorar la calidad de vida de quienes se ven afectados. Aunque en esta revision nos centraremos mas las metodologias de la deteccion mamaria en cualquier etapa para despues proponer alguna forma en la que los metodos descritos aqui puedan ayudar en la deteccion temprana.

A continuacion se describen algunas de las características de los tumores, tanto benignos como malignos, para familiarizarnos con algunas de las características de estas anomalias, asi como las repercusiones que pueden tener en el cuerpo humano y una vision general sin entrar en detalles.

### 4.1.1. Tumores benignos

Un tumor benigno es un crecimiento anormal de células que se encuentra en un área específica del cuerpo y generalmente no representa una amenaza para la vida. A diferencia de los tumores malignos (cáncer), los tumores benignos no tienen la capacidad de invadir tejidos circundantes ni de propagarse a otras partes del cuerpo a través del torrente sanguíneo o el sistema linfático.

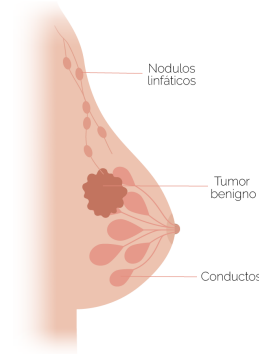


Figura 4.1: Esquema de tumor benigno  
*Suelen presentar estructuras bien definidas  
 y no suelen presentar un peligro para la salud*

En la **Figura 4.1** se muestra un esquema de tumor benigno, en el cual resaltan algunas características de la mama y en especial la geometría regular de este tipo específico de anomalías.

Estos tumores suelen ser encapsulados, lo que significa que están rodeados por una membrana o barrera de tejido conectivo. Además, tienden a crecer de manera lenta y controlada. A menudo, los tumores benignos se descubren de manera incidental durante exámenes médicos de rutina o se identifican debido a síntomas locales que pueden estar relacionados con su crecimiento, como dolor o presión en el área afectada.

En la mayoría de los casos, los tumores benignos se pueden tratar de manera efectiva mediante la extirpación quirúrgica, y rara vez reaparecen después de la extracción. Aunque generalmente no son una amenaza para la vida, algunos tumores benignos pueden causar síntomas o problemas de salud si crecen lo suficiente y ejercen presión sobre órganos o tejidos circundantes.

Sin embargo el motivo principal por el que se muestran relevantes en esta investigación es que este tipo de anomalías presenta unas características excelentes para que algunas investigaciones puedan ser llevadas a cabo sobre ellas. La mayor parte de las anomalías benignas suelen tener geometrías circulares que las vuelven fácilmente reconocibles con los distintos tipos de métodos de detección que existen, de esta forma, la utilidad en este estudio es el de poder reconocer patrones en el análisis que se realizara puesto que estudiar tumores malignos suele conllevar un grado algo más alto de dificultad por lo general.

### 4.1.2. Tumores malignos

Los tumores malignos, a menudo conocidos como cáncer, son crecimientos anormales de células que tienen la capacidad de invadir tejidos cercanos y propagarse a otras partes del cuerpo a través del torrente sanguíneo o el sistema linfático.

Estos tumores representan una amenaza significativa para la salud, ya que pueden interferir con el funcionamiento normal de los órganos y tejidos. Además, pueden causar síntomas como dolor, pérdida de peso inexplicada, fatiga y cambios en la piel.

Los tumores malignos a diferencia de los tumores benignos se producen debido al crecimiento descontrolado de distintos tipos de células dentro de la mama y presentan geometrías más bien anormales poco convencionales, dado que crecen sobre cualquier clase de tejido en la mama, sin embargo dada su naturaleza, al momento de analizar este tipo de anomalías en las mamas, la principal fuente de error es la de detectar sobre qué zonas de la mama se ha esparcido el tumor.

En la **Figura 4.2** se puede observar un esquema de los tumores benignos en donde se resalta la irregularidad geométrica que presentan este tipo de afecciones dentro de la mama.

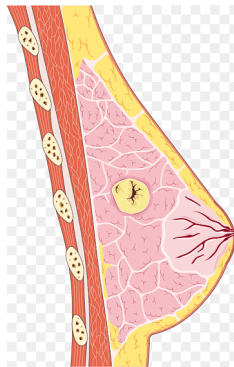


Figura 4.2: Estructura de tumor maligno

*Presentan formas irregulares, presentan un riesgo para la salud dada su capacidad de crecer descontroladamente y moverse a distintas zonas del cuerpo*

El tipo de cáncer de mama que se presente depende del tipo de tumores que se hayan vuelto cancerosos, sin embargo los tipos de cáncer de mama más comunes son:

- *Carcinoma ductal infiltrante*: Las células cancerosas en este caso se generan en los conductos mamarios [3]
- *Carcinoma lobulillar infiltrante*: Las células cancerosas se originan en los lobulillos [3]

Aunque cabe mencionar que estos no son los únicos tipos de afecciones, aunque sí representan a la mayoría de los casos.

Finalmente, cabe mencionar que existen ciertos factores que incrementan la posibilidad de que pueda existir una afección en forma de tumor mamario, como el género, la edad, factores genéticos y de densidad del tejido, mientras que por otro lado factores asociados al estilo de vida, como el sobrepeso incrementan la posibilidad de padecer cáncer de mama.

## 4.2. Metodos de deteccion

La medicina moderna se beneficia enormemente de los avances en tecnología de imagen médica, permitiendo a los profesionales de la salud explorar y evaluar de manera no invasiva el interior del cuerpo humano.

Los métodos de análisis de imágenes médicas son fundamentales en el diagnóstico, la planificación del tratamiento y el seguimiento de diversas afecciones.

Desde la radiografía hasta la resonancia magnética y la tomografía computarizada, estas técnicas proporcionan una ventana a los tejidos y órganos internos, ayudando a los médicos a comprender mejor la salud de sus pacientes.

En esta exploración, revisaremos algunos de los métodos de análisis de imágenes médicas.

Exploraremos cómo funcionan estas tecnologías, sus aplicaciones en el campo de la medicina, especialmente en el campo de las detecciones de anomalías mamarias.

Desde la mamografía para la detección temprana del cáncer de mama hasta la ecografía para seguimiento del desarrollo de ciertas afecciones comunes, estas técnicas están transformando la atención médica y brindando un mayor entendimiento de la anatomía y la función del cuerpo humano.

Analizaremos además, cómo los métodos de análisis de imágenes médicas se han convertido en herramientas esenciales en la práctica clínica moderna y en la búsqueda de una atención médica más precisa y efectiva.

Pero además algunas consideraciones adicionales sobre estas, como sus contrapartes negativas y por que es factible buscar en la medida de lo posible buscar metodos eficientes que puedan reducir al minimo los inconvenientes que las tecnicas actuales generan.

### 4.2.1. Ultrasonido

La ecografía mamaria utiliza ondas sonoras para crear imágenes en tiempo real de los tejidos internos de la mama. Estas ondas sonoras rebotan en las estructuras mamarias y se traducen en imágenes en un monitor.

Se utiliza como complemento a la mamografía y a veces a la resonancia magnética. Puede ser especialmente útil en casos de evaluación de masas palpables, en mujeres jóvenes con tejido mamario denso y en situaciones donde se necesita mayor claridad sobre la naturaleza de una lesión detectada en otras pruebas.

La ecografía mamaria es útil para evaluar lesiones que no son visibles en una mamografía, como tumores pequeños o quistes. También es segura y no involucra radiación ionizante.

En la **Figura 4.5** se muestra una imagen obtenida mediante el uso de la ecografía, en contraposición con los demás métodos expuestos, en donde se pueden apreciar de primera mano sus ventajas y desventajas. **Limitaciones y Consideraciones :**

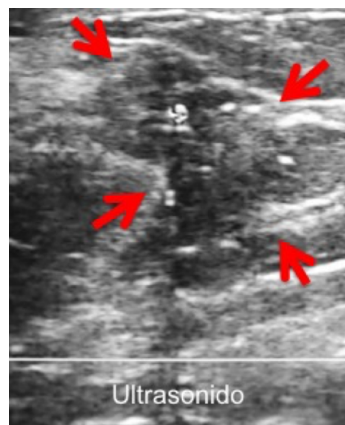


Figura 4.3: Ecografía mamaria

*En la imagen se pueden apreciar el nivel de resolución de las ecografías, en donde además se muestran distintas anomalías detectadas*

La ecografía mamaria puede ayudar a diferenciar entre tumores sólidos y quísticos, y proporciona información sobre la vascularización de las lesiones. Puede ser particularmente útil para guiar procedimientos de biopsia.

**Limitaciones y Consideraciones :** Aunque la ecografía mamaria es valiosa, tiene algunas limitaciones. Puede generar más falsos positivos y falsos negativos en comparación con otras pruebas, y su eficacia puede variar según el operador y la calidad del equipo.

Se utiliza a menudo en mujeres jóvenes con tejido mamario denso, donde la mamografía puede ser menos eficaz. También se puede emplear durante el seguimiento

de mujeres con alto riesgo de cáncer de mama. Se utiliza en combinación con mamografías y, en algunos casos, con resonancia magnética, para obtener una evaluación completa de la salud mamaria.

En resumen, la ecografía mamaria es una herramienta valiosa en la detección y evaluación de tumores mamarios, especialmente en situaciones específicas y en grupos de mujeres donde otras pruebas pueden tener limitaciones. Su uso se adapta a casos particulares y a menudo se integra con otras modalidades de imagen para una evaluación más completa.

### 4.2.2. Mamografía

Las mamografías son una herramienta crucial en la detección temprana del cáncer de mama, una enfermedad que afecta a millones de mujeres en todo el mundo. Este procedimiento de imagen médica utiliza rayos X de baja dosis para obtener imágenes detalladas de los tejidos mamarios. A lo largo de las décadas, las mamografías han demostrado ser efectivas en la detección precoz del cáncer de mama, lo que aumenta significativamente las posibilidades de tratamiento exitoso y supervivencia.

La mamografía implica la compresión suave pero firme del seno entre dos placas de rayos X. Esta compresión asegura que el tejido mamario se extienda uniformemente, lo que permite obtener imágenes más claras y reduce la dosis de radiación necesaria. Las imágenes resultantes se llaman mamogramas y se interpretan por radiólogos especializados en imágenes de mama.

**Detección Temprana :** La principal ventaja de las mamografías es su capacidad para detectar pequeñas anomalías en los senos antes de que sean palpables. Esto permite el diagnóstico temprano y el tratamiento oportuno del cáncer de mama, cuando las tasas de supervivencia son más altas.

**Reducción de Mortalidad :** Numerosos estudios han demostrado que las mamografías regulares en mujeres mayores de 40 años pueden reducir la mortalidad por cáncer de mama. La detección temprana permite tratamientos menos invasivos y más efectivos.

**Evaluación de Anomalías :** Además de la detección del cáncer, las mamografías también pueden utilizarse para evaluar anomalías en los senos, como quistes o masas benignas, proporcionando información esencial para el diagnóstico y tratamiento adecuado.

Las pautas varían según las organizaciones de salud, pero generalmente se recomienda que las mujeres comiencen a realizarse mamografías de detección a partir de los 40 o 50 años, dependiendo de los factores de riesgo individuales. Las mamografías de detección suelen realizarse cada uno o dos años, aunque pueden ser más frecuentes en caso de antecedentes familiares de cáncer de mama u otros factores de riesgo.

A pesar de sus beneficios, las mamografías no están exentas de controversia. Algunos críticos argumentan que pueden dar lugar a falsos positivos (resultados anormales que no se traducen en cáncer) y falsos negativos (cánceres que se pasan por alto). Además, existe preocupación sobre la exposición a la radiación, especialmente en mujeres jóvenes. Sin embargo, los beneficios de la detección temprana generalmente superan los riesgos en la población adecuada.

En la **Figura 4.3** se puede observar una mamografía, en la cual se puede notar la forma en la que la radiación interactúa con distintos tipos de tejidos de la mama,

pero ademas, la forma en la que ciertos tipos de anomalias son detectadas.

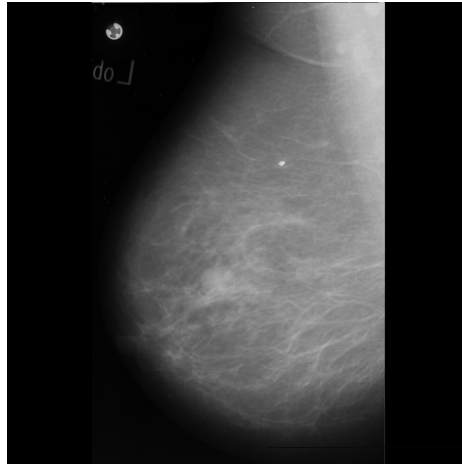


Figura 4.4: Mamografía mamaria

*Se muestra la deteccion de una microcalcificacion en una imagen mamografica*

En conclusión, las mamografías desempeñan un papel esencial en la detección temprana y el tratamiento exitoso del cáncer de mama. Es importante que las mujeres se informen sobre las recomendaciones de detección y hablen con sus médicos para determinar el plan de cribado adecuado para su situación individual. La detección temprana puede salvar vidas y mejorar la calidad de vida de las pacientes con cáncer de mama.

### 4.2.3. Resonancia magnetica

Junto a las mamografías, la resonancia magnetica puede ser usada para detectar anomalías mamarias mediante la producción de imágenes que son de gran ayuda para el diagnóstico.

La resonancia magnética utiliza campos magnéticos y ondas de radio para crear imágenes detalladas de los tejidos internos del cuerpo. En el caso de la detección de tumores mamarios, la RM puede proporcionar imágenes más detalladas y precisas que otras modalidades de imagen, como la mamografía o la ecografía.

En la **Figura 4.4** podemos observar la resolución de las imágenes obtenidas mediante el uso de las resonancias magnéticas, a la vez en que podemos observar el detalle con el que se muestran los distintos tipos de tejido de la mama así como las anomalías mamarias.

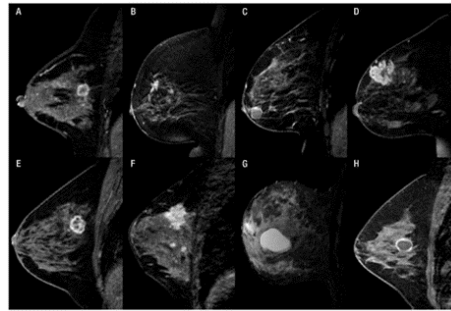


Figura 4.5: Imagen obtenida mediante resonancia magnetica  
*Se muestran distintos tipos de mamas, algunas con anomalías obtenidas con el método de la resonancia magnetica*

Se utiliza en situaciones específicas, como en mujeres con alto riesgo de cáncer de mama, para evaluar implantes mamarios, para evaluar la extensión del cáncer de mama ya diagnosticado y en casos de detección de lesiones no detectadas por otras pruebas.

**Mejora en la Sensibilidad** : La RM mamaria es particularmente sensible para detectar tumores en mujeres con tejido mamario denso, donde la mamografía puede ser menos efectiva. Además, puede identificar lesiones más pequeñas que podrían pasar desapercibidas en otras pruebas.

**Características de la Imagen** : Puede además proporcionar imágenes detalladas de los vasos sanguíneos, lo que ayuda a diferenciar entre tumores malignos y benignos. También puede ayudar en la detección de múltiples tumores y evaluar la extensión del cáncer en casos ya diagnosticados.

**Limitaciones y Consideraciones** : Aunque la RM mamaria es muy sensible, también puede dar lugar a falsos positivos. Además, no es la primera línea de detección

para todas las mujeres debido a sus costos y la baja disponibilidad de este tipo de equipos en los hospitales.

En resumen, la resonancia magnética mamaria es una herramienta valiosa en la detección y evaluación de tumores mamarios, especialmente en situaciones específicas donde otras pruebas pueden no ser tan efectivas. Sin embargo, su uso suele estar indicado en casos particulares y no como una prueba de rutina para todas las mujeres.



# 5 Segmentacion de imagen

Ahora que se ha realizado una pequeña revision a cerca de algunos de los metodos con las que las imagenes son procesadas para la posterior deteccion caracterizacion de tumores, es importante, antes de introducirnos al estudio de las variedades diferenciables, revisar cuales han sido hasta el momento algunas de los metodos usados para esta tarea.

## 5.0.1. Segmentacion de imagen

En un articulo publicado por la Sociedad Venezolana de Farmacologia Clinica y Terapeutica, publicado en 2021, "Diagnostico y caracterizacion del cancer mamario". [4]

Se presenta una recopilacion de investigaciones sobre el estado de las investigaciones en el campo de las anomalias mamarias, en el contexto de el campo medico clinico pero tambien un breve resumen a cerca de las investigaciones y el estado actual de los metodos de deteccion y caracterizacion de las anomalias mamarias. [4] Las investigaciones en el campo de la imageneologia se extiende a metodos de deteccion en ultrasonidos, resonancia magnetica y mamografias, sin embargo, dadas las caracteristicas de la tesis presente, se estudiara unicamente aquellas investigaciones en las que los desarrollos se hayan dado usando imagenes mamograficas.

Para analizar una imagen se usan distintos métodos de segmentación que se complementan entre si, entre los que se encuentran metodos basados en discontinuidad y basados en similitudes. [5]

- *Basado en discontinuidad* : En donde podemos encontrar metodos de Thresholding y metodos basados en segmentacion de bordes.
- *Basado en similitud*: Aqui se encuentran los metodos en los que se enfoca en la deteccion de regiones en las que se encuentran los tumores.

Adicionalmente, el estudio titulado 'Segmentation Techniques for Breast Cancer Imaging Modalities- A Review' [6] Se proporciona una revision detallada sobre tecnicas de segmentacion para mejorar el contraste y eliminar informacion no deseada. En ese sentido, se ha observado que los detalles de los senos también son eliminados a veces en el proceso eliminación del músculo pectoral.

Ademas se exponen algunas de las características de las investigaciones llevadas a cabo en la bibliografía usual. Para ello, se categorizan los metodos con los que se trabaja usualmente y se exponen como sigue a continuacion:

- *Segmentacion basada en aproximacion de region*:En este tipo de segmentacion se mide el peso entre las conexiones de los pixeles en una imagen para de esta forma
- *Segmentacion basada en borde*:Usa el reconocimiento entre los bordes de distintas regiones para analizar las zonas de interes [6]
- *Segmentacion basada en Thresholding*: Segmenta las regiones de la mama usando un criterio de seleccion(por ejemplo, discrimina las intensidades mas bajas para los pixeles de una imagen). [6]

De forma que en la **Figura 5.1** podemos encontrar como se establece la clasificacion de los metodos de segmentacion de imagen basados en la forma en la que estos metodos actuan dentro de la imagen

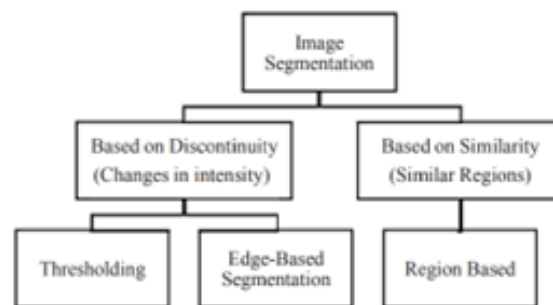


Figura 5.1: Clasificacion de segmentacion de imagen

*En la imagen se expone la segmentacion de los metodos utilizados para el procesamiento de las imagenes*

Cada una de ellas con sus respectivas ventajas y deventajas, en cuyo caso, algunas de las mas grandes ventajas que se enumera es la de la alta resolusion de algunas de las tecnicas, la rapidez y la mejora del contraste en contraposicion con las desventajas, entre las cuales se destaca el costo computacional, la sensibilidad al ruido, la seleccion de la semilla inicial para iniciar el proceso o la incapacidad de aplicar alguna de las tecnicas en concreto, en algunas muestras (por ejemplo, la de las muestras con alta densidad).

## 5.0.2. Estado del arte

A continuacion se muestra una recopilacion de algunos de los articulos en los que se muestran algunos de los tipos de segmentacion de imagen

## Basada en regiones

En el artículo titulado 'Automatic Segmentation of Breast Lesions for Interaction in Ultrasonic Computer-Aided Diagnosis', encontramos un método de segmentación de imágenes basado en regiones, el cual funciona como sigue:

- Se construye una gráfica de la imagen que se analizará y se define el peso de los bordes  $\omega_{ij}$  que conecta a dos píxeles adyacentes  $v_i, v_j$  como:

$$\omega_{ij} = |v_i - v_j|$$

[7] De forma que al principio, cada uno de los píxeles es parte de un subgrafo en sí mismo

- Se ordenan a todos los bordes detectados en orden decreciente y se establece  $q=1$
- Elegimos el  $q$ -ésimo borde de entre los bordes ordenados y si los bordes conectan a dos diferentes subgrafos  $G_i$  y  $G_j$  entonces se toma el CNR (Contrast Noise-Ratio)  $Dif(G_i, G_j)$  y la mínima diferencia interna  $MInt(G_i, G_j)$ , en donde:

$$Dif(G_i, G_j) = |u(G_i) - u(G_j)|$$

Donde  $u(G)$  es el promedio de la intensidad de los píxeles en  $G$ , mientras que por otro lado

$$MInt(G_i, G_j) = \min(Int(G_i) + \lambda(G_i), Int(G_j) + \lambda(G_j))$$

Donde  $INT(G)$  es la diferencia interna que es computada por la desviación estándar de los píxeles en  $G$  y  $\lambda(G)$  es una función de umbralización definida como:

$$\lambda(G) = \frac{k}{|G|} \left(1 + \frac{1}{\alpha * \beta}\right); \beta = \frac{\mu(G)}{\sigma(G)}$$

con  $k, \alpha$  parámetros positivos  $|G|$  el tamaño del subgrafo  $G$ , en donde además se escoge  $k = 2000$  y  $\alpha = 0.02$  de acuerdo a experimentos.

Finalmente, para encontrar las regiones en las que la imagen está separada, se sigue el criterio

$$Dif(G_i, G_j) \leq MInt(G_i, G_j)$$

El criterio se repite para distinta cantidad de subgrupos de la imagen de tal forma que todas las regiones dentro de la imagen son detectadas.

En la **Figura 5.2** podemos observar la forma en la que el método de segmentación por regiones actúa en una imagen en la que se puede observar un tumor, la idea general es la de aislar la anomalía para poder reconocerla fácilmente o en su defecto, caracterizar algunas de las propiedades que esta región posee.

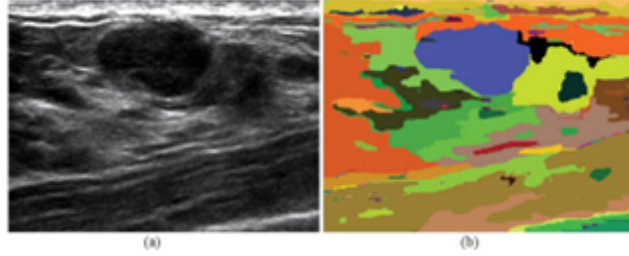


Figura 5.2: Segmentacion por region

*Se muestra un ejemplo de como una imagen es segmentada dependiendo de las características de cada region*

### Basada en Thresholding

Por otro lado, en el artículo titulado 'Detection of masses in mamography breast cancer images using MATH method', se propone un enfoque para detectar eficazmente las masas en imágenes mamográficas de cáncer de mama mediante el uso del método MHAT (Modified Histogram based Adaptive Thresholding).

Este método se basa en la aproximación de píxeles y usa métodos de probabilizar para segmentar regiones. Usando un clasificador

$$G(x, y) = \begin{cases} 1 & \text{si } f(x, y) > T \\ 0 & \text{si } f(x, y) \leq T \end{cases} \quad (5.1)$$

En donde

$$T = T[(x, y), p(x, y), f(x, y)]$$

Cuyo primer elemento  $(x, y)$  es la posición de un píxel, el segundo elemento  $p(x, y)$  es el de alguna propiedad local del punto elegido (por ejemplo el promedio del nivel de grises de los píxeles de alrededor) y el tercer elemento  $f(x, y)$  es el nivel de grises del punto. [8]

De esta forma, podemos definir límites locales o globales de la imagen, modificando los parámetros de la función  $T$ , agregándolos al método o ignorando algunas de sus variables.

De forma tal en la que si tenemos una imagen en la que queremos segmentar dos regiones distintas  $R_1$  (el fondo) y  $R_2$  (la zona del tejido), la región de umbralización vendrá definida como:

$$T = T_1(p(x, y), f(x, y)) + T_2(p(x, y), Ne(x, y))$$

En donde  $Ne$ , es el valor de la intensidad de los vecinos circundantes.

De esta forma están definidos variables globales (en el primer término) al realizar mediciones únicamente sobre la imagen completa sin ningún tipo de segmentación y variables locales (en el segundo término) cuando consideramos la vecindad de un píxel en un momento dado.

Con el fin de retener bordes significativos, es introducido un limite adaptativo

$$T = \mu + c\sigma$$

En donde  $\mu$  es la media y  $\sigma$  la desviación estandar del histograma de la muestra y además 'c' es una constante de normalización.

Finalmente se mide el promedio del peso de las intensidades en el histograma de la imagen procesada y se localizan los puntos mas bajos y mas altos para evitar las irregularidades y mejorar el ajuste de los bordes.

El algoritmo se probó con 100 imágenes mamográficas y los resultados experimentales muestran que el método de detección tiene una sensibilidad del 98.3% a 0.78 falsos positivos con precisión del 99% por imagen. El estudio evaluó el desempeño del MHAT comparado con respecto a imágenes de referencia. [8]

En esta revisión, se presentan algunos estudios y revisiones de técnicas de segmentación de imágenes. A pesar que han sido desarrolladas muchas técnicas, no todas son útiles para todos los tipos de imágenes. La segmentación segmenta la imagen según algunas similitudes, el algoritmo de crecimiento de regiones es clásico para la segmentación dinámica de las imágenes médicas. Esta revisión presenta una base para las técnicas existentes y desarrolla una técnica híbrida como propuesta para trabajos futuros.

En la **Figura 5.3** podemos observar como una imagen es analizada usando un método de thresholding en el que las zonas menos luminosas son elegidas para conservar unicamente aquellas en las que las intensidades de los pixeles son claramente mas oscuras, de forma similar, en el metodo descrito anteriormente se usa un tipo de factor de segmentacion de tal forma que se pueda determinar la region en la que el tumor actua.

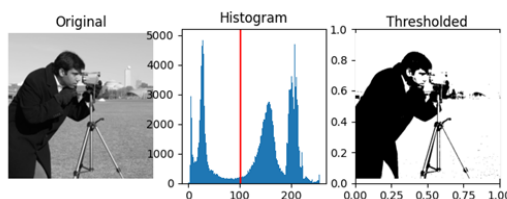


Figura 5.3: Segmentación basada en thresholding

*Se presenta un ejemplo del funcionamiento de los métodos de segmentación de imágenes basadas en thresholding*

En la una gran parte de las investigaciones en las que se analizan los tumores mamarios, se usan distintos tipos de metodos de segmentacion de imagenes, en los cuales son usados distintos tipos de 'factores' para asegurar que los tumores sean detectados de forma adecuada, dado que independientemente de que el metodo elegido para segmentar la imagen sea bueno, siempre existe la posibilidad de que el metodo falle si un solo criterio es utilizado, por lo que hace falta combinar mas de uno de los metodos de clasificacion de anomalias mamarias.

Para esto hace falta explorar que tipo de factores pueden ser considerados dado el metodo con el que se analizan las imagenes, aprovechando las ventajas y desventajas de cada uno de los metodos de imageneologia (ultrasonido, mamografia, resonancia magnetica, etc.)

En el articulo titulado 'Tumor and microcalcification characterization using Entropy, Fractal Dimension and intensity values statistical analysis in mammography' [9] Se expone el uso de la entropia de la imagen y la dimension fractal de tal que se puede caracterizar a las microcalcificaciones dentro de las imagenes mamograficas de tal forma que teniendo en cuenta estos elementos, se pueda usar un metodo de thresholding para aislar a microcalcificaciones , en donde ademas los resultados muestran que la entropia (que es normalmente entendida como el grado de desorden de un sistema ) puede ser aprovechada(Por lo menos en las microcalcificaciones ) dado su alto valor notado en las mediciones, aunque el estudio tambien muestra que existen ciertos problemas al detectactar algunos tipos especificos de tumores, como lo puede ser los tumores denso-glandulares [9]

En el articulo 'Segmentation of breast masses in digital mammograms using adaptive median filters and texture analysis' [10]se expone un proceso con el cual se pretende detectar anomalias mamarias con la ayuda de un filtro adaptativo mediano y un metodo llamado 'CLAHE' basado de ecualizacion del histograma de la imagen para posteriormente eliminar aquellos pixeles que por sus caracteristicas estemos seguros que no forman parte del tumor, para de esta forma eliminar elementos no deseados dentro de la imagen como el musculo pectoral y/o etiquetas dentro de la imagen; el metodo se basa en realizar particiones dentro de la imagen y analizar las texturas de cada una de las particiones de donde ademas se expone que las regiones mas brillantes y ademas las de densidad mas uniforme son aquellas que pueden potencialmente ser una region sospechosa, el resultado final son regiones delimitadas por curvas cerradas de tal forma que la unica dificultad es la de elegir la curva que represente mejor al area de la anomalia. [10]

En el articulo 'Detection and identification of mass structure in digital mammograms' [11] se expone un metodo basado en thresholding con el que se detectan las anomalias basandose en la continuidad de la escala de de grises dentro de la imagen con lo cual se logra detectar anormalidades conservado de manera precisa sus bordes, finalmente una vez detectadas las anomalias, se procede a caracterizar a la region del tumor, obteniendo caracteristicas como el area, perimetro,compacticidad, ancho, alto, factor de isotropia y excentricidad de la masa extraida. [11]

En el artículo 'A novel statical approach for detection of suspicios region in digital mammograms' [12] se presenta un metodo basado en estadistica en el que se mide distintos tipos de entropia de las regiones analizadas y ademas un factor de comparacion con contra el cual se mide la posibilidad de haber captado una anomalia [12]



# 6 Variedades diferenciables

## 6.1. Variedades diferenciables

Las variedades diferenciables se encuentran en una gran cantidad de areas de la fisica, especialmente en teorias como las de la 'relatividad general' en donde nos ayudan a describir deformaciones en el espacio-tiempo, pero ademas, podemos utilizarlas en el campo de la fisica medica para describir propiedades que se encuentran implícitas dentro de las imagenes (en nuestro caso, las imagenes mamograficas) dado que algunos de sus elementos son usados para caracterizar regiones.

Las variedades son la generalizacion de nuestra idea sobre curvas y superficies a objetos de dimensiones arbitrarias.

Por ejemplo una curva en 3 dimensiones se puede parametrizar por un unico numero 't' como:  $C(x(t),y(t),z(t))$  mientras dos numeros u,v,parametrizan una superficie como:  $S(x(u,v),y(u,v),z(u,v))$ , de esta forma una curva y una superficie son homeomorfos a  $R$  y  $R^2$  respectivamente.

Una variedad en general es un espacio topologico el cual es localmente homeomorfo a  $R^n$  aunque puede no serlo globalmente. [13]

Se pueden considerar ciertas aplicaciones que lleven ciertas partes de  $X$  en  $R^n$  por medio de aplicaciones que denominamos 'cartas', . [14]

En ese sentido, una imagen puede entenderse como una matriz  $M \times N$  de tal forma que podemos asignarle a cada punto una coordenada especifica  $(x, y)$ , tal que a cada punto se le puede asignar un parametro que pueda a ayudarnos a comprender el comportamiento de estos pixeles, por ejemplo la funcion de escala de grises  $I(P_i)$  por lo que tenemos entonces una funcion que mapea a la imagen

$$\nu((x, y), I(p_i)) : R^{N \times M} \rightarrow R^{N \times M}$$

De modo intuitivo una superficie en  $R^3$  puede construirse tomando partes de un plano y deformandolo de tal forma que no queden ni aristas ni vortices, pero en general podriamos decir que cualquier superficie n-dimensional puede ser construida a partir de elementos que sean homeomorfos a superficies k-dimensionales, con  $k < n$ . [15]

**Definicion** : Una variedad  $M$  de dimension  $n$ , o  $n$ -variedad es un espacio topologico si sigue las siguientes propiedades: [16]

- $M$  Es un espacio de Hausdorff
- para cualquier punto  $p \in M$  existe un vecindario  $p \in U$  tal que sea homeomorfo a un subconjunto abierto  $V \subset R^m$
- $M$  tiene una base contable de conjuntos abiertos

Un espacio Hausdorff cumple la siguiente definición.

**Propiedad Hausdorff** : Sean  $x$  e  $y$  dos puntos distantes de un espacio metrico,  $(E,d)$ . Existen entonces dos bolas abiertas  $B(x,r)$  y  $B(y,s)$ , tal que: [17]

$$B(x, r) \cap B(y, s) = \emptyset$$

En la **Figura 6.2** se observa el esquema de dos bolas matematicas que no comparten elementos en comun.

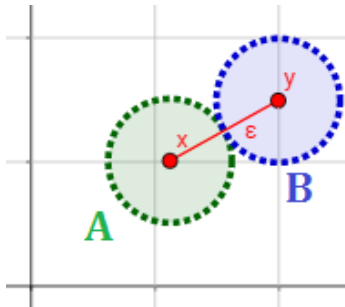


Figura 6.1: Bolas disjuntas

*Dos conjuntos que se mantienen distantes en un espacio métrico, no comparten elementos en común*

Esta eleccion es particularmente necesaria dado que cada una de las 'bolas' necesita contener informacion independiente dado que el caso en el que se encontraran sobrepuestas una con otra, una medicion podria llevar a fallos dentro del analisis.

La segunda propiedad nos dice en palabras sencillas que dado un punto  $P$  en  $M$ , este tiene un vecindario que es homeomorfo a un disco n-dimensional, y finalmente la ultima propiedad nos menciona que  $M$ , puede estar cubierta de muchos de estos vecindarios.

### 6.1.1. Metricas

Intuitivamente una metrica podria entenderse simplemente con la nocion de una medida o distancia entre distintos elementos, algo que nos permite en primera instancia precisar la nocion de proximidad; una idea que esta implicitamente presente en los conceptos fundamenteles de la topologia y el analisis matematico, como en el caso del calculo.

Un caso sencillo que podemos analizar cuando tratamos con este tipo de ideas es el de la distancia entre puntos en la recta real (que todos conocemos), por ejemplo, es sencillo conocer la distancia entre dos puntos: propongase la distancia entre el 3 y el 7

$$\text{dist}(3, 7) = 7 - 3 = 4$$

La idea de la metrica esta aplicada particularmente en este problema en el que unicamente introducimos ideas de la aritmetica(aunque este tipo de conceptos a menudo requieren generalizarse aplicando herramientas matematicas mas complejas).

Sin embargo, esta idea precisa de un par de suposiciones primordiales para funcionar de la forma en la que lo ha hecho como en el ejercicio que hemos resuelto anteriormente y esta es que hemos resuelto una metrica en una dimension y con la forma usual de un espacio euclideo, o de una forma mas general, en el espacio  $\mathbb{R}$  (de los reales) , como lo es la recta real, mientras que a una dimension mayor, la idea de metrica cambia un poco a una idea mas general, para un espacio en el que existan dos dimensiones,  $(x, y)$ ,o en otras palabras  $\mathbb{R}^2$  la metrica ahora se modifica un poco, siendo ahora la forma correctara medir las metricas, como: [17]

$$M = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Podemos notar ademas, que descrita asi, la metrica nos indica un cierto grado de dificultad para movernos a traves del espacio, es decir, si nos movemos en linea recta la metrica tendra el valor mas bajo posible, mientras que si nos movemos a traves de obstaculos, la metrica empezara a tomar valores cada vez mayores.

En la **Figura 6.3** encontramos un esquema de los que representa una metrica, en la que podemos observar que dependiendo de las restricciones que existan dentro del espacio, seria mas 'sencillo' o mas 'complicado' alcanzar un cierto punto dentro del mapa, es decir, es claro que el camino que no posee restricciones (verde) es el mas corto, mientras que los demas 'caminos' poseeran valores mas altos.

A partir de aqui el caso se vuelve algo mas general, y es que para el caso en el que hablamos de dimensiones mas altas, la metrica toma una forma bastante similar a la del espacio en  $\mathbb{R}^2$ ; podemos definir entonces al espacio  $\mathbb{R}^n$  de una manera simple que podamos entender, como el espacio formado por los vectores unitarios  $(x^1, x^2, \dots, x^n)$  siendo la metrica usual para el espacio en  $\mathbb{R}^n$  dada como: [18]

$$M^2 = (x_2^1 - x_1^1)^2 + \dots + (x_2^n - x_1^n)^2$$

Este claramente es un ejemplo de la metrica usual que es comprendida intuitivamente como la nocion de una medida, sin embargo, cabe recalcar que las metricas no siempre

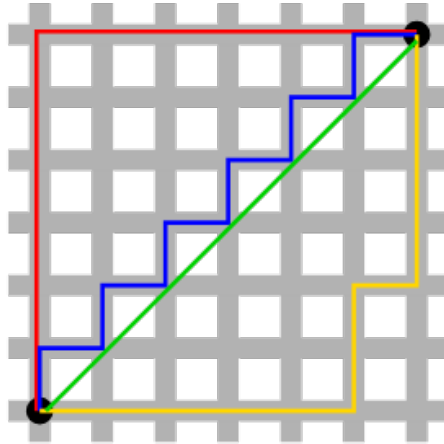


Figura 6.2: Métricas en el espacio  
*Las métricas pueden entenderse como la noción de medida  
entre puntos y puede tomar formas arbitrarias  
siguiendo algunas reglas generales*

tienen por que tener esta misma definición, basta con que sigan una serie de normas que enumeraremos mas adelante, de forma que las métricas de un espacio pueden ser tan contraintuitivas dependiendo de como sean definidas, su importancia principal radica en los espacios matemáticos que puedan ser definidos una vez definida esta característica.

A partir de aquí podemos obtener algunas propiedades esenciales a cerca de los espacios métricos, de donde enumeramos algunas de sus propiedades: [18]

- $M \geq 0$
- $M(x, y) = M(y, x)$
- $M(x, x) = 0$
- $M(x, y) \leq M(x, z) + M(z, y)$

La primera propiedad nos dice que todas las métricas deben ser positivas, esto, dado que en el mundo real, no tiene sentido hablar de distancias negativas.

La segunda propiedad nos dice esencialmente que la métrica es invariante y por lo tanto conmutativa, esto quiere decir que el orden de los factores no altera el producto y por lo tanto, no hay distinción entre medir una distancia de un punto A a un punto B, que medir esta misma propiedad en forma inversa, del punto B al punto A.

La tercera definición en palabras simples expresa de forma sutil que dentro de este espacio, existe el elemento neutro, el cual pensarse como la acción de medir la distancia de un punto a si mismo.

La cuarta propiedad es a menudo conocida como la desigualdad triangular y expresa que la menor "distancia" para recorrer entre dos puntos, es la línea recta. [18]

A partir de aquí podemos obtener una primera definición.

**Definición:** sea  $E$  un conjunto no vacío y  $d$  una distancia definida en él. Al par  $(E, d)$  se le denomina espacio métrico. [18]

De esta forma  $(\mathbb{R}^n, d_n)$  donde consideramos a  $d_n$  como la distancia Euclídea, es un espacio métrico, además de que podemos identificar en general, que cualquier espacio de la forma  $(\mathbb{R}^{n-k}, d_{n-k})$  es un subespacio de  $(\mathbb{R}^n, d_n)$  [17]

Esta definición será de ayuda más adelante para comprender algunas de las implicaciones que tienen los espacios métricos en la investigación desarrollada en este artículo. Los elementos de un espacio métrico pueden ser de naturaleza arbitraria, tanto como números, puntos, vectores, matrices, funciones, conjuntos, entre otros. De forma equivalente, las métricas pueden estar definidas de forma arbitraria dependiendo de la situación y el tipo de objetos matemáticos con los que se está trabajando, siempre y cuando se cumplan las propiedades anteriormente definidas [18]

Las definidas anteriormente son algunas propiedades básicas que debe poseer una métrica, sin embargo por sí sola es un tanto irrelevante, hace falta entonces el uso complementario de otro tipo de herramientas para un mejor entendimiento de la utilidad potencial que posee estas (además de otras) características que podemos obtener del análisis.

### 6.1.2. variedades diferenciables y metricas

ahora falta hacer una conexion entre los espacios metricos y las variedades diferenciables, esto se logra con el siguiente teorema: [19]

**Teorema** : Sea  $M$  una variedad diferenciable Hausdorff, entonces:

- $M$  admite particiones de la unidad
- $M$  es paracompacta
- $M$  es metrizable

Estamos seguros de que tratamos con una variedad ya que consideramos que existe un espacio metrico asociado a la imagen y ademas que la variedad es Hausdorff (esta propiedad nos dice que cuales quiera dos bolas arbitrarias son disjuntas) puesto que cualquier espacio metrico es Hausdorff y vemos ademas que los detalles sobre las metricas se pueden extender a cualquier region de la imagen sin ningun tipo de restriccion. [19]

Esto es, dicho vagamente. Una vez que una unidad de medida es elegida, el espacio Euclideo  $E^2$  se transforma en un espacio metrico. Que cumple las propiedades con las que hemos estado trabajando, es decir, es Hausdorff y tiene una base contable de conjuntos abiertos, la eleccion de un origen y ejes mutuamente perpendiculares, establecen un homeomorfismo  $\psi : E^2 \rightarrow R^2$ . [16]

Y por tanto estructura de variedad diferenciable, de donde podemos apoyarnos extrapolando algunas de estas propiedades a la zona elegida de los tumores y por lo tanto la consideramos a dicha zona como localmente Euclidea.

Podemos notar que estas características definidas arriba dan lugar a que las imágenes mamograficas sean metrizable ya que pueden entenderse como una variedad diferenciable que cumple la propiedad Hausdorff. De esta forma, las propiedades definidas en el estudio de las metricas y los espacios metricos toman sentido

Finalmente, podemos el proceso de analisis de imagenes viene definido por una serie de transformaciones a lo largo del proceso que siguen el siguiente esquema:

En la **Figura 6.4** se observa el proceso que se sigue para analizar una imagen mamografica, desde la obtencion de la imagen, hasta el indicador o factor que caracteriza o puede caracterizar a las anomalias dentro de la imagen mamografica. de donde

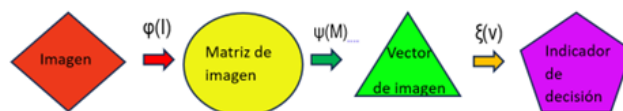


Figura 6.3: Proceso de análisis de imágenes mamográficas

ademas tomamos varios factores que al analizar, nos ayudan a comprender mejor el comportamiento de los tumores benignos



## 7 Metodología

Los datos usados para el análisis llevado a cabo en la investigación que se presenta a continuación fueron descargados de la pagina 'the mini-Mias database of mammograms'

En la **Figura 7.1** se muestra la pagina de donde se obtuvieron las imágenes mamograficas que posteriormente serian analizadas.

The mini-MIAS database of mammograms

Figura 7.1: Pagina del dataset  
*Base de datos de donde se tomaron las mamografias*

Posteriormente se llevo a cabo el análisis de las imágenes que serian aptas para realizar la investigación propuesta, basándonos en la información provista por el dataset, la cual contenía diversas muestras de mamografias con distintos tipos de patologías, con el objetivo de estudiar el comportamiento de la métrica Euclídea para caracterizar distintas regiones elegidas.

En la **Figura 7.1** se muestra la información de la imagen que contiene el dataset y a su vez algunas características de las anomalías detectadas en ellos.

La tabla anterior nos muestra características de las imágenes mamograficas contenidas en el dataset.

Cada una de las columnas de la tabla corresponde con la siguiente información respectivamente.

1. Código de referencia de la imagen
2. Característica del tejido de fondo
  - F Tejido Graso
  - G Graso-Glandular
  - D Denso-Glandular
3. Clase de anormalidad presente
  - CALC Calcificacion

mdb001	G	CIRC	B	535	425	197
mdb002	G	CIRC	B	522	280	69
mdb003	D	NORM				
mdb004	D	NORM				
mdb005	F	CIRC	B	477	133	30
mdb005	F	CIRC	B	500	168	26
mdb006	F	NORM				
mdb007	G	NORM				
mdb008	G	NORM				
mdb009	F	NORM				
mdb010	F	CIRC	B	525	425	33
mdb011	F	NORM				
mdb012	F	CIRC	B	471	458	40
mdb013	G	MISC	B	667	365	31
mdb014	G	NORM				
mdb015	G	CIRC	B	595	864	68
mdb016	G	NORM				
mdb017	G	CIRC	B	547	573	48

Figura 7.2: Datos de las imágenes

*Cada dato incluía algunos datos generales de las anomalías detectadas en las mamografías junto con información adicional*

- CIRC Masa circunscrita
  - SPIC Masa espiculada
  - MISC Otra
  - ARCH Distorcion arquitectural
  - ASYM Asimetrica
  - NORM Normal
4. Severidad de la anormalidad
    - B Benigno
    - M Maligno
  5. Coordenada x del centro del tumor en la imagen
  6. Coordenada y del centro del tumor en la imagen
  7. Radio aproximado de un circulo en píxeles encerrando la anormalidad

Elegimos aquellas muestras en las que las anomalías fueran benignas, esto debido a la geometría particular que presentan, siendo que estas presentan formas mayormente circulares muy bien definidas que nos pueden ayudar a observar como funciona el modelo para que podamos obtener información necesaria y relevante que en dado caso, nos ayuden a extender los propósitos de la investigación llevada a cabo a otros tipos de anomalías como pueden ser los tumores malignos, que presentan formas mas irregulares y complejas.

Posteriormente usando el software CERN ROOT versión 5.34/38 basado en C++, se analizo cada una de las imágenes.

Las imágenes poseen unas dimensiones de  $1024 \times 1024$  píxeles sobre las cuales se plasma un plano coordenado y se usan las coordenadas proporcionadas por el dataset para localizar el centro de la anomalía y además se usa el radio para delimitar la zona cubierta por el tumor

En la **Figura 7.3** se muestra un ejemplo de una mamografía contenida en el dataset en la que se muestra una anomalía que puede ser fácilmente detectable dadas distintas características de la anomalía, como la geometría, la densidad, entre otras cuestiones.

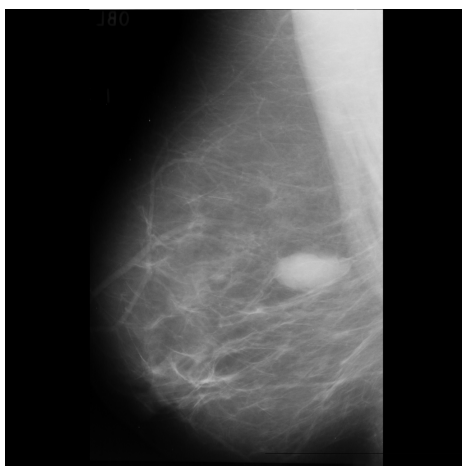


Figura 7.3: Mamografía

*Se muestra una mamografía contenida en el dataset en el que se puede notar un tumor benigno formado*

Algunos de los datos contenidos en el dataset se muestran en la **Tabla 7.1**

Dato(Mdb)	Eje X	Eje Y	Radio
001	535	425	197
002	522	280	69
005	477	133	30
010	525	425	33
012	471	458	40

Tabla 7.1: Datos del dataset

*Encontramos las coordenadas 'X', 'Y' y el radio 'R' de cada una de los tumores en las imágenes.*

*El resto de los datos se encuentra contenido en el apéndice A*

En el que se describen las claves de los datos seleccionados, las coordenadas 'x' e 'y' del centro del tumor y el radio 'r' del mismo.

De esta forma localizamos a los tumores dentro de las imágenes.

En la **Figura 7.4** se muestra la forma en la que los tumores en esta tesis, son encerrados para poder analizar única y específicamente dicha zona, en contraposición con una zona similar de tejido sano.

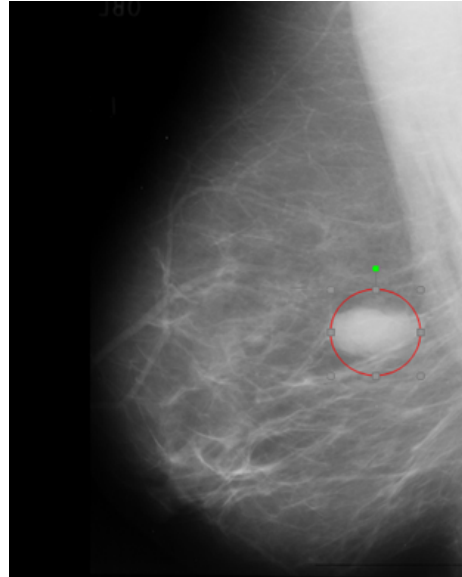


Figura 7.4: zona analizada del tumor

*Se muestra un ejemplo de como las areas de los tumores fueron seleccionadas*

Una vez localizado el tumor se analiza píxel por píxel únicamente a los píxeles contenidos en la zona anteriormente delimitada.

Se estableció una escala de intensidades de entre cero y uno para clasificar las intensidades lumínicas de cada uno de los píxeles

En la **Figura 7.5** observamos un esquema en el que se ha elegido una escala de grises para transformar el brillo a cada uno de los píxeles de la imagen en valores numéricos.



Figura 7.5: Escala de tonos

*Aquí se muestra el criterio con el que se analizaron las imágenes para convertirlas en datos que pudieran ser analizados fácilmente*

El objetivo fue el de analizar cada una de las intensidades de los píxeles individuales para poder transformar la imagen a una matriz de intensidades que pudiera ser fácilmente analizada.

En la **Figura 7.6** se muestra la transformación de imagen a la matriz de píxeles dada la condición expuesta arriba en la que se usa un escala de grises para dar lugar a la transformación.

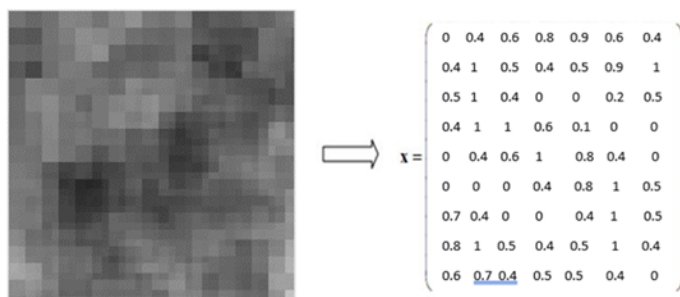


Figura 7.6: De imagen a matriz de datos

*Bosquejo de como se transformo a cada uno de las imágenes en su matriz de intensidades*

Una vez hecho esto, tenemos una matriz  $\mathbf{X} \in R^{N \times M}$  y procedemos a 'aplanar' los datos pasando así a tener un vector  $\vec{V} \in R^{N \times M}$  que podemos analizar de una forma mucho mas sencilla.

Una vez caracterizadas las zonas de los tejidos de los tumores y con el objetivo de caracterizar a las zonas de tejido sano dentro de las mamografías, se realizo exactamente el mismo procedimiento aplicado a zonas arbitrarias dentro de la mama, con la condición de que las zonas analizadas deberían tener exactamente las mismas dimensiones que las de la zona del tumor, es decir, usamos el mismo radio de los tumores de cada imagen para analizar también el tejido sano. De esta forma nos aseguramos de analizar la misma cantidad de datos.

En la **Tabla 7.2** se muestran algunos ejemplos de los datos elegidos para analizar el el tejido sano de las muestras y se presenta a continuación:

La tabla de datos de las zonas analizadas en el tejido sano, se encuentran en el apéndice A.

En donde se describen algunas de las características de las zonas que fueron usadas para realizar el análisis que se describe aquí, se incluyo, al igual que con las zonas de los tumores analizados, la clave con la que se identifican los datos, así como las coordenadas 'X' e 'Y' del tejido sano en el que se realizo el análisis así como el radio de la zona analizada 'R' la cual es la misma que la de los tumores.

De esta forma nos aseguramos de encontrar de forma precisa las zonas de tejido sano a analizar dentro de las imágenes.

Dato(Mdb)	Eje X	Eje Y	Radio
001	575	749	197
002	409	701	69
005	537	569	30
010	398	618	33
012	339	681	40

Tabla 7.2: Datos del tejido sano

*Encontramos las coordenadas 'X', 'Y' y el radio 'R' de cada una de zonas del tejido sano elegidas arbitrariamente.*

*El resto de los datos se encuentra contenido en el apéndice A*

En la **Figura 7.7** se muestra la selección del tejido sano contra la que se comparo los resultados obtenidos de los tejidos de los tumores analizados.

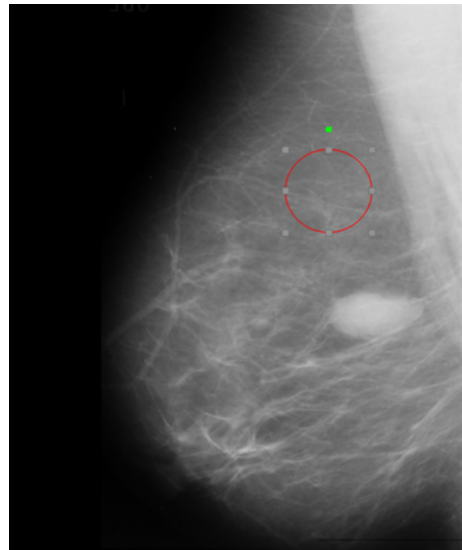


Figura 7.7: Zona analizada del tejido sano

*Se muestra el criterio con el que las zonas sanas de la mama fueron elegidas para su análisis.*

Una vez logrado esto, conseguimos dos vectores, el 'vector de intensidades del tumor ( $\vec{T}$ ) y el 'Vector de intensidades del tejido sano ( $\vec{S}$ )'

Procedemos entonces a obtener las métricas de cada uno de los vectores analizados, donde la métrica usada viene definida como:

$$M(m) = \sum (m(\vec{v}))$$

con

$$m(\vec{v}) = v_i * v_i$$

Obteniendo los datos de las métricas del tumor  $M_T(\vec{m})$  y las métricas del tejido sano  $M_S(\vec{m})$ . Para de esta forma poder realizar comparaciones entre ambos set de datos analizados, usando el indicador

$$P = \frac{M_S(\vec{m})}{M_T(\vec{m})}$$

Es notorio que podrían existir ciertas anomalías en las que las métricas no pudieran distinguir de forma adecuada la diferencia entre el tejido sano y el tejido del tumor, por lo que un análisis adicional es requerido

Este análisis consiste en encontrar ajustes para los histogramas con el que podamos describir generalidades de del tejido sano, como su media y (el rango en el que se presentan) su desviación estandar( $\sigma$ ), para de esta forma poder realizar analisis adicionales en los que podamos distinguir de forma mas clara y concisa que parte del tumor analizado tiene probabilidad de ser tejido sano, debido a que muchos tumores presentan formas mas bien ovaladas y que los métodos de ajuste no son tan rigurosos en en esta investigación.

De esta forma nos aseguramos de que estamos analizando estrictamente al tumor.

Una ventaja adicional que presenta este segundo análisis es que en caso de querer extender el uso de este método al estudio de los tumores malignos, tendríamos rápidamente una forma básica y sencilla de hacer distinciones preliminares de entre el tejido sano y el tejido del tumor, que en este caso serian mas irregulares.



# 8 Resultados y analisis

A continuación de presentan los resultados obtenidos luego de haber aplicado la metodología que se ha descrito anteriormente.

## 8.1. Resultados

Al analizar cada una de las imagenes obtenemos dos sets de datos, de los tejidos del tumor y de las del tejido sano.

De las cuales obtenemos un histograma para cada una de las zonas analizadas dentro de cada imagen.

En la **Figura 8.1** se muestra el histograma obtenido del tumor de una imagen mamografica, en donde podemos observar que la mayor parte de las intensidades de los pixeles tienden a poseer altas intensidades.

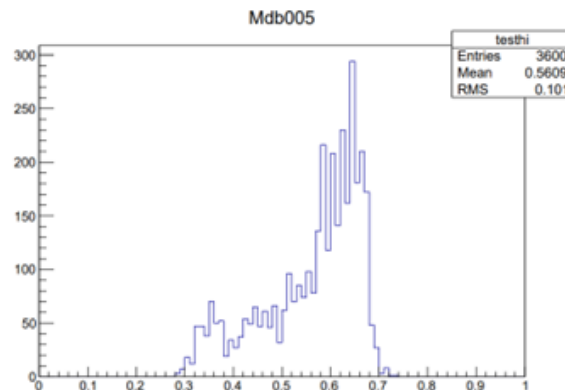


Figura 8.1: Histograma del tejido del tumor

*Se muestra un uno de los histogramas en donde se analiza las intensidades de los pixeles del tejido del tumor*

Así como los histogramas del tejido sano pertenecientes a cada una de las muestras de imagenes analizada.

En la **Figura 8.2** observamos la distribución de las intensidades de los píxeles del tejido sano de la zona seleccionada, en donde podemos observar que la mayor parte

posee intensidades mas bien bajas, sin embargo, se presenta de forma uniforme tal que todas o casi todas las observaciones se quedan confinadas en un rango especifico de intensidades

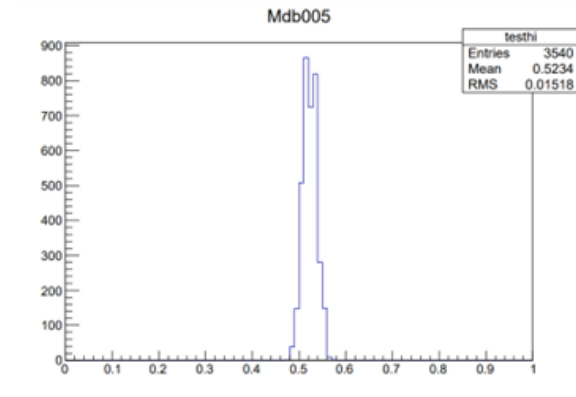


Figura 8.2: Histograma del tejido sano

*Se muestra un histograma de las intensidades de los pixeles del tejido sano*

Posteriormente procedimos a ordenar cada uno de los datos de las zonas especificas de las imagenes analizadas, obteniendo entonces las metricas, pero conservando el radio por si este pudiera ser requerido para una revision mas a fondo en el que se encontrara alguna anomalia.

Se presentan entonces algunos de los datos obtenidos, en la **tabla 8.1** en la que podemos notar algunos de los comportamientos de las métricas, tanto para las zonas que contienen tejido de tumores, como para las zonas que contienen tejido sano elegido en cada una de las imágenes.

Dato(Mdb)	Tumor	Tejido sano	Radio
001	52,487	30,048	197
002	11,229	1,841	69
005	1,169	970	30
010	2,154	1,594	33
012	2,388	1,672	40
⋮	⋮	⋮	⋮

Tabla 8.1: Metricas de los tejidos

*Se presentan los datos de las imagenes analizadas asi como las metricas de cada uno de las imagenes analizadas.*

*El resto de los resultados se encuentran en el apendice A*

Finalmente se realizo una comparacion entre las Metricas del Tejido Sano (MTS) contra las Metricas del Tejido del Tumor (MTT), en donde se uso un criterio de comparacion dado como:

$$P = \frac{MTS}{MTT}$$

En la **Figura 8.3** se muestra la distribucion del parametro P, para cada una de las imagenes, de tal forma que cada uno de los puntos representa el analisis realizado a una imagen individual, en este sentido, el eje x representa la imagen analizada mientras que el eje y representa el valor del parametro P para cada imagen analizada.

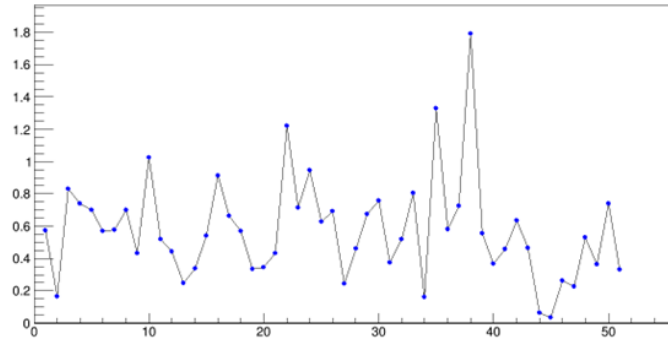


Figura 8.3: comparacion de metricas

*Cada uno de los puntos representa la comparacion de las metricas del tejido sano, contra el tejido del tumor, de cada una de las imagenes analizadas*

De tal forma que fuera facil notar el comportamiento a gran escala de las MTS como del las MTT evitando asi algun tipo de problema causado por los valores 'crudos' de los datos dados por el tamaño de los tumores y otros datos.

Asi pues, tenemos entonces un par de posibilidades para la comparacion entre las MTT contra las MTS

1.  $MTT > MTS$
2.  $MTT \leq MTS$

En el primer caso, tendríamos que el estimador P definido arriba se mostraria como:

$$P \leq 1$$

y por tanto en la grafica estos resultados deberian encontrarse en un rango entre 0 y 1,

Por otro lado, en el segundo caso, el estimador P, se mostraria como:

$$P > 1$$

De esta manera es sencillo ver el comportamiento a gran escala de cada uno de los datos, cabe mencionar que lo esperable es que dado que las zonas con tumores tienden a poseer las escalas mas 'luminosas', pero ademas las que contienen una densidad mas uniforme, deberian por tanto devolver metricas cuyos valores fueran mayores que las de el tejido sano.

Sin embargo podemos notar de inmediato, que existen anomalías dentro de el análisis que acabamos de realizar, por lo que notamos que es factible realizar un par de análisis adicionales para asegurarnos de que podamos corregir este tipo de datos que no suelen ser intuitivos.

Para ello podemos analizar que sucede con las métricas del tejido de los tejidos cuando las comparamos una a una, los resultados se muestran a continuación:

En la **Figura 8.4** se muestra la distribución de la correlación de las métricas, del tejido sano contra las métricas de los tumores, en donde podemos notar además, que las muestras parecen tener cierta relación entre si, la cual esta representada por el ajuste lineal.

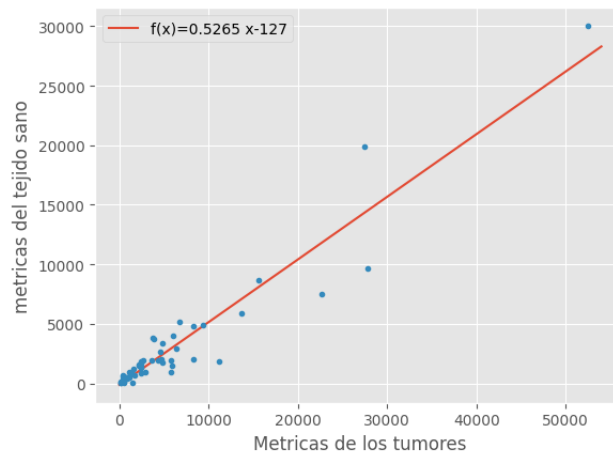


Figura 8.4: Métricas coorelacionadas

*Se muestra una comparación entre las métricas de los tumores, contra las métricas del tejido sano de cada una de las muestras*

$$f(x) = 0.5265x - 127$$

Podemos notar inmediatamente que existe cierta correlación lineal entre las métricas de ambos sets de datos, en la que se muestra que por lo general que el valor de las métricas de los tumores son en promedio el doble en comparación con las métricas del tejido sano, lo cual ajusta a los que se puede notar al revisar los resultados del parámetro P, en donde notamos que al menos 31 de las 50 imágenes tienen un indicador P por debajo de 0.6, lo cual refuerza el resultado anterior, sin embargo cabe resaltar por ahora que la cantidad de datos es insuficiente para determinar el comportamiento real a gran escala, pues harían falta la revisión de más casos.

Por otro lado, el modelo usado para realizar el ajuste lineal nos arroja que existe muy poca correlación de los datos, lo cual refuerza la sentencia anterior.

Finalmente, tomamos los histogramas de cada uno de los tejidos y realizamos un ajuste como se muestra a continuacion.

En la **Figura 8.5** se muestra el ajuste que se le realizo al histograma de intensidades de una muestra de tejido sano, cuyo proceso se repitio para cada una de las muestras, tanto de tejido sano como del tejido del tumor.

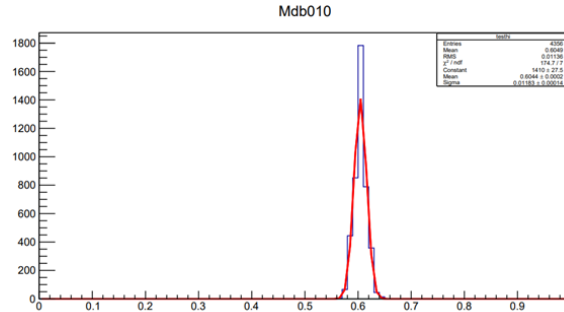


Figura 8.5: Ejemplo de ajuste de histograma

*Se muestra como se realizo el ajuste de uno de los histogramas del tejido sano*

Una vez hechos todos los ajustes, anotamos las medias y las desviaciones estandar ( $\sigma$ ) de cada uno de los ajustes realizados, pues nos interesa conocer si existe alguna característica adicional que podamos usar a nuestro favor para distinguir entre las zonas analizadas del tumor contra aquellas que hacen referencia unicamente al tejido sano.

El la **Tabla 8.2**, se muestran los resultados de haber obtenido las medias y las desviaciones de cada una de las muestras unicamente del tejido sano.

Muestra	Media	$\sigma$
001	0.3804	0.1721
002	0.3101	0.018338
005	0.5234	0.01223
010	0.6049	0.01183
012	0.511	0.01576
⋮	⋮	⋮

Tabla 8.2: Medias y sigma del tejido sano

*Se observan los resultados de el analisis del tejido sano.*

*El resto de los datos se encuentra contenido en el apendice A*

Mientras que en la **Tabla 8.3** se muestra los resultados de haber obteido las medias y las desviaciones estandar de los ajustes de cada una de las muestras de los tumores.

De donde ademas en la **Tabla 8.4** podemos observar cual es el promedio de los parametros anteriormente mencionados divididos por el tipo de tejido del que se este hablando.

Muestra	Media	$\sigma$
001	0.3804	0.1721
002	0.3101	0.018338
005	0.5234	0.01223
010	0.6049	0.01183
012	0.511	0.01576
$\vdots$	$\vdots$	$\vdots$

Tabla 8.3: Medias y sigma de los tumores

*Se observan los resultados del analisis de cada uno de los tumores.*

*El resto de los datos se encuentran contenidos en el apendice A*

Tejido	Media promedio	$\sigma$ promedio
Tumor	0.6773	0.03923
Tejido sano	0.4664	0.06240

Tabla 8.4: Promedios de las medias y las sigmas de los datos

*Se observan los resultados de promediar la media y la desviacion estandar de cada uno de los ajustes de los datos*

Finalmente, podemos observar como se comportan las medias tanto de los tumores como del tejido sano

En la **Figura 8.6** se muestra un diagrama de cajas en donde se expone el comportamiento de los parametros de la media de cada uno de los tejidos analizados, en donde se pueden observar ciertas regularidades a excepcion de algunos datos 'anomalos' dentro de la caracterizacion de los tumores, que se explicara en el capitulo siguiente.

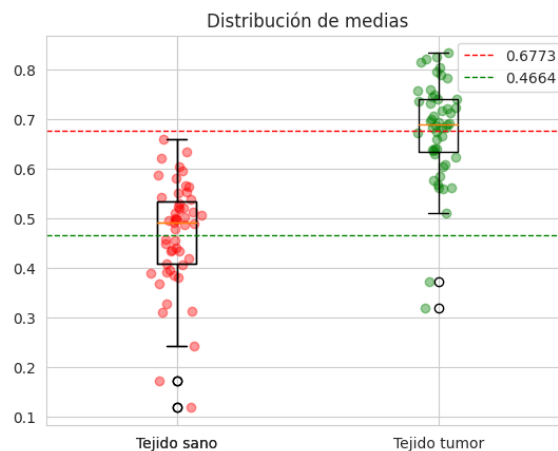


Figura 8.6: Medias de los tejidos

*Se muestra el comportamiento de las medias de cada uno de los datos analizados*

Aqui podemos darnos cuenta de que existen algunos tipos de anomalías en las que dadas las métricas de los tejidos, podemos inmediatamente conocer que una región tiene potencial para tratarse de un tumor, es decir, como podemos notar en la imagen de arriba, una media alta en los tejidos (en particular  $\mu > 0.7$ ) puede significar una anomalía.

Tomado en cuenta estos 3 valores, podemos ser capaces de caracterizar una buena cantidad de anomalías mamarias dentro de las imágenes mamográficas.

Después de haber obtenido los resultados correspondientes al análisis aplicado a las mamografías, podemos entonces discutir los resultados.

En primera instancia, encontramos que las métricas medidas de los tejidos, pueden usarse como un factor indicador en el análisis de la detección de tumores.

Esto se debe a que los tumores por lo general presentan las zonas con tipos de tejido que interactúan de mejor forma con los rayos X y por lo que suelen presentarse como las zonas más luminosas dentro de las mamografías, por lo que las métricas, que cuenta las intensidades de zonas del tejido mamario, puede ser usado como un indicador en la detección de anomalías mamarias, con lo que si bien, es poco probable que se pueda detectar al 100% un tumor, es también cierto que junto a otros indicadores y métodos de detección, puede volverse una herramienta que sea capaz de analizar y simplificar cálculos más complejos usado como estimador.

En la mayoría de las investigaciones con un enfoque en el análisis de las imágenes mamográficas, se usan distintos indicadores para realizar el análisis posterior usando machine learning o usando redes neuronales, sin embargo, el foco principal de las investigaciones es el de proponer estimadores que funcionen como ventana para que las diversas zonas de una imagen puedan ser comprendidas por la computadora y caracterizadas como tumores.

Una de las limitaciones más grandes que podemos encontrar al momento de analizar los datos obtenidos es la de la baja cantidad de datos analizados, pues existen patrones que pueden ser algo difíciles de interpretar debido a la falta de más elementos que podamos ingresar al análisis realizado, sin embargo, algunos otros datos podrían decirse que son concluyentes.

Es de esperarse de forma intuitiva, como se ha mencionado anteriormente, que el factor  $P = \frac{MTS}{MTT}$  sea siempre menor a 1, dadas las características que presenta de la interacción entre los tumores y los rayos x, sin embargo, podemos encontrar algunas anomalías en los datos marcados con las etiquetas [Mdb021,Mdb126 ,Mdb190,Mdb195]

Los cuales pareciera sugerir que existen zonas de tejido sano que contienen las zonas más luminosas dentro de la imagen. Aun así, estos resultados parecen ser anómalos y no concordar con el patrón que siguen la gran mayoría de los datos analizados.

Por lo cual, podríamos realizar una breve revisión a estos datos en la búsqueda de

algun tipo de anomalía que explique si existe algún error en las mediciones o por otro lado, es una generalidad que debería ser tomada en cuenta en futuras investigaciones.

**Mdb021** En esta muestra se nos presenta un tumor del tipo graso-glandular, no se encontró ningún indicio de elemento fuera de lugar por lo menos dentro de la imagen, por lo que se puede concluir que a grandes rasgos, pueden y de hecho existen anomalías de este mismo tipo en muestras que se encuentren fuera del dataset de prueba.

**Mdb126** En el caso de la muestra presente, se pudo encontrar al analizar la imagen, que la zona seleccionada para el análisis del tumor se encuentra fuera de rango, es decir, 'la bola' con la que está caracterizado este tumor embolsa también a una buena cantidad de datos nulos, por lo cual es predecible que se presente cualquier tipo de anomalía en análisis de este tipo.

**Mdb190** Esta muestra nos presenta a un tumor del tipo graso-glandular espicular que pudiera encontrarse en etapas tempranas, además dada la geometría un tanto irregular que presenta, es difícil centrar el estudio dentro del tumor a analizar

**Mdb195** Esta muestra, presenta un tumor de tipo graso-espicular en que presumiblemente podría encontrarse en etapas tempranas de desarrollo, con lo que la caracterización se vuelve difícil, además de que cabe resaltar que dentro de la muestra en donde se observa el tumor, existen zonas mucho más luminosas que pudieran deberse a factores de edad o a factores de salud. Con los que se vuelve aún más difícil su detección y caracterización por medio del estudio llevado a cabo en esta tesis.

Estos resultados a su vez no resultan por lo general ser tan reveladores en relación con el estudio realizado actualmente, dado que tan solo el 5.1% de los datos analizados resultaron mostrar comportamientos distintos a los esperados, aunque cabría la posibilidad de que estos resultados sean particulares del dataset en el que se ha estado trabajando, por lo que, para futuras investigaciones, se deberían usar una base de datos más grande y diversa.

Además, es coherente mencionar que este factor que es tomado en cuenta actualmente posee suficientes características que podrían potencialmente catalogarlo como un factor de decisión en métodos en el método MATH o algún otro basado en Thresholding.

Aun así, existen algunos factores que no han sido puestos a prueba en esta tesis y que tienen una clara correlación con los datos obtenidos, finalmente podría ser interesante, catalogar a los grupos de análisis por algunos de los factores más relevantes como podría ser la edad, el país o el estilo de vida para poder analizar características puntuales con los que la posibilidad de evitar errores sistemáticos aumentase de tal forma que se pueda extender este método a la detección de anomalías más severas e incluso en distintas partes del cuerpo.

## 9 Conclusiones

A lo largo de la investigación presentada se establecieron algunas bases matemáticas con la finalidad de describir características en el análisis de imágenes de mamografías que puedan ayudar a caracterizar las regiones en las que encontramos anomalías mamarias

Se midió una característica dentro de las imágenes como lo es la métrica Euclídea, aplicado a tumores benignos, y finalmente se observó una comparación de ambos sets de datos para extraer características.

De las 50 imágenes analizadas, existieron únicamente 4 anomalías, en el que se encuentra que el resultado de la métrica del tejido sano es mayor que la del tejido tumoral, sin embargo en una de las muestras, el tumor aparece incompleto, lo cual conduce a un error en la medición realizada, al analizarse detenidamente, pudimos encontrar que el factor decisivo para que las muestras presentaran esta anomalía se encuentra en la baja densidad de los tejidos de los tumores, o por el contrario, una alta densidad de la mama en el tejido sano.

Por lo que es lógico pensar que algunos factores como la propia estructura de la mama en cada individuo analizado tenga que ver en este resultado anómalo, la cantidad de tejido adiposo, tejido glandular y tejido muscular pueden volver propensos a un sujeto a que el estudio de las métricas como se realizó en este estudio, no pueda concretarse de manera correcta a la esperada.

Sin embargo, es importante notar que estos resultados anómalos representan (por lo menos en este estudio) un porcentaje bastante bajo de los sujetos analizados, representando únicamente el 6% de los resultados anómalos, por lo que se demuestra que esta podría ser una buena forma de filtrar resultados antes de aplicar estudios más complejos y costosos ya sea en cuestión monetaria o en cuestión de tiempo de cómputo.

Adicionalmente complementar el estudio aquí analizado con técnicas más avanzadas en el análisis de imágenes como en una segunda revisión de los datos anómalos puede conducir a encontrar información más precisa que ayude de mejor manera a los usuarios, y se hace énfasis en el uso de más de un método de detección de anomalías no solo en mamografías, sino en el análisis de imágenes obtenidas a través de métodos complementarios (como imágenes obtenidas por ecografías)

### 9.0.1. Retos y Metas a futuro

Aunque se ha logrado un avance excepcional en el tema de las detecciones mamarias, es importante reconocer que un buen diagnostico requiere de distintas herramientas para funcionar de manera correcta, puesto que cada uno de los métodos que son utilizados para la obtención de las imágenes tienen ventajas que pueden ser cubiertas por otro tipo de métodos de obtención de imágenes, en el caso de las mamografías, una de las principales desventajas es la de la incapacidad de detectar anomalías en mamas que contengan dentro de si tejido muy denso o por el contrario anomalías de baja densidad.

Algunos de los retos mas comunes que se encuentra en los artículos investigados es el de buscar parámetros que ayuden a caracterizar mejor a los tumores y en general a las anormalidades dentro de el estudio de las imágenes mamograficas, además métodos con los que se pueda reducir el ruido natural presente en las imágenes mamograficas, sin intervenir en la información presente dentro de la misma imagen, pues muchas veces ocurre que al momento de realizar el pre-procesamiento, se pierden algunas de las características de los tumores dada la naturaleza de las mamografías.

Dentro de las investigaciones que se exponen se hace énfasis en el uso de las herramientas que ofrece la tecnología, muchas de los métodos que se han desarrollado a lo largo de los años se complementan de tal forma que solventan muchos de los fallos que tienen en si mismos.

Esto, unido al hecho de que se pueden usar herramientas de machine learning o redes neuronales de una forma mucho mas simple hoy en día en comparación a los últimos 20 años, ha hecho que el nivel de especialización de los métodos descritos sea tal que muchos de los métodos propuestos requieren factores mas precisos con los que trabajar de tal manera que la sensibilidad de los métodos sea la suficiente para detectar las anomalías mamarias de forma temprana. [6]

De forma alternativa al uso de una sola metodología, es recomendable combinar los métodos de extracción de imágenes (mamografía, ultrasonidos, resonancia magnética) y explotar sus particulares ventajas, ya que en general, las imágenes obtenidas por cada uno de estos métodos contienen ventajas que se complementan entre si. [6]

# Bibliografía

- [1] OMS, “Cancer de mama,” tech. rep., <https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer>, marzo 2024.
- [2] Y. L. M. Sanchez, “Epidemiología del cancer de mama,” tech. rep., Instituto Mexicano del Seguro Social, 24 de octubre de 2022.
- [3] C. p. e. C. y. l. P. d. E. División de Prevención y Control del Cáncer, “Cancer de mama,” tech. rep., <https://www.cdc.gov/spanish/cancer/breast/basic;info/what-is-breast-cancer.htm>, julio2023.
- [4] M. V. Sandra Vargas, “Diagnóstico y caracterización de cáncer mamario en seres humanos: Una revisión,” 21 de agosto de 2021.
- [5] B. K. Gayathri and P. Raajan, “A survey of breast cancer detection based on image segmentation techniques,” *International Conference on Computing Technologies and Intelligent Data Engineering*, 2016.
- [6] A. S. V. R. H. M. Dabass, J., “Segmentation techniques for breast cancer imaging modalities- a review,” *9th International Conference on Cloud Computing, Data Science Engineering (Confluence) - IEEE Explore Digital Library.*, 2009.
- [7] L. L. X. L. Qinghua Huang, Feibin Yang, “Automatic segmentation of breast lesions for interaction in ultrasonic computer-aided diagnosis,” *Elsevier*, 1 September 2015.
- [8] S. G. Patel, B.C. and D. Soni, “Detection of masses in mammographic breast cancer images using modified histogram based adaptive thresholding (mhat) method,” *International Journal Biomedical Engineering and Technology*.
- [9] e. C.H. Zepeda Fernandez, E. Moreno Barbosa, “Tumor and microcalcification characterization using entropy, fractal dimension and intensity values statistical analysis in a mammography,” *arXiv e-prints*, 2021.
- [10] M. H. M. Nasseer M. Basheer, “Segmentation of breast masses in digital mammograms using adaptive median filtering and texture analysis,” *International Journal of Recent Technology and Engineering*, 2013.

- [11] Y. Prakash Bethapudi, E.Sreenivasa Reddy, "Detection and identification of mass structure in digital mammogram," *International Journal of Computer Applications*, 2013.
- [12] G. A.-A. Z.A. Abo-Eleneen, "A novel statistical approach for detection of suspicious regions in digital mammogram," *Egyptian Mathematical Society*, 2013.
- [13] M. Nakahara, *geometry, topology and physics*. Institut of physics publishing, 2003.
- [14] J. v. Varilly, *MA-870:Geometria Diferencial*. Escuela de matematica, Universidad de costarrica, 2014.
- [15] D. P. L. Sorin, *geometria y Topologia*. Universidad de Murcia.
- [16] W. M. Boothby, *An introduction to differentiable manifolds and riemannian geometry*. Samuel Ellenberg and Hyman Earns, 1975.
- [17] J. M. D. Moreno, *Introduccion a la topologia de los espacios metricos*. SERVICIO DE PUBLICACIONES UNIVERSIDAD DE CADIZ, 1998.
- [18] E. L. Lima, *espaços métricos*. Projeto Euclides, 1993. 3a edicion.
- [19] F. E. Gordejuela, "Topologia diferencial."

# A Tablas Completas

Dato(Mdb)	Eje X	Eje Y	Radio
001	535	425	197
002	522	280	69
005	477	133	30
010	525	425	33
012	471	458	40
013	667	365	31
015	595	864	68
017	547	573	48
019	693	477	49
021	493	125	49
025	674	443	79
030	522	676	43
032	388	742	66
063	546	463	33
069	462	406	44
080	432	149	20
097	612	297	34
099	714	340	23
104	357	365	50
107	600	621	111
121	492	434	87
126	191	549	23
127	523	551	48
132	522	788	52
142	347	636	26
144	233	994	29
145	669	543	49
150	351	661	62
152	675	486	48
160	536	519	61
163	391	365	50
165	537	490	42
167	574	657	35
188	406	617	61
190	512	621	31
191	594	516	41
193	399	563	132
195	725	129	26
198	568	612	93
199	641	177	31
207	571	564	19
218	519	629	8
219	546	756	29
222	398	427	17
223	523	482	29
227	504	467	9
236	276	824	14
248	378	601	10
290	337	353	45
314	518	191	39
315	516	447	93

Tabla A.1: Datos del tejido del tumor  
*Encontramos las coordenadas 'X', 'Y' y el radio 'R' de cada una de los tumores en las imagenes*

Dato(Mdb)	Eje X	Eje Y	Radio
001	575	749	197
002	409	701	69
005	537	569	30
010	398	618	33
012	339	681	40
013	731	963	31
015	527	949	68
017	599	681	48
019	546	753	49
021	775	373	49
025	557	735	79
030	497	877	43
032	458	850	66
063	544	766	33
069	595	799	44
080	550	574	20
097	740	854	34
099	650	604	23
104	400	754	50
107	725	224	111
121	630	824	87
126	270	724	23
127	620	724	48
132	130	544	52
142	330	784	26
144	288	324	29
145	810	874	49
150	200	524	62
152	420	859	48
160	440	839	61
163	780	204	50
165	560	804	42
167	635	810	35
188	450	894	61
190	484	94	31
191	513	787	41
193	705	334	132
195	603	689	26
198	480	784	93
199	540	804	31
207	520	750	19
218	330	734	8
219	560	846	29
222	555	744	17
223	650	84	29
227	585	774	9
236	260	188	14
248	325	760	10
290	480	778	45
314	273	761	39
315	597	830	93

Tabla A.2: Datos del tejido sano

*Encontramos las coordenadas 'X', 'Y' y el radio 'R' de cada una de zonas del tejido sano elegidas arbitrariamente*

Dato(Mdb)	Tumor	Tejido sano	Radio
001	52,487	30,048	197
002	11,229	1,841	69
005	1,169	970	30
010	2,154	1,594	33
012	2,388	1,672	40
013	2,402	1,367	31
015	8,286	4,771	68
017	4,836	3,388	48
019	4,671	2,010	49
021	3,730	3,821	49
025	9,337	4,855	79
030	4,340	1,918	43
032	8,311	2,066	66
063	2,908	983	33
069	2,446	1,324	44
080	433	395	20
097	2,202	1,458	34
099	921	526	23
104	5,783	1,934	50
107	27,800	9,612	111
121	13,725	5,917	87
126	503	614	23
127	2,707	1,934	48
132	3,919	3,712	52
142	1,184	742	26
144	1,402	971	29
145	5,2929	1,451	49
150	6,382	2,936	62
152	5,981	4,025	48
160	6,778	5,145	61
163	4,770	1,778	50
165	3,642	1,888	42
167	1,551	1,250	35
188	5,732	925	61
190	442	587	31
191	4,569	2,657	41
193	27,488	19,910	132
195	383	687	26
198	15,621	8,673	93
199	1,740	640	31
207	816	372	19
218	122	77	8
219	1,117	519	29
222	557	35	17
223	1,414	46	29
227	168	43	9
236	543	123	14
248	188	99	10
290	2,386	866	45
314	2,489	1,839	39
315	22,616	7,459	93

Tabla A.3: Metricas de los tejidos

*Se presentan los datos de las imagenes analizadas asi como las metricas de cada uno de las imagenes analizadas*

Muestra	Media	$\sigma$
001	.3804	0.1721
002	.3101	0.018338
005	.5234	0.01223
010	.6049	0.01183
012	.511	0.01576
013	.5961	0.01933
015	.4961	0.5639
017	.5639	0.0237
019	.4562	0.521
021	.6337	0.0213
025	.4406	0.03339
030	.5125	0.04991
032	.3272	0.1031
063	.4776	0.03135
069	.408	0.2261
080	.4968	0.01068
097	.5656	0.01488
099	.4983	0.01592
104	.4367	0.02786
107	.4883	0.01422
121	.4328	0.003845
126	.5387	0.01367
127	.4576	0.0201
132	.5801	0.01479
142	.5289	0.01244
144	.5418	0.01458
145	.3881	0.01992
150	.4348	0.01723
152	.6607	0.0139
160	.5865	0.02015
163	.4191	0.04637
165	.5199	0.02343
167	.5014	0.01402
188	.2418	0.05207
190	.3908	0.01278
191	.6223	0.02509
193	.5204	0.06711
195	.4993	0.01106
198	.4904	0.04539
199	.4066	0.03308
207	.5071	0.0293
218	.5504	0.01146
219	.3853	0.02892
222	.171	0.01621
223	.1172	0.02225
227	.368	0.009791
236	.3958	0.02874
248	.4866	0.01273
290	.3126	0.0561
314	.5533	0.01044
315	.4497	0.02668

Tabla A.4: Medias y sigma del tejido sano  
*Se observan los resultados de  
el analisis del tejido sano*

Muestra	Media	$\sigma$
001	.0.5111	0.0.1156
002	.7484	0.02615
005	0.5609	0.03516
010	0.707	0.04167
012	0.6094	0.03654
013	0.7892	0.04578
015	0.6654	0.07276
017	0.722	0.0222
019	0.6953	0.05575
021	0.6824	0.04679
025	0.6932	0.0428
030	0.7605	0.03554
032	0.7376	0.06688
063	0.8162	0.03836
069	0.5862	0.04238
080	0.5683	0.05589
097	0.6842	0.03414
099	0.6829	0.02091
104	0.7841	0.02157
107	0.7962	0.04181
121	0.7501	0.04687
126	0.6402	0.03
127	0.637	0.05052
132	0.5615	0.02815
142	0.6605	0.03819
144	0.6038	0.01462
145	0.6891	0.0464
150	0.7008	0.04041
152	0.8059	0.01911
160	0.675	0.04435
163	0.732	0.04477
165	0.7411	0.02131
167	0.5604	0.02144
188	0.714	0.04923
190	0.3193	0.01369
191	0.8226	0.02047
193	0.7229	0.05142
195	0.3712	0.04742
198	0.7164	0.054
199	0.6885	0.04725
207	0.7583	0.03141
218	0.6403	0.02942
219	0.5775	0.03808
222	0.6928	0.03467
223	0.6388	0.03821
227	0.7414	0.03138
236	0.8271	0.01076
248	0.6737	0.01908
290	0.6229	0.06376
314	0.6295	0.05415
315	0.8337	0.02658

Tabla A.5: Medias y sigma de los tumores  
*Se observan los resultados de  
el analisis de cada uno de los tumores*

