



**INAOE**

## **Detección de Microcalcificaciones utilizando Discriminantes Lineales de Fisher**

por

**Gabriela Alejandra Rodríguez Ruiz**

Tesis sometida para obtener el grado de **Maestra en Ciencias** en la especialidad de **Ciencias Computacionales** en el Instituto Nacional de Astrofísica, Óptica y Electrónica.

Supervisada por:

**Dr. Jesús A. González Bernal**  
**Dr. Leopoldo Altamirano Robles**

Octubre 2009, Tonantzintla, Puebla

© INAOE 2009

Derechos reservados

El autor otorga al INAOE el permiso de reproducir y distribuir copias de su totalidad o en partes de esta tesis





## **Resumen**

Este trabajo tiene como objetivo fundamental desarrollar un método para la detección de microcalcificaciones en mastografías digitales, basado en técnicas de aprendizaje computacional y visión por computadora; el cual permita identificar microcalcificaciones de diversas formas, tamaños, tonalidades y además sea capaz de identificarlas no sólo en senos grasos sino también en senos densos sin incrementar el número de falsos positivos. Para lograr el objetivo planteado, el tipo de tejido por el que mayormente está formado el seno (densidad) es tomado en cuenta, lo que permite detectar microcalcificaciones aún en senos donde la diferencia de contraste entre el tejido y la microcalcificación es mínima (senos densos). En la etapa de segmentación los discriminantes lineales de Fisher (FLD) se utilizan para segmentar las posibles microcalcificaciones del seno, a partir de estas regiones, se extraen características (morfológicas y de intensidad de gris) y mediante un clasificador se reduce el número de falsos positivos.

Para probar el método, se utilizaron las bases de mastografías ISSSTEP y MIAS, para la evaluación del desempeño del método se utilizó el análisis ROC y el análisis FROC.



## **Abstract**

The main objective of this thesis is to develop a new method for detecting microcalcifications in digital mammograms, using machine learning and computer vision techniques; the method detects different shapes, sizes and intensities of microcalcifications, and also it's able to recognize them not only in fat breast but also in dense breast. To achieve this objective, the breast tissue is taken into account to detect microcalcifications even in dense breast, where the contrast difference between tissue and microcalcification is almost null. In the segmentation step the Fisher Linear Discriminants are used to segment the possible microcalcifications, to reduce the false positives generated in this step, some characteristics (morphological and intensity) are extracted from these regions. The method is tested; using ISSSTEP and MIAS databases, the ROC and FROC analysis were used as performance measures.



## **Agradecimientos**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el apoyo económico proporcionado para mis estudios de maestría, bajo el número de registro 212417.

A Dios por la vida que me ha dado, por ayudarme en todas mis necesidades y por siempre estar conmigo.

A mis asesores, el Dr. Jesús González Bernal y el Dr. Leopoldo Altamirano Robles, por compartir sus conocimientos científicos y valiosas sugerencias a lo largo del desarrollo de mi tesis. Gracias por su paciencia, su apoyo y sobre todo por creer en mí.

A mis revisores, el Dr. Jesús Ariel Carrasco Ochoa, Dr. Gustavo Rodríguez Gómez y Dr. Luis Villaseñor Pineda, por todo su tiempo, sugerencias y su gran paciencia.

A mi familia, por brindarme todo su apoyo, aliento y comprensión en todos los momentos importantes.





## **Dedicatorias**

A mis padres Fabiola y Norberto, por todo el amor y apoyo que me han brindado a lo largo de mi vida.

A mi hermana Naye, por su cariño y confianza además de su gran sentido del humor.



# Índice General

|  |          |
|--|----------|
| <b>1. Introducción</b>                                     | <b>1</b> |
| 1.1. Antecedentes  | 1        |
| 1.2. Definición del problema                               | 2        |
| 1.3. Objetivos   | 3        |
| 1.4. Metodología   | 4        |
| 1.5. Organización de la tesis                              | 5        |
| <br>   |          |
| <b>2. Fundamentos</b>                                      | <b>7</b> |
| 2.1. Conceptos Médicos                                     | 7        |
| 2.1.1. Cáncer de Seno                                      | 7        |
| 2.1.2. Mastografías  | 8        |
| 2.1.3. Densidad Mamográfica                                | 9        |
| 2.1.4. Calcificaciones en el seno                          | 10       |
| 2.2. Aprendizaje Automático                                | 12       |
| 2.2.1. Análisis ROC  | 16       |
| 2.3. Visión por computadora                                | 19       |
| 2.3.1. Imagen Digital                                      | 19       |
| 2.3.2. Histograma de una imagen                            | 20       |
| 2.3.3. Segmentación  | 21       |
| 2.3.4. Evaluación de algoritmos de segmentación            | 21       |
| 2.3.5. Extracción de características                       | 24       |
| 2.4. Técnicas para la reducción de dimensión de un espacio | 25       |
| 2.4.1. Eigenvectores y Eigenvalores                        | 26       |
| 2.4.2. Análisis de Componentes Principales                 | 27       |
| 2.4.3. Discriminantes Lineales de Fisher                   | 28       |

|  |        |
|--|--------|
| <b>3. Estado del Arte</b>  | 31     |
| 3.1. Introducción  | 31     |
| 3.2. Métodos basados en regiones   | 33     |
| 3.3. Métodos basados en contornos  | 35     |
| 3.4. Métodos basados en agrupamiento y basados en umbrales   | 36     |
| 3.5. Métodos de segmentación supervisada   | 38     |
| 3.6. Discusión   | 40     |
| <br><b>4. Detección de Microcalcificaciones utilizando FLD</b>   | <br>43 |
| 4.1. Introducción  | 43     |
| 4.2. Clasificación de la densidad del seno   | 45     |
| 4.3. Espacios discriminantes y segmentación de microcalcificaciones en mastografías digitales mediante FLD | 48     |
| 4.3.1. Bases de imágenes de micro-calcificaciones y de tejido del seno                                     | 49     |
| 4.3.2. Espacios discriminantes para detectar microcalcificaciones  | 49     |
| 4.3.3. Segmentación de microcalcificaciones utilizando FLD   | 51     |
| 4.4. Reducción de Falsos Positivos   | 52     |
| <br><b>5. Experimentos y Resultados</b>  | <br>55 |
| 5.1. Bases de Mastografías   | 55     |
| 5.2. Experimentos de la etapa de clasificación de la densidad del seno                                     | 58     |
| 5.2.1. Experimentos utilizando la base de datos MIAS   | 60     |
| 5.2.2. Experimentos utilizando la base ISSSTEP   | 66     |
| 5.3. Experimentos de la etapa de segmentación de las microcalcificaciones                                  | 71     |
| 5.3.1. Experimentos utilizando la base MIAS  | 73     |
| 5.3.2. Experimentos utilizando la base ISSSTEP   | 77     |
| 5.4. Experimentos de la etapa de reducción de falsos positivos   | 81     |
| 5.4.1. Experimento utilizando la base MIAS   | 82     |
| 5.4.2. Experimento utilizando la base ISSSTEP  | 84     |
| 5.5. Resultados del desempeño global del método propuesto  | 87     |
| 5.6. Discusión   | 88     |

|                        |    |
|------------------------|----|
| <b>6. Conclusiones</b> | 91 |
| 6.1. Sumario           | 91 |
| 6.2. Aportación        | 93 |
| 6.3. Conclusiones      | 94 |
| 6.4. Trabajo Futuro    | 94 |
| <b>Bibliografía</b>    | 97 |



## Índice de Figuras

|   |    |
|---|----|
| 2.1. Mastografía  | 8  |
| 2.2. Proyecciones de mastografía  | 9  |
| 2.3. Densidad mamográfica   | 10 |
| 2.4. Microcalcificaciones   | 11 |
| 2.5. Árbol de decisión  | 13 |
| 2.6. Red Neuronal Multicapa   | 15 |
| 2.7. Matriz de Confusión  | 17 |
| 2.8. Análisis de una imagen   | 19 |
| 2.9. Imagen digital   | 20 |
| 2.10. Clasificación de píxeles  | 22 |
| 2.11. Clasificación de regiones   | 23 |
| 2.12. Eigenvectores y eigenvalores  | 27 |
| 2.13. Análisis de Componentes Principales   | 28 |
| 2.14. Discriminantes Lineales de Fisher   | 29 |
| 3.1. Diagrama de Flujo de algoritmos de detección de microcalcificaciones           | 32 |
| 4.1. Esquema general del método de detección de microcalcificaciones utilizando FLD | 44 |
| 4.2. Microcalcificaciones en senos densos y grasos                                  | 46 |
| 4.3. Esquema del algoritmo para clasificar densidad del seno                        | 46 |
| 4.4. Segmentación del seno  | 47 |
| 4.5. FLD para detectar microcalcificaciones   | 48 |
| 4.6. Representación de una imagen como vector columna                               | 50 |
| 4.7. Segmentación de microcalcificaciones utilizando FLD                            | 52 |
| 5.1. Mastografías de la base MIAS   | 56 |

|   |    |
|---|----|
| 5.2.Mastografías base ISSSTEP .....   | 57 |
| 5.3. Esquema validación cruzada .....   | 58 |
| 5.4.Esquema de los experimentos de la clasificación de la densidad del seno .....                     | 59 |
| 5.5. Matrices de confusión del experimento 1 utilizando la base MIAS .....                            | 60 |
| 5.6. Matrices de confusión del experimento 2 utilizando la base MIAS .....                            | 62 |
| 5.7. Matrices de confusión del experimento 3 utilizando la base MIAS .....                            | 64 |
| 5.8. Matrices de confusión del experimento 1 utilizando la base ISSSTEP .....                         | 66 |
| 5.9. Matrices de confusión del experimento 2 utilizando la base ISSSTEP .....                         | 69 |
| 5.10. Matrices de confusión del experimento 3 utilizando la base ISSSTEP .....                        | 70 |
| 5.11. Microcalcificaciones detectadas en la base MIAS ,tomando en cuenta densidad<br>del seno .....   | 74 |
| 5.12. Evaluación de detección de microcalcificaciones utilizando la base MIAS. ....                   | 76 |
| 5.13. Microcalcificaciones detectadas en la base ISSSTEP tomando en cuenta densidad<br>del seno ..... | 79 |
| 5.14. Evaluación de detección de microcalcificaciones utilizando la base ISSSTEP. . .                 | 80 |



## Índice de Tablas

|   |    |
|---|----|
| 3.1. Ventajas y desventajas de métodos de detección de microcalcificaciones . . . . .   | 40 |
| 5.1. Resultados del experimento 1 de la etapa de clasificación de la densidad del seno<br>utilizando la base MIAS . . . . .       | 61 |
| 5.2. Resultados del experimento 2 de la etapa de clasificación de la densidad del seno<br>utilizando la base MIAS . . . . .       | 63 |
| 5.3. Resultados del experimento 3 de la etapa de clasificación de la densidad del seno<br>utilizando la base MIAS . . . . .       | 65 |
| 5.4. Resultados del experimento 1 de la etapa de clasificación de la densidad del seno<br>utilizando la base ISSSTEP. . . . .     | 67 |
| 5.5. Resultados del experimento 2 de la etapa de clasificación de la densidad del seno<br>utilizando la base ISSSTEP. . . . .     | 69 |
| 5.6. Resultados del experimento 3 de la etapa de clasificación de la densidad del seno<br>utilizando la base ISSSTEP. . . . .     | 71 |
| 5.7. Resultados del experimento de la etapa de segmentación utilizando la base MIAS. . . . .                                      | 73 |
| 5.8. Resultados de la comparación realizada con el método basado en wavelets y<br>umbrales, utilizando la base MIAS . . . . .     | 75 |
| 5.9. Resultados del experimento de la etapa de segmentación utilizando la base<br>ISSSTEP . . . . .                               | 78 |
| 5.10. Resultados de la comparación realizada con el método basado en wavelets y<br>umbrales, utilizando base la ISSSTEP . . . . . | 79 |
| 5.11. Resultados del experimento de la etapa de reducción de falsos positivos<br>utilizando la base MIAS. . . . .                 | 83 |

|  |    |
|--|----|
| 5.12. Resultados del experimento de la comparación de la etapa de reducción de falsos positivos utilizando la base MIAS. ....    | 84 |
| 5.13. Resultados del experimento de la etapa de reducción de falsos positivos utilizando la base ISSSTEP .....                   | 85 |
| 5.14. Resultados del experimento de la comparación de la etapa de reducción de falsos positivos utilizando la base ISSSTEP ..... | 86 |
| 5.15. Resultados del desempeño global del método propuesto .....   | 87 |

## **Lista de Acrónimos**

ACR - Colegio Americano de Radiología, del inglés American College of Radiology.

ANFIS - Sistema de inferencia neuro-difuso adaptativo, del inglés Adaptive Network based Fuzzy Inference System.

FLD - Discriminantes Lineales de Fisher, del inglés Fisher Linear Discriminant.

FN - Falso Negativo.

FP - Falso Positivo.

FPI - Falsos Positivos por imagen, del inglés False Positive per Image.

FPR - Proporción de Falsos Positivos, del inglés False Positive Rate.

FROC – Característica operativa del receptor de respuesta libre, del inglés Free-response Receiver Operating Characteristic.

GRNN - Red neuronal de regresión generalizada, del inglés Generalized Regression Neuronal Network.

ISSSTEP - Instituto de Seguridad y Servicios Sociales de los Trabajadores al servicio de los poderes del Estado de Puebla.

LoG - del inglés Laplacian of Gaussian.

MIAS - Sociedad de análisis de imágenes mamográficas, del inglés Mammographic Image Analysis Society.

PCA - Análisis de componentes principales, del inglés Principal Components Analysis.

ROC – Característica operativa del receptor, del inglés Receiver Operating Characteristic.

ROI - Región de interés, del inglés Region Of Interest.

SRDM - del inglés Surrounding Region Dependence Method.

SVM - Máquinas de vectores de soporte, del inglés Support Vector Machine.

TPR - Proporción de Verdaderos Negativos, del inglés True Positive Rate.

VN - Verdadero Negativo.

VP - Verdadero Positivo.

# **Capítulo 1**

## **Introducción**

### **1.1. Antecedentes**

El cáncer de seno ocupa actualmente uno de los primeros lugares a nivel mundial como causa de muerte para el sexo femenino, *The American Cancer Society* estima que en el año 2009, 465,000 mujeres alrededor del mundo morirán a causa de este tipo de cáncer; ver [20].

En México cada año mueren 3,500 mujeres por cáncer de mama, convirtiéndose en la primera causa de decesos y el segundo tumor más frecuente en este grupo de la población; ver [41].

Entre más avanzado esté el cáncer de seno menos probabilidad de sobrevivir tiene la paciente. Debido a que ninguna mujer está absuelta de padecer este tipo de enfermedad, los médicos recomiendan realizar periódicamente ciertas pruebas médicas (como la mastografía), con el fin de detectar cualquier posible indicio de cáncer en etapas tempranas, previniendo que éste llegue a un estado avanzado en el que la vida de la paciente esté en riesgo.

Existen diversas pruebas médicas para detectar el cáncer de seno tempranamente, una de ellas es el auto examen del seno, con el cual la mujer mensualmente se realiza una palpación en los senos con la finalidad de encontrar anormalidades. El problema del auto examen es que sólo se pueden detectar anormalidades o tumores mayores a 1 centímetro.

Otra herramienta muy utilizada para el diagnóstico temprano de cáncer de seno son las mastografías, las cuales pueden mostrar anormalidades del seno incluso cuando éstas aún no son palpables. Uno de los indicadores de posible cáncer de seno en sus primeras etapas son las microcalcificaciones, éstas son pequeñas acumulaciones de calcio de 0.1 mm a 1 mm de ancho; ver [28].

La detección de microcalcificaciones es una tarea difícil, éstas tienen muy poco contraste con el tejido que las rodea, por lo que es muy fácil confundirlas con tejido glandular. Por otra parte, debido a su diminuto tamaño, en ocasiones el médico las pasa por alto o las confunde con artificios (fallas en la mastografía). Por este motivo, desde hace ya varios años, se ha tratado de automatizar esta tarea haciendo uso de la visión por computadora y el aprendizaje computacional.

Los algoritmos que se han desarrollado para la detección de microcalcificaciones se pueden diferenciar por la manera en que segmentan o identifican las microcalcificaciones, éstos se pueden clasificar en uno de los siguientes grupos: basados en regiones, contornos, agrupamiento o basados en modelos. Sin embargo, hasta el momento, ninguno de los algoritmos desarrollados segmenta de manera eficaz las microcalcificaciones, provocando la generación de un alto número de falsos negativos (regiones no identificadas como microcalcificaciones que en realidad sí lo son) y aún más falsos positivos (regiones identificadas como microcalcificaciones que en realidad no lo son).

## **1.2. Definición del problema**

La tarea de detectar microcalcificaciones está muy lejos de ser algo trivial, desde hace más de dos décadas se han diseñado técnicas para la detección y clasificación automática de microcalcificaciones, pero esta tarea es complicada debido principalmente a lo complejo de su naturaleza: a la diversa iluminación que presentan las mastografías y la poca distinción que hay entre las microcalcificaciones y el tejido que las rodea. A continuación se

enumeran algunos de los principales problemas a los que nos enfrentamos en la tarea de detección de microcalcificaciones:

1. Las microcalcificaciones son muy pequeñas, su tamaño está en el rango de 0.1-1.0 mm.
2. La iluminación de las mastografías es muy variada, ya que ésta depende de la cantidad de rayos X que pasan a través del seno, por lo que utilizar un umbral fijo en el proceso de segmentación para diversas bases de imágenes no funciona.
3. Existen microcalcificaciones de diversas formas.
4. Dependiendo del tejido por el que mayormente esté compuesto el seno (densidad del seno), la diferencia de contraste entre las microcalcificaciones y el tejido que las rodea varía. Para senos formados por más tejido graso (senos grasos) el contraste entre una microcalcificación y el tejido que la rodea es muy notorio, pero en senos formados en su mayoría por tejido glandular (senos densos), la diferencia de contraste es casi nula.
5. Las microcalcificaciones suelen confundirse con artificios (fallas en la mastografía).
6. Existen microcalcificaciones con diferentes niveles de contraste debido principalmente a la composición química de éstas.

### **1.3. Objetivos**

El objetivo general de la tesis es desarrollar un algoritmo de segmentación basado en Discriminantes Lineales de Fisher para la detección de microcalcificaciones que tome en cuenta la densidad del seno y que permita reducir el número de falsos negativos sin incrementar de manera excesiva el número de falsos positivos.

Para alcanzar el objetivo propuesto se identificaron los siguientes objetivos particulares:

- Desarrollar un algoritmo para identificar la densidad del seno.

- Desarrollar un algoritmo para la detección de microcalcificaciones basado en Discriminantes Lineales de Fisher, que tome en cuenta la densidad del seno y además sea capaz de detectarlas sin importar su forma y tamaño.
- Reducir el número de falsos positivos mediante la extracción de características descriptivas de las microcalcificaciones y el uso de un clasificador.

## **1.4. Metodología**

La metodología propuesta para alcanzar los objetivos presentados en esta tesis se muestra a continuación:

1. **Identificar la densidad del seno.** Separar el seno del fondo de la mastografía utilizando el algoritmo *Fuzzy c-means*, extraer características que permitan identificar el tejido por el que mayormente está formado el seno (graso o glandular) y mediante un algoritmo de aprendizaje clasificar la densidad del seno (graso o denso).
2. **Generar espacios discriminantes utilizando Discriminantes Lineales de Fisher (FLD).** Para cada una de las dos clases de seno según su densidad (seno graso y seno denso), generar un espacio discriminante mediante FLD que permita diferenciar entre imágenes de microcalcificaciones e imágenes de tejido del seno.
3. **Segmentar microcalcificaciones utilizando espacios discriminantes creados mediante FLD.** Utilizar una ventana de tamaño fijo que analice toda la región del seno, proyectar la región dentro de la ventana al espacio discriminante correspondiente a la clasificación del seno, e identificar si esa región en el seno corresponde a una microcalcificación o a tejido del seno. Repetir el proceso hasta que la ventana haya recorrido y analizado todo el seno.



4. **Reducir el número de falsos positivos de las microcalcificaciones segmentadas.**

Extraer características de las posibles microcalcificaciones segmentadas y mediante el uso de un clasificador reducir el número de falsos positivos.

5. **Evaluación del método propuesto.** Utilizar análisis ROC y FROC para evaluar el desempeño obtenido por nuestro método. Comparar con otro método de detección de microcalcificaciones utilizando las mismas bases de mastografías.

## **1.5. Organización de la tesis**

La tesis está estructurada de la siguiente manera, en el capítulo 2 se presentan los fundamentos teóricos sobre el área médica, aprendizaje computacional, visión por computadora y técnicas para la reducción de dimensiones y extracción de características (Análisis de Componentes Principales, Discriminantes Lineales de Fisher).

En el capítulo 3 se presenta un estudio del estado del arte referente a algoritmos y técnicas para la detección y clasificación de microcalcificaciones utilizadas en los últimos años. En el capítulo 4 se presenta la metodología propuesta en esta tesis para la detección de microcalcificaciones en mastografías. En el capítulo 5 se describen los experimentos realizados para evaluar el desempeño el método propuesto, y finalmente en el capítulo 6 se presentan las conclusiones y perspectivas de esta tesis.



## **Capítulo 2**

### **Fundamentos**

#### **2.1. Conceptos Médicos**

##### **2.1.1. Cáncer de seno**

En términos generales, cualquier tipo de cáncer ocurre cuando las células en alguna parte del cuerpo humano comienzan a crecer sin control. Estas células cancerosas se expanden hasta desarrollar una alteración en los tejidos formando un tumor, el cual puede ser benigno o maligno. El término cáncer de seno, se refiere a un tumor maligno que se ha desarrollado a partir de células situadas en el seno. Usualmente comienza en las células de las glándulas productoras de leche o en los pasajes que transportan la leche hacia el pezón; ver [20].

Dependiendo del tamaño del tumor en el seno, éste puede ser detectado mediante la palpación realizada en el auto examen o mediante una mastografía. El diagnóstico mediante mastografías permite detectar el cáncer de seno en etapas tempranas incluso cuando el tumor no es palpable. Cuando los tumores se encuentran en una etapa inicial, son fáciles de tratar y si el cáncer de mama se detecta a tiempo, generalmente es curable. De aquí la importancia que ha tomado el diagnóstico temprano por medio de la mastografía.

### **2.1.2. Mastografías**

La mastografía es una técnica de rayos X (vea la figura 2.1), en la que se toma una serie de placas radiográficas en diferentes ángulos del seno. Este estudio médico tiene la finalidad de detectar anomalías o cambios en el seno. Las mastografías también son conocidas como mamografías y el equipo con el que se obtienen se llama mastógrafo.

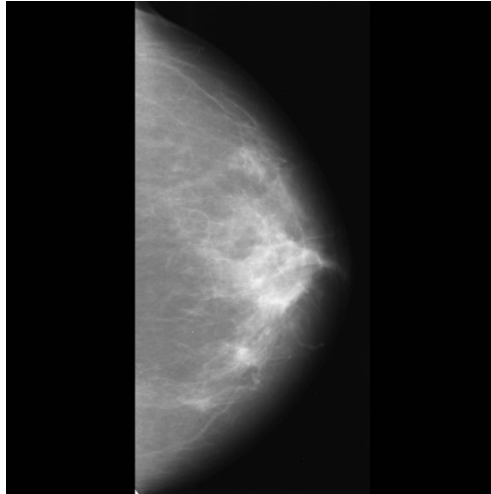


Figura 2.1. Ejemplo de una mastografía.

Un estudio de mastografías generalmente está compuesto por 2 tipos de proyecciones estándar que son: la medio-lateral oblicua y la cráneo-caudal (vea la figura 2.2). La proyección medio-lateral se toma a un ángulo de  $45^\circ$  permitiendo analizar el perfil del seno (proyección vertical) y observar un mayor volumen del tejido mamario. Por otra parte la proyección cráneo-caudal se toma desde la parte superior del seno, es decir, es una proyección horizontal del mismo. Esta proyección se utiliza para la evaluación de la superficie interna de la mama y provee mejor detalle de imagen, ya que permite comprimir más al seno. También se pueden usar un número de proyecciones adicionales en casos en que el médico quiera asegurarse de la ubicación exacta de un posible tumor; ver [40].

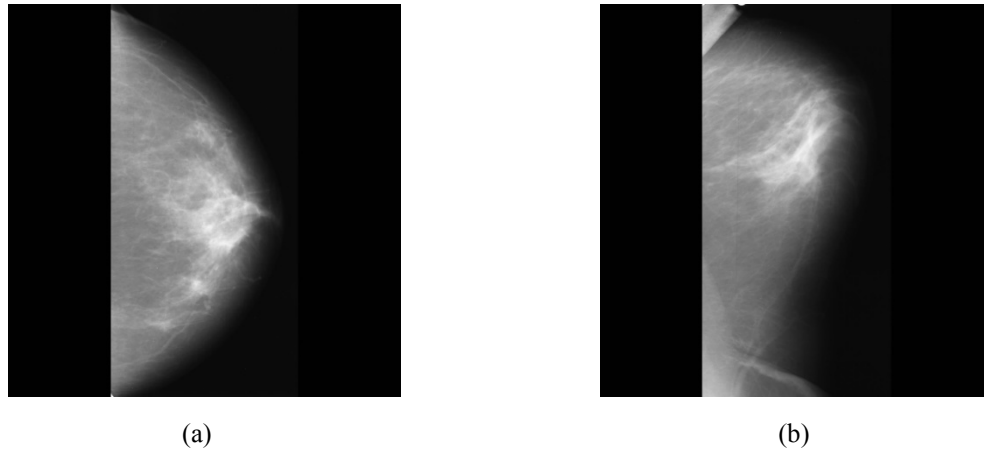


Figura 2.2. Ejemplos de proyecciones de mastografías, (a) medio-lateral oblicua y (b) cráneo-caudal.

La mastografía es útil para evaluar y analizar diferentes estructuras y anomalías asociadas al cáncer de seno como son las masas, microcalcificaciones, densidad del seno, distorsión de la arquitectura normal del tejido del seno, la asimetría entre el seno izquierdo y el derecho, entre otras; ver [28].

### 2.1.3. Densidad Mamográfica

En términos radiográficos el seno está conformado principalmente por dos tipos de tejido: tejido fibroglandular (perteneciente a la glándula mamaria) y tejido graso. El tejido fibroglandular es una mezcla de tejido conectivo fibroso y tejido glandular, el resto del seno es tejido graso. Los rayos X traspasan más fácilmente al tejido graso que al tejido fibroglandular, por lo que las regiones grasas del seno aparecen más oscuras mientras que las regiones pertenecientes al tejido fibroglandular tienen una tonalidad brillante o con mayor iluminación. A la cantidad de tejido fibroglandular en el seno se le conoce como densidad mamográfica. Si el seno contiene más tejido fibroglandular que tejido graso éste es catalogado como denso; ver [56].

Desde hace ya varios años, se han propuesto diversas clasificaciones para los senos de acuerdo a su densidad, entre los más importantes se encuentra el propuesto por Wolfe, el propuesto por ACR (American College of Radiology) y el método utilizado en la base de imágenes MIAS (Mammographic Image Analysis Society); ver [32,44].

Los dos primeros métodos clasifican la densidad del seno en 4 categorías con la diferencia de que el primero se basa en una estimación visual y el segundo en evaluaciones cuantitativas.

Las 4 categorías propuestas por Wolfe y ACR son: seno completamente graso, seno graso con escaso tejido fibroglandular, densidad heterogénea del seno y seno completamente denso. Por otro lado, el método utilizado en la base de imágenes MIAS clasifica la densidad del seno en tres categorías: denso-glandular, graso-glandular y graso (vea la figura 2.3).

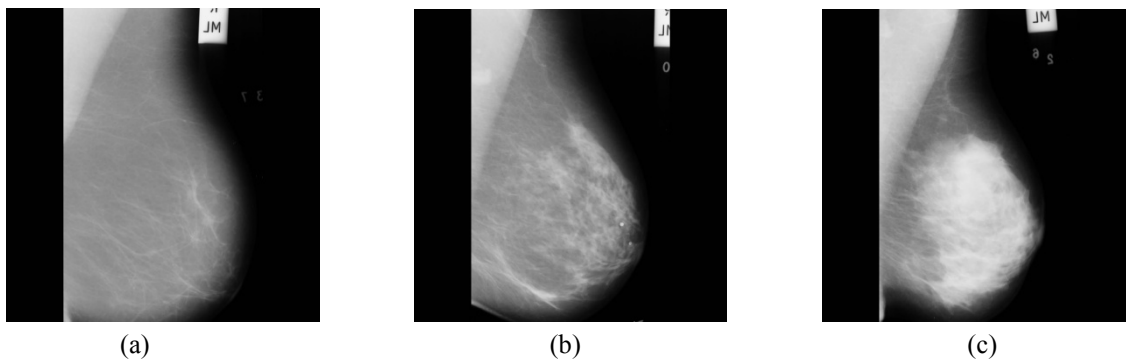


Figura 2.3. Clasificación de densidad propuesto por MIAS (a) graso, (b) graso-glandular y (c) denso-glandular.

#### **2.1.4. Calcificaciones en el seno**

Las calcificaciones del seno, son cristales compuestos generalmente de fosfato de calcio, carbonato de calcio, oxalato de calcio o fosfato de magnesio. Éstas se desarrollan en estructuras benignas o malignas del tejido del seno pero usualmente no es posible deducir su malignidad basándose únicamente en su composición química. Existen dos tipos de calcificaciones: macrocalcificaciones y microcalcificaciones; ver [13].

- **Macrocalcificaciones:** Generalmente son cambios degenerativos en el seno como resultado de antiguas lesiones, inflamaciones o por el envejecimiento de las arterias del seno. Este tipo de calcificaciones aparecen como manchas blancas grandes en el seno y no requieren de seguimiento ya que generalmente no están asociadas con el cáncer.
- **Microcalcificaciones:** Son pequeñas acumulaciones de calcio que miden entre .01 mm a 0.1mm, éstas aparecen como pequeños puntos blancos en la mastografía (vea

la figura 2.4). Algunas de las microcalcificaciones surgen a consecuencia del crecimiento anormal de las células del seno, por lo que éstas pueden ser un indicador de cáncer; ver [17].

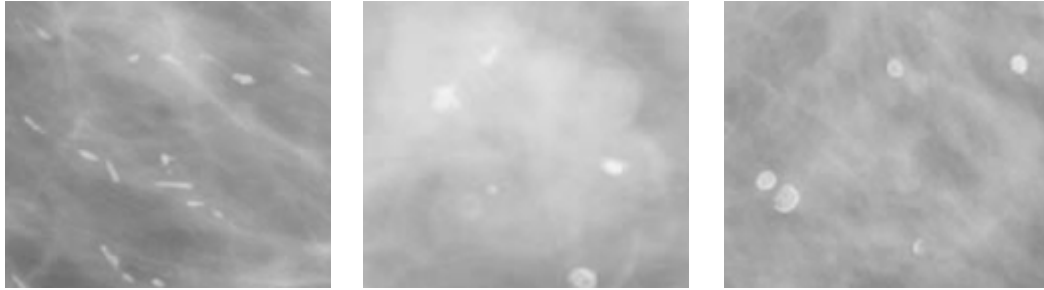


Figura 2.4. Imágenes de microcalcificaciones.

Al momento de analizar las calcificaciones del seno, se deben tomar en cuenta sus características, ya que las diferentes combinaciones de éstas pueden ser signos de cambios benignos o revelar un proceso maligno que aún no es palpable. Algunas de las características que se utilizan para evaluar la malignidad de una región (entre otras) son: la forma, el tamaño, la densidad, el número y su distribución; ver [13,17].

- Forma: La forma de las microcalcificaciones es una de las principales características que los médicos toman en cuenta. Existen microcalcificaciones de diversas formas como son: redondas, lineales, gruesas, en bastones, puntiformes, irregulares, granulares, curvilíneas o ramales. Las microcalcificaciones con mayor grado de malignidad son las granulares, ramales e irregulares.
- Tamaño: Las calcificaciones encontradas en una mastografía pueden medir a lo más 0.5mm, pero por lo general, las que se consideran sospechosas tienen un tamaño entre 0.01 y 0.1 mm.
- Densidad: Por lo general cuando las microcalcificaciones están asociadas a un proceso benigno su tonalidad es homogénea y de alta intensidad, mientras que las generadas por un proceso maligno son no homogéneas y tienen bajo contraste. El contraste de una microcalcificación depende de la cantidad de rayos X que pasen a través de ella dependiendo de su composición química.

- Número: En el seno se puede presentar una gran cantidad de calcificaciones, pero si en 1 centímetro cúbico aparecen más de 5 microcalcificaciones éste es un indicador de que son malignas.
- Distribuciones: Existen cinco tipos de agrupaciones para las microcalcificaciones, las cuales son: dispersas, lineales, dispersas en segmentos, ubicadas en regiones y grupos múltiples. Si la distribución de las microcalcificaciones es lineal o dispersas en segmentos, estas son altamente sospechosas.

Dada la importancia de la detección temprana de microcalcificaciones para así reducir el riesgo de muerte de mujeres por cáncer de seno, se han creado herramientas computacionales para analizar mastografías digitales. Algunos de los componentes principales para este análisis son los algoritmos de clasificación provenientes del área de aprendizaje automático y el análisis de imágenes por computadora, como se describe a continuación.

## 2.2. Aprendizaje Automático

Uno de los principales objetivos del aprendizaje automático es crear programas capaces de generalizar comportamientos a partir de información suministrada en forma de ejemplos.

El aprendizaje automático se define como: *“un programa de computadora se dice que aprende de la experiencia  $E$  con respecto a una clase de tareas  $T$  y medida de desempeño  $D$ , si su desempeño en las tareas en  $T$ , medidas con  $D$ , mejoran con la experiencia  $E$ ”*; ver [31].

Un tema de gran interés en el aprendizaje computacional es la predicción de clases. Donde una clase específica es asignada o atribuida a un objeto y cada objeto está representado por un número finito de atributos o características que lo describen. La clasificación tiene como objetivo obtener una descripción precisa o modelo para cada clase, utilizando los valores de los atributos de los ejemplos almacenados o de entrenamiento.

El modelo obtenido a partir de estos ejemplos tiene la finalidad de clasificar objetos cuyas clases se desconozcan. Para lograr un modelo fiable es necesario contar con un



número suficiente de ejemplos de entrenamiento, asegurando que los patrones obtenidos son válidos y no fueron sesgados debido a irregularidades o errores en los datos.

Existen diversas técnicas para generar estos modelos como son: árboles de decisión, redes neuronales, reglas, análisis discriminante, basado en instancias, regresión lineal entre muchas otras más.

A continuación se describen de manera muy general algunas de las técnicas de predicción utilizadas en esta tesis.

- **Árboles de decisión**

Los árboles de decisión permiten representar el conocimiento obtenido a partir de un conjunto de ejemplos. Debido a su sencillez y a su amplia aplicación, los árboles de clasificación o decisión son uno de los métodos más utilizados en el aprendizaje computacional; ver [31,53].

Un árbol de decisión está formado por nodos internos, ramas y nodos terminales u hojas. Un nodo interno representa un atributo del objeto, y cada rama perteneciente a este nodo interno representa uno de los posibles valores para ese atributo, teniendo una rama para cada posible valor de dicho atributo. Los nodos hojas o nodos terminales representan las clases a las que puede pertenecer el objeto (vea la figura 2.5).

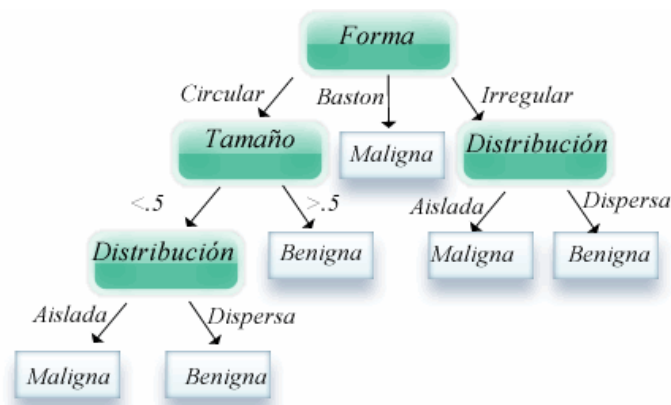


Figura 2.5. Ejemplo de un árbol de decisión para clasificar una microcalcificación en benigna o maligna.

La etapa de aprendizaje consiste en construir el árbol de decisión a partir de los ejemplos de entrenamiento, el espacio formado por los ejemplos se divide de manera recursiva aplicando reglas de decisión simples (i.e. en ID3 se utiliza una medida conocida como

ganancia de información, basada en el cálculo de la entropía de cada atributo, para elegir el siguiente nodo), mientras que la etapa de clasificación consiste en realizar una serie de preguntas acerca de los valores de los atributos del objeto. Para clasificar un objeto se comienza por el nodo raíz y se sigue por las ramas que contesten a las preguntas de los nodos intermedios, hasta llegar a un nodo terminal u hoja. El valor de este nodo hoja es la clase asignada al objeto. Por ejemplo en el árbol de decisión de la figura 2.5, siguiendo la rama del extremo derecho, se genera la siguiente regla: si la forma de la microcalcificación es irregular y además su distribución es dispersa entonces la microcalcificación es benigna.

Existen diferentes algoritmos basados en árboles de decisión, algunos de los más conocidos son: ID3, J4.8, C4.5, CLS y SPRINT; ver [31,53].

- **Aprendizaje basado en reglas**

A diferencia de los árboles de decisión, en el aprendizaje basado en reglas se toman las clases de manera aislada. Su objetivo es cubrir la mayor cantidad de ejemplos pertenecientes a una clase y excluir la mayor cantidad de ejemplos que no pertenezcan a esa clase; ver [6, 31].

Las reglas se pueden definir a partir de árboles de decisión o pueden ser definidas directamente. La estructura de una regla básica sería la siguiente:

**Si** (atributo(s) = valor(es) específico(s)) **entonces**

Clasificación del Objeto = Clase asociada al valor(es) del atributo(s)

Una vez que las reglas han sido definidas a partir de un conjunto de ejemplos de entrenamiento, éstas se utilizan para clasificar nuevas instancias.

En general los algoritmos basados en este tipo de aprendizaje generan reglas que tienden a ser compactas y que son fáciles de entender, por estas razones hay una preferencia por el aprendizaje basado en reglas sobre los árboles de decisión.

Algunos algoritmos que utilizan aprendizaje basado en reglas son PRISM, NNge, Decisión Table, PART; ver [52, 53].

- **Redes Neuronales**

Una red neuronal es un modelo o técnica inspirada en sistemas neuronales biológicos que abarca problemas de clasificación de objetos mediante el conocimiento obtenido a partir de un conjunto de ejemplos o instancias; ver [2].

La principal característica de este modelo es la estructura que se utiliza para el procesamiento de la información, está formado por un gran número de elementos (neuronas) interconectados entre sí para resolver de manera grupal un problema en específico; ver [6].

Estas neuronas pueden estar distribuidas en múltiples capas, y cada capa de neuronas envía información a la siguiente capa de neuronas.

Existen diversas arquitecturas de redes neuronales, pero generalmente se cuenta con una capa de entrada, un cierto número de capas intermedias que están ocultas y una capa de salida (vea la figura 2.6). Cada capa está formada por un número de neuronas que se encargan de procesar la información.

La capa de entrada recibe información del exterior, las capas intermedias tienen la tarea de realizar el proceso de clasificación o predicción y finalmente la capa de salidas proporciona el resultado de la clasificación al exterior.

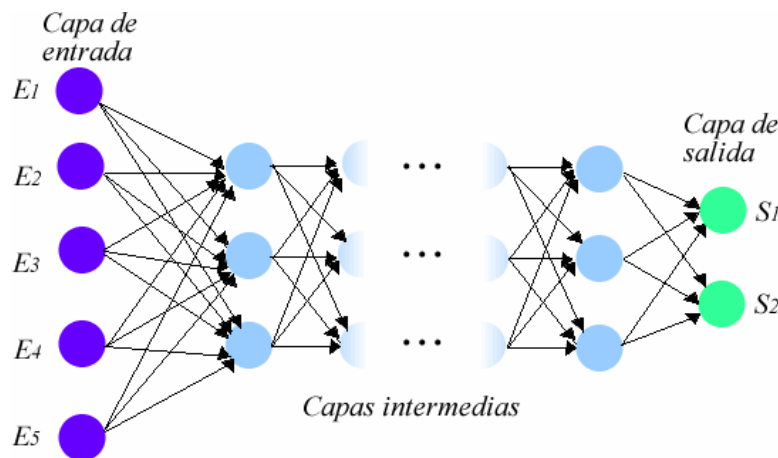


Figura 2.6. Esquema de una red neuronal, donde los círculos representan neuronas y las flechas conexiones entre ellas.

Cada neurona está dedicada a procesar información proveniente de las neuronas situadas en la capa anterior, utiliza sus entradas como parte de una función y si el resultado de dicha función es mayor a un umbral, ésta se activa emitiendo una señal a las neuronas situadas en la siguiente capa. En caso de que el resultado de la función no supere el umbral no envía ninguna señal manteniéndose inactiva.

Finalmente, para la capa de salidas, algunas neuronas de esta capa se activarán (1) y otras se quedarán inactivas (0). El resultado del conjunto de las neuronas de la capa de salida puede ajustarse o no con el deseado para el conjunto de entradas a la red. Algunos tipos de redes neuronales son: perceptron, backpropagation, multicapas, ADALINE, entre otros; ver [6, 25].

- **Aprendizaje basado en instancias**

La característica substancial de los algoritmos basados en instancias se localiza en la fase de entrenamiento (aprendizaje), donde simplemente se almacenan los ejemplos; ver [6, 31].

En general la mayoría de los métodos de aprendizaje se basan en encontrar modelos o descripciones generales a partir de los ejemplos de entrenamiento. Estos modelos se utilizan para clasificar objetos cuyas clases se desconocen. A diferencia de estos métodos, los algoritmos basados en instancias solo recopilan y almacenan los ejemplos en la etapa de aprendizaje. En la etapa de clasificación se calcula la similitud (medida de distancia) de la nueva instancia con los ejemplos almacenados y se le asigna la clase de los objetos más parecidos. Debido a que en la fase de aprendizaje los ejemplos no son analizados ni procesados con el fin de encontrar patrones, este tipo de algoritmos también son conocidos como perezosos.

Algunos algoritmos que utilizan aprendizaje basado en instancias son K-Nearest Neighbor (KNN), IbK, KStar, entre otros; ver [24, 25, 31, 53].

### **2.2.1. Análisis ROC**

El análisis ROC (*Receiver Operating Characteristic*) sirve como herramienta para evaluar el desempeño de un clasificador, este análisis se obtiene a partir de una matriz de

confusión de un clasificador binario, donde las instancias solo pueden ser positivas o negativas; ver [21].

A continuación se describe como obtener una matriz de confusión y las medidas de evaluación utilizadas en el análisis ROC para evaluar el desempeño de un clasificador.

- **Matriz de confusión**

La matriz de confusión muestra el desempeño de un algoritmo de clasificación sobre los datos de prueba, e indica cuántos errores cometió en cada clase y con qué clase las confundió. La matriz de confusión está formada por el número de instancias clasificadas correctamente y el número de instancias etiquetadas de manera errónea; ver [7, 21, 38].

Generalmente las columnas de la matriz representan el número de predicciones de cada clase, mientras que cada fila representa el número de instancias de la clase real (vea la figura 2.7).

|                |          | Clasificados como |          |
|----------------|----------|-------------------|----------|
|                |          | Positivo          | Negativo |
| Valores Reales | Positivo | VP                | FN       |
|                | Negativo | FP                | VN       |

Figura 2.7. Matriz de confusión de dos clases.

- **Verdadero Positivo (VP):** Es el número de predicciones correctas de que una instancia es positiva.
- **Falso Negativo (FN):** Es el número de predicciones incorrectas de que una instancia es negativa (siendo que realmente es positiva).
- **Verdadero Negativo (VN):** Es el número de predicciones correctas de que una instancia es negativa.
- **Falso Positivo (FP):** Es el número de predicciones incorrectas de que una instancia es positiva (siendo que realmente es negativa).

A partir de la matriz de confusión se extraen los siguientes parámetros utilizados para evaluar el resultado de algoritmos de clasificación.

- **Exactitud (*Accuracy*):** Proporción del número total de predicciones positivas:

$$Accuracy = \frac{VP + VN}{VP + FN + FP + VN} \quad (2.1)$$

- **Proporción de VP (*True Positive Rate*):** También conocido como **sensibilidad**, es la proporción de casos positivos que fueron correctamente identificados, y se calcula con la siguiente ecuación:

$$TPR = \frac{VP}{VP + FN} \quad (2.2)$$

- **Proporción de FN (*False Negative Rate*):** La proporción de casos positivos que fueron incorrectamente clasificados como negativos:

$$FNR = \frac{FN}{VP + FN} \quad (2.3)$$

- **Proporción de VN (*True Negative Rate*):** También conocido como **especificidad**, es la proporción de casos negativos que fueron correctamente identificados.

$$TNR = \frac{VN}{FP + VN} \quad (2.4)$$

- **Proporción de FP (*False Positive Rate*):** Proporción de casos negativos que fueron incorrectamente clasificados como positivos:

$$FPR = \frac{FP}{FP + VN} \quad (2.5)$$

En los sistemas orientados a la detección de anomalías o enfermedades médicas, los valores más significativos son las proporciones de verdaderos positivos (TPR) y de falsos positivos (FPR). El valor de TPR indica la cantidad de anomalías que se

detectaron correctamente, mientras que el FPR muestra la cantidad de ejemplos que se catalogaron como anomalía pero que realmente no lo son.

## 2.3. Visión por computadora

El análisis e interpretación de imágenes digitales mediante el uso de las computadoras ha tenido un gran auge en las últimas décadas, la visión por computadora ha sido utilizada en diversas áreas como medicina, robótica, tecnología industrial, tecnología militar, ingenierías, entre muchas otras más.

El objetivo principal de la visión por computadora es identificar patrones en una imagen digital que permitan clasificar los objetos (regiones) existentes en ésta. El propósito es que la computadora sea capaz de obtener información a partir de una imagen para analizar e interpretar sus características; ver [22,30].

El esquema general de las operaciones utilizadas en la visión por computadora parte de una imagen digital a la que se le puede realizar un pre-procesamiento para mejorar la calidad visual de la imagen, realizar la segmentación en donde se realiza la extracción de estructuras (regiones, contornos) de la imagen, el análisis de los objetos (textura, forma, tamaño) y finalmente la clasificación de los objetos, donde se determinan las diferentes clases de objetos en la imagen (vea la figura 2.8).

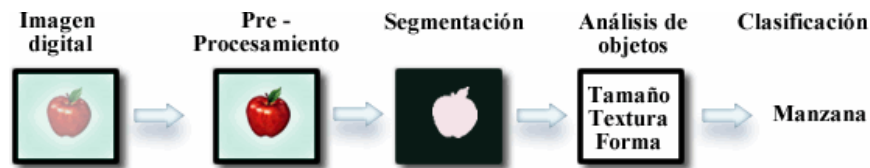


Figura 2.8. Esquema general utilizado en visión por computadora.

### 2.3.1. Imagen Digital

Una imagen se define como una "función bidimensional,  $f(x,y)$ , donde  $x$  y  $y$  son coordenadas espaciales y la amplitud  $f$  es llamada intensidad o nivel de gris en ese punto. Cuando  $x$ ,  $y$ , y  $f$  son todos enteros, es decir cantidades discretas, se le llama imagen digital"; ver [22].

Una imagen digital está compuesta por un número finito de elementos, cada uno de los cuales cuentan con un valor de intensidad de color y una posición particular, a estos elementos se les llama píxeles (vea la figura 2.9). El número máximo de píxeles a lo largo y a lo ancho representan el tamaño de la imagen (resolución espacial).

El número de bits utilizados para almacenar información del color en cada píxel define la resolución de la imagen (resolución radiométrica). Usualmente, para almacenar la información de un píxel se emplea 1 byte (8 bits), esto proporciona capacidad de almacenar hasta 256 niveles de escala de grises, donde el valor 0 corresponde al nivel más oscuro y el 255 al nivel más claro.

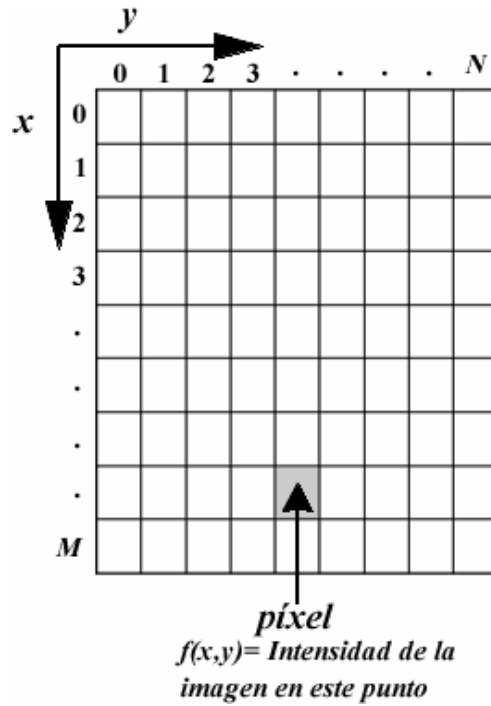


Figura 2.9. Representación de una imagen digital.

### 2.3.2. Histograma de una imagen

El histograma de una imagen es una representación gráfica de la distribución de tonalidades en una imagen digital, si éste tiene un total de  $L$  posibles niveles de intensidad, se define mediante la siguiente función discreta:

$$h(r_k) = n_k \quad (2.6)$$



Donde  $r_k$  es el K-ésimo nivel de intensidad en el intervalo  $[0, L-1]$  y  $n_k$  es el número de píxeles en la imagen cuyo nivel de intensidad es  $r_k$ ; ver [22].

El valor de  $L$  es el número de colores que pueden ser representados mediante el número de bits que definen al píxel, es decir, si cada píxel se representa con 8 bits, existirán 256 tonos de gris, en este caso  $L = 256$ , si cada píxel se representa con 16 bits, existirán 65,536 tonos de gris y en este caso  $L = 65,536$  y así sucesivamente,  $L$  varía dependiendo del número de bits que representan a los píxeles.

### 2.3.3. Segmentación

La segmentación es una de las etapas más complicadas en el análisis de imágenes, ésta consiste en dividir la imagen en regiones con características similares. El objetivo principal de la segmentación es identificar y localizar objetos específicos en una imagen; ver [39,48].

*“Sea  $R$  la región que incluye la imagen completa, se puede definir a la segmentación como un proceso que divide a  $R$  en  $k$  subconjuntos o sub-regiones disjuntas no vacías  $R_k$  cumpliendo con las siguientes reglas”:*

- $\bigcup_{i=1}^k R_i = R$  para  $i = 1, 2, \dots, k$
- $R_i \cap R_j = \emptyset$  para todo  $i$  y  $j$ ,  $i \neq j$ ,

### 2.3.4. Evaluación de algoritmos de segmentación

El problema de evaluar la eficacia de un algoritmo de segmentación/detección en mastografías, es que aunque se han propuesto diversas métricas de evaluación aún no existe una que sea estándar. A continuación se describe la técnica más utilizada para la evaluación de algoritmos de segmentación y la más utilizada con algoritmos de detección de microcalcificaciones en mastografías.

- **Análisis ROC**

El análisis ROC también se puede aplicar para evaluar algoritmos de **segmentación**, donde se realiza la evaluación a **nivel píxel**; ver [8,38].

El análisis ROC utilizado para evaluar clasificadores (mencionado en la sección 2.2.1) no difiere mucho del que se utiliza para evaluar algoritmos de segmentación.

Cada píxel de la imagen se trata como una instancia en el proceso de clasificación, de esta manera un píxel pertenece a la región de interés (en nuestro caso a una microcalcificación) o no pertenece a ella (tejido del seno). Por lo tanto, el píxel segmentado por el algoritmo de detección es comparado con el mismo píxel de la imagen segmentada manualmente por un experto, resultando entonces como un píxel bien clasificado o un píxel mal clasificado (vea la figura 2.10).

Un píxel se considera como bien clasificado si en ambas imágenes pertenece a la misma región, mientras que un píxel mal clasificado es aquel que pertenece a diferentes regiones en cada imagen. Para nuestro problema, si en la imagen segmentada por el experto el píxel pertenece a la región de microcalcificación, éste es un Verdadero Positivo (VP), y si el píxel pertenece a la región del tejido del seno, éste es un Verdadero Negativo (VN).

Si en la imagen manualmente segmentada el píxel pertenece a la región de microcalcificación y en la imagen segmentada por el algoritmo pertenece a la clase tejido del seno, ese píxel se considera como Falso Negativo (FN). De manera contraria, si el píxel pertenece a la región de tejido del seno en la imagen manualmente segmentada y en la imagen segmentada por el algoritmo el píxel pertenece a una microcalcificación, éste píxel se cataloga como Falso Positivo (FP).

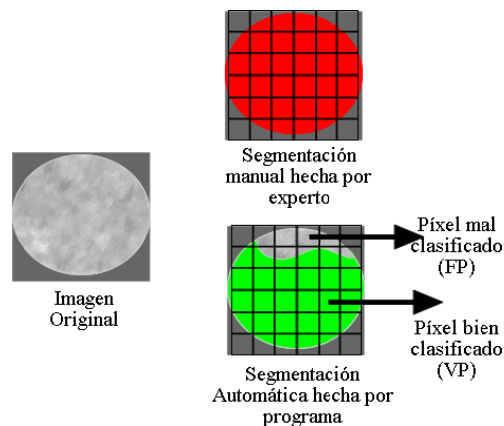


Figura 2.10. Ejemplo de píxeles bien y mal clasificados.

- **Análisis FROC**

El análisis FROC (*Free-response Receiver Operating Characteristic*) se aplica para la evaluación de algoritmos de **detección** de anomalías en mastografías, donde no se evalúa cada píxel, sino que se evalúa a **nivel de región**; ver [8,10].

Las métricas utilizadas para reportar el desempeño de un algoritmo de detección son la Proporción de Verdaderos Positivos (TPR; vea la ecuación 2.7) y el promedio de Falsos Positivos por Imagen (mastografía), (FPI; vea la ecuación 2.8).

$$TPR = \frac{\text{Número de Marcas de Verdaderos Positivos}}{\text{Número de Microcalcificaciones}} \quad (2.7)$$

$$FPI = \frac{\text{Número de Marcas de Falsos Positivos}}{\text{Número de mastografías}} \quad (2.8)$$

Donde una marca de un Verdadero Positivo (VP), se refiere a una marca realizada por el algoritmo de detección que corresponde a la ubicación de una microcalcificación; mientras que una marca de un Falso Positivo (FP), es una marca hecha por el algoritmo de detección que no corresponde a la ubicación de una microcalcificación (vea la figura 2.11).

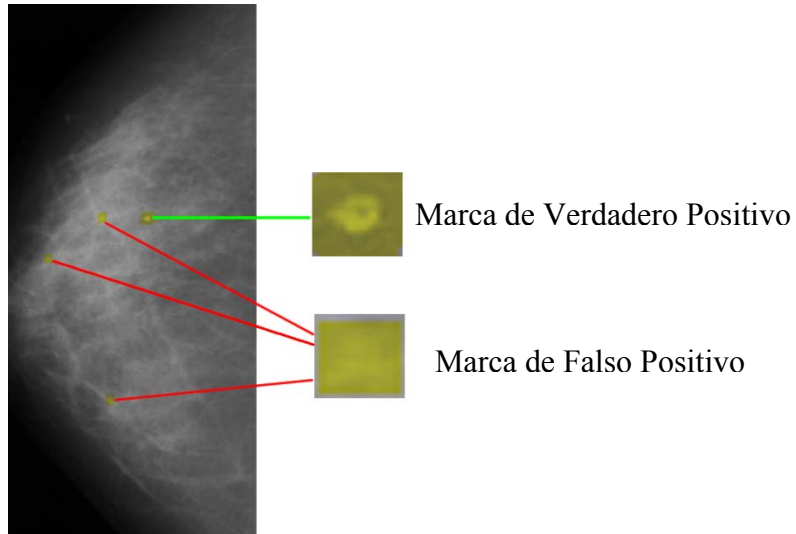


Figura 2.11. Ejemplo de Marcas de Verdaderos y Falsos Positivos.

Es necesario tener en cuenta que para aplicar este tipo de métodos para la evaluación de algoritmos de segmentación y detección, se necesita contar con imágenes que hayan sido marcadas previamente por el experto en el área, de tal forma que se pueda identificar que región pertenece a una microcalcificación y que región no.

### 2.3.5. Extracción de características

Para poder clasificar regiones u objetos en imágenes, es necesario extraer un conjunto de características del objeto o región, que permitan diferenciarlo de otras clases de objetos.

A continuación se describen las características utilizadas en la etapa de reducción de falsos positivos propuesta en esta tesis, dando una breve descripción de cada una de ellas y la forma en que se calculan.

- **Características basadas en la forma de la región**

- **Área:** Es el número total de píxeles en la región , y es calculada mediante la siguiente ecuación:

$$A = |R| \quad (2.9)$$

Donde  $R$  es conjunto de píxeles

- **Diámetro Equivalente:** Especifica el diámetro de un círculo con la misma área que la región
- **Perímetro:** Contiene la distancia alrededor del borde de una región, se obtiene calculando la distancia entre cada par de píxeles unidos que formen parte del contorno de la región
- **Área Convexa:** Escalar que especifica el número de píxeles que están dentro del polígono convexo de una región en una imagen binaria.
- **Excentricidad de la elipse:** Es el radio de la distancia entre el foco de la elipse más pequeña que contiene a la región y la longitud del eje mayor de ésta, este valor se encuentra entre 0 y 1.
- **Área rellena:** Es el área de una región, que toma en cuenta el área de los huecos existentes en ella.
- **Orientación:** Ángulo del eje x y el eje mayor de una elipse de tamaño menor que contenga la región.

- **Extensión:** Indica la proporción de píxeles que se encuentran en la región y además se encuentran dentro del rectángulo más pequeño que contenga a la región.
- **Solidez:** Define la cantidad de píxeles que se encuentran dentro de una región y el área convexa de ésta.
- **Longitud del eje mayor de la región:** Indica la longitud en píxeles del eje mayor de una elipse de tamaño más pequeño en la que cabe la región.
- **Longitud del eje menor de la región:** Indica la longitud en píxeles del eje menor de la elipse de menor tamaño donde cabe la región.
- **Características basadas en el valor de intensidad del píxel**
  - **Nivel mínimo de intensidad:** Escalar que indica el valor del píxel con la intensidad más baja de una región.
  - **Nivel medio de intensidad:** Escalar que indica la media de todos los valores de intensidad de los píxeles dentro de la región.
  - **Nivel máximo de intensidad:** Escalar que indica el valor del píxel con la intensidad más alta de una región.
  - **Desviación Estándar:** Proporciona el nivel de contraste de la región.
  - **Frecuencia:** Cantidad de píxeles que se encuentran dentro de un rango de intensidad de color.

## **2.4. Técnicas para la reducción de la dimensión de un espacio**

Un problema muy estudiado en diversas áreas de las ciencias computacionales es la identificación de patrones a partir de conjuntos de datos de grandes dimensiones; ver [5, 9, 47].

La tarea de encontrar patrones en datos con dimensiones elevadas no es nada fácil, una solución es encontrar un espacio de menor dimensión, en donde los datos se expresen con un menor número de características que permitan describirlos de manera precisa. Existen varios métodos para lograr esto, entre los más utilizados se encuentran el Análisis de Componentes Principales (PCA) y los Discriminantes Lineales de Fisher (FLD). Antes de

describir cómo funcionan estos métodos es necesario recordar algunos conceptos de álgebra lineal: los eigenvectores y eigenvalores.

### 2.4.1. Eigenvectores y Eigenvalores

Un eigenvector se define de la siguiente manera: “Si  $A$  es una matriz de  $n \times n$ , entonces un vector  $v$  diferente de cero en  $R^n$  se denomina eigenvector de  $A$  si  $Av$  es un múltiplo escalar de  $v$ ”, es decir,

$$Av = \lambda v \quad (2.10)$$

Para algún escalar  $\lambda$ . El escalar  $\lambda$  se denomina eigenvalor de  $A$  y se dice que  $v$  es un eigenvector de  $A$  correspondiente a  $\lambda$ ; ver [3].

En otras palabras, un eigenvector es aquel vector que al aplicarle una transformación lineal no modifica su dirección. Estos vectores no se ven alterados por la transformación o pueden ser el resultado del producto de ellos mismos por un escalar. A este escalar se le conoce como eigenvalor.

Por ejemplo sea  $A$  la matriz:

$$A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$$

Un eigenvector de la matriz  $A$  es el vector  $v$ :

$$v = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Utilizando la ecuación 2.10:

$$Av = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 4v$$

Así, el eigenvalor del eigenvector  $v$  es  $\lambda = 4$ , en la figura 2.12 se observa gráficamente el eigenvector  $v$  y su transformada.

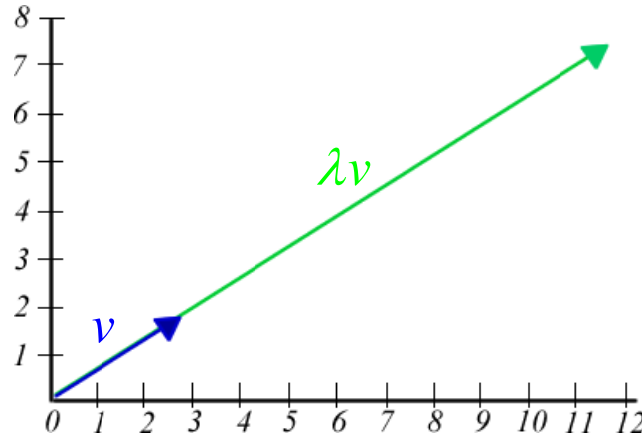


Figura 2.12. Representación gráfica del eigenvector  $v$  y su transformación, con este ejemplo se puede observar que aplicando la transformación al vector, la magnitud de éste cambia pero su dirección es la misma.

### 2.4.2. Análisis de Componentes Principales (PCA)

El análisis de componentes principales es una técnica estadística para la reducción de la dimensión de un espacio (número de variables o características). En el área de aprendizaje automático, a partir de  $N$  ejemplos representados por  $n$  variables se analiza si es posible representar adecuadamente esa información por un número menor de variables construidas a partir de combinaciones lineales de las variables originales; ver [26, 43].

La idea básica de este método es que a partir de un conjunto de datos, PCA encuentra el mejor conjunto de direcciones de proyección  $U_{opt}$ , en que se maximice la **varianza o dispersión total** de los datos (vea la figura 2.13) tal que:

$$U_{opt} = \arg \max_U |U^T S_t U| \quad (2.11)$$

Donde  $S_t$  es la matriz de dispersión o covarianza resultante de los ejemplos de entrenamiento  $(X_1, X_2, \dots, X_N)$ , la cual está definida de la siguiente forma:

$$S_t = \sum_{i=1}^N (X_i - \mu)(X_i - \mu)^T \quad (2.12)$$

Siendo  $\mu$  la media de los datos.

Para hacer la reducción de espacio, utilizando la matriz de transformación  $U_{opt}$  se seleccionan los  $m$  eigenvectores de  $S_t$  correspondientes a los  $m$  eigenvalores más altos, siendo  $m < n$ . Así, los datos son proyectados en el sub-espacio formado por los componentes principales con mayor longitud, es decir, en dirección donde la varianza total de los datos es mayor.

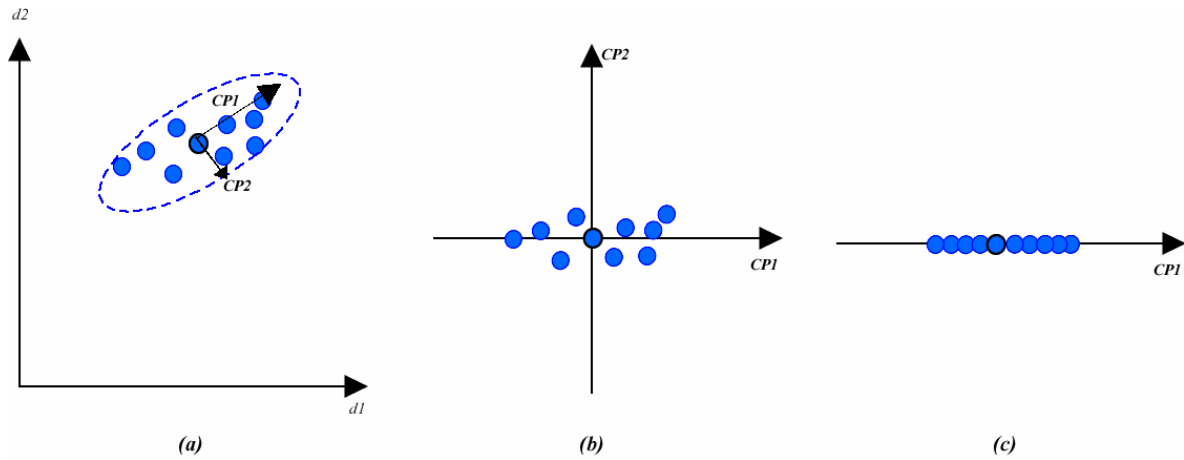


Figura 2.13. Reducción del espacio mediante PCA, a) datos representados en un espacio 2-dimensional, b) Datos representados en espacio formado por los 2 componentes principales ortogonales (CPs), ordenados por varianza c) proyección de los datos en el eje del CP con eigenvalor mayor, es decir eje que representa mayor varianza de los datos.

La desventaja de este método es que no diferencia entre datos etiquetados pertenecientes a diferentes clases ya que se basa en la varianza general de todos los datos.

### 2.4.3. Discriminantes Lineales de Fisher (FLD)

En el área de aprendizaje computacional y reconocimiento de patrones, la finalidad de realizar cualquier análisis discriminante, es encontrar propiedades que permitan diferenciar objetos que pertenezcan a diferentes clases. El objetivo de los Discriminantes Lineales de Fisher, es realizar una reducción de dimensión preservando la mayoría de la información discriminante de las clases; ver [23, 45, 51].

Si las clases pueden ser separadas mediante la combinación lineal de características, entonces se puede utilizar un modelo discriminante lineal como es el caso de FLD.

A diferencia de PCA, FLD selecciona la transformación lineal de los datos donde el radio de la dispersión entre clases (*between-class scatter matrix*) sea maximizado y el radio



de la dispersión dentro de las clases (*within-class scatter matrix*) sea minimizada (vea la figura 2.14).

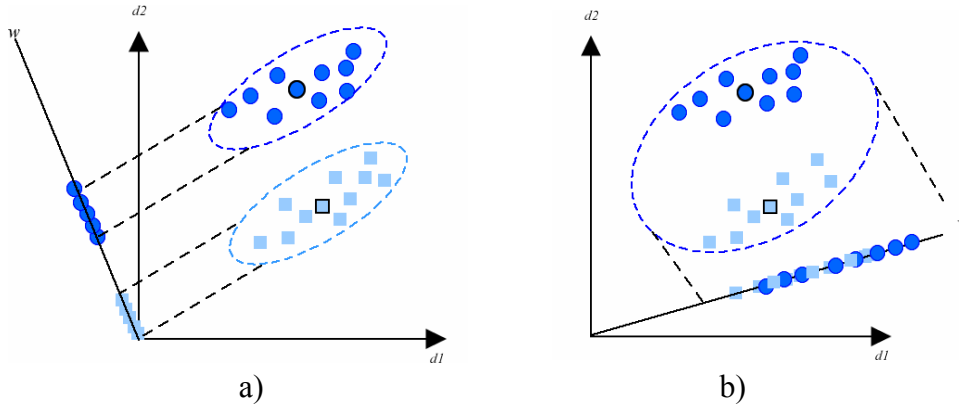


Figura 2.14. (a) Proyección mediante FLD que maximiza la distancia entre las medias proyectadas de las clases y minimiza la dispersión proyectada dentro de las clases ( b) Proyección mediante PCA que no sirve para discriminar entre clases ya que solamente toma la dispersión total de los datos.

Dado un conjunto de  $N$  ejemplos  $\{x_1, x_2, \dots, x_N\}$ , los cuales están representados en un espacio  $n$ -dimensional, y donde cada uno pertenece a una de las  $k$  clases  $\{X_1, X_2, \dots, X_k\}$ . La matriz de dispersión entre clases está definida como:

$$S_b = \sum_{i=1}^k N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (2.13)$$

Y la matriz de dispersión dentro de las clases está definida por

$$S_w = \sum_{i=1}^k \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T \quad (2.14)$$

Donde  $\mu$  es la media general y  $\mu_i$  es la media de la clase  $i$  la cual está definida por

$$\mu_i = \frac{1}{n_i} \sum_{x_j \in X_i} x_j \quad i = 1, 2. \quad (2.15)$$

Por lo tanto, FLD selecciona la transformación lineal  $U$  que maximice la distancia entre las medias de las clases proyectadas y que minimice la varianza dentro de las clases proyectadas.

$$U_{opt} = \arg \max_U \frac{|U^T \cdot S_b \cdot U|}{|U^T \cdot S_w \cdot U|} \quad (2.16)$$

Si  $S_w$  es una matriz de rango completo la transformación lineal óptima  $U_{opt}$  está formada por los  $m$  eigenvectores con los eigenvalores más altos del producto de  $(S_w^{-1})(S_b)$ , es decir,

$$S_b S_w^{-1} u_i = \lambda_i u_i \quad (2.17)$$

Una vez que se ha encontrado la transformación lineal óptima, los ejemplos se proyectan a este nuevo espacio de la siguiente manera,

$$W = U_{opt}^T (x_i - \mu) \quad (2.18)$$

Cuando se desea clasificar un objeto nuevo cuya clase se desconoce, este objeto se proyecta al espacio, se mide la distancia Euclidiana con cada una de la medias de las clases proyectadas y se le asigna la clase de la media a la que se encuentre más cercano.

## **Capítulo 3**

### **Estado del Arte**

#### **3.1. Introducción**

El cáncer de seno es una de las principales causas por la que miles de mujeres en todo el mundo mueren cada año. Se ha demostrado que algunos tipos de microcalcificaciones son indicadores de la presencia de cáncer de seno en etapas tempranas, cuando el tumor canceroso aún no es visible. Sin embargo, detectar estas pequeñas acumulaciones de calcio, con la apariencia de pequeños puntos brillantes en las mastografías, está muy lejos de ser un proceso trivial. Su diminuto tamaño, sus diversas formas y el poco contraste que existe entre ellas y el tejido que las rodea, son algunos de los principales problemas para detectar microcalcificaciones.

A nivel internacional, diferentes grupos de trabajo han desarrollado diversas técnicas y algoritmos para la detección y clasificación de microcalcificaciones. La meta de los algoritmos de detección es ayudar al radiólogo a localizar anomalías (microcalcificaciones, tumores, quistes) en la mastografía. La mayoría de estos algoritmos de detección, constan de 2 etapas principales; ver [8].

La primera etapa se enfoca en detectar regiones sospechosas con una tasa alta en la proporción de verdaderos positivos (TPR). En la segunda etapa el objetivo es reducir el número de falsos positivos sin disminuir drásticamente la TPR. Si además de detectar (o segmentar) microcalcificaciones, éstas son clasificadas como benignas o malignas, este tipo de algoritmos son llamados de diagnóstico y éstos pueden partir de regiones de interés (ROIs) (segmentadas manualmente para después clasificarlas), o pueden realizar las dos etapas de manera automática para primero detectar y obtener las ROIs a partir de la imagen obtenida y posteriormente clasificar (vea la figura 3.1).

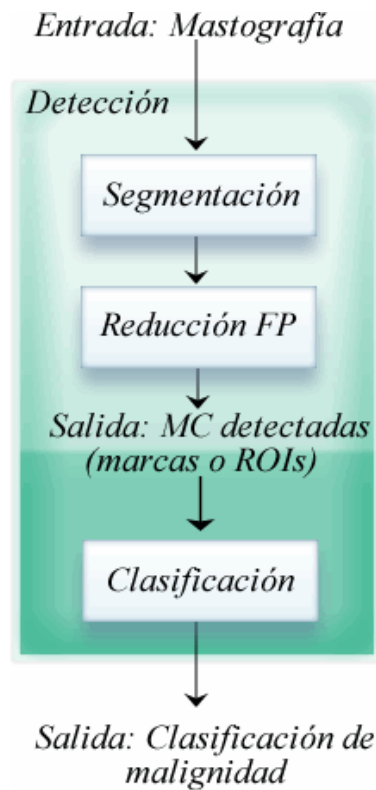


Figura 3.1. Diagrama de Flujo que muestra las principales etapas de los algoritmos de detección y de diagnóstico de anomalías en el seno.

Las técnicas de segmentación pueden ser de dos tipos: supervisadas o no supervisadas; ver [19]. La segmentación supervisada, también conocida como segmentación basada en modelos, depende de conocimiento previo acerca de lo que se va a segmentar. La información previa se utiliza para determinar si regiones específicas están presentes dentro de la imagen o no. La segmentación no supervisada consiste en dividir la imagen en un

conjunto de regiones que están formadas por píxeles que tienen propiedades similares, tales como nivel de gris, textura o color. Los enfoques clásicos para realizar segmentación no supervisada se pueden dividir en tres grandes grupos:

- Métodos basados en región; los cuales dividen la imagen en regiones homogéneas y espacialmente conectadas.
- Métodos basados en contorno; los cuales se basan en encontrar los bordes o límites de las regiones.
- Métodos de agrupamiento y basados en umbrales, donde se agrupan aquellos píxeles que tienen propiedades similares y pueden resultar en regiones no conectadas.

### 3.2. Métodos basados en Regiones

La segmentación basada en regiones depende del principio de homogeneidad, que significa que existe al menos una característica que permanece uniforme (más/menos una medida de error) para todos los píxeles de una región. **Morrow et al.** utilizan un procesamiento de imágenes basado en regiones, esta técnica adapta las características de la región y resalta estas características con respecto a su alrededor, sin importar el tamaño o forma de éstas. Morrow considera regiones que se traslapan, su método utiliza cada píxel en la imagen como una semilla que crece para convertirse en región. El crecimiento y forma de la región se adapta de acuerdo a la variación local del nivel de gris. Posteriormente se calcula el contraste de cada región con respecto a su fondo. Este método tiene la ventaja de que al tomar en cuenta el contraste de cada región con el fondo, genera buenos resultados en imágenes de mastografías difíciles de segmentar (senos densos), la desventaja es que es costoso en tiempo y espacio; ver [33].

**Shen et al.** utilizan un método basado en crecimiento de regiones, este método requiere de dos variables, el tamaño de la ventana y la diferencia absoluta entre el píxel procesado y el píxel semilla. Una vez que el algoritmo de crecimiento de regiones termina, si el promedio de la intensidad de la región obtenida es mucho mayor a la región que lo rodea, el

píxel es clasificado como un píxel de la microcalcificación. Este método busca un cambio de intensidad para detectar la microcalcificación pero utilizando este criterio es fácil ver que en senos muy densos, donde no existe una gran diferencia de intensidad entre el tejido y la microcalcificación, no funciona adecuadamente; ver [42].

**Wallet et al.** proponen un método que consta de dos etapas principales, en la primera etapa utilizan un filtro para atenuar algunas señales de la imagen (high boost). En la segunda etapa utilizan un umbral (calculado con el promedio del nivel de gris de los píxeles de la región de interés) y aplican una técnica basada en crecimiento de regiones para extraer posibles microcalcificaciones. Aunque este método también mostró buen desempeño en detectar microcalcificaciones en senos densos, en comparación con otros métodos basados en umbralización; genera muchos falsos positivos. Esto puede ser causado por el umbral que utilizan para la selección de la semilla original; ver [35, 49].

**Woods et al.** utilizan un umbral local en la etapa de pre-procesamiento para restar la intensidad promedio de una ventana de 15x15 píxeles alrededor del píxel que se está analizando y posteriormente utilizan crecimiento de regiones para unir píxeles que comparten propiedades (similares) y así formar regiones; ver [11,54].

**Kim y Park** proponen un método que calcula el histograma de segundo orden o matriz de coocurrencia a partir de dos regiones de tamaño fijo que rodean a un píxel, (SRDM). La matriz de coocurrencia permite considerar no solamente la distribución de las intensidades de una imagen, sino también las posiciones de píxeles que tienen valores de intensidad casi iguales. Mediante este proceso identifican características de textura de la región de interés con el fin de detectar si en esa región existe o no microcalcificación. Este trabajo mostró ser efectivo para localizar microcalcificaciones en mastografías con diferentes densidades en el seno. Esto debido a que toma en cuenta el contraste basado en homogeneidad difusa que utiliza información global y local de la región. De manera general, los métodos basados en regiones tienen problemas en la selección adecuada de la semilla y la determinación del parámetro de paro de crecimiento de las regiones, así como el alto costo en tiempo y memoria; ver [27,46].

### 3.3. Métodos basados en Contornos

Identificar regiones basándose en información de los bordes no es una tarea sencilla, debido a que los algoritmos de detección de bordes generalmente no poseen la habilidad del sistema de visión humano para completar bordes discontinuos utilizando experiencia e información contextual. Por esta razón, algunas veces los bordes detectados tienen espacios y pueden no necesariamente formar un conjunto de curvas conectadas cerradas que rodean diferentes regiones.

**Fu et al.** diseñaron un método para la detección de microcalcificaciones basado en dos etapas. En la primera etapa utiliza un modelo matemático para calcular la ubicación y la forma de las microcalcificaciones. Este modelo se basa en aplicar la transformada Top-Hat para detectar bordes de objetos que diferencian en luminosidad del fondo que los rodea. Posteriormente aplican Sobel y Canny para segmentar los bordes resaltados. En la segunda etapa se extraen 61 características (de textura y espaciales) y mediante SFS (sequential forward search) se seleccionan solamente las características que mejor describan las microcalcificaciones. Por último, utilizan una red neuronal de regresión general (GRNN) y una Máquina de Vectores de Soporte (SVM) para reducir el número de falsos positivos. Una de las desventajas de este método es que en imágenes donde la diferencia de luminosidad entre la microcalcificación y su fondo no excede cierto umbral, las microcalcificaciones no pueden ser detectadas, como sería el caso de senos que tienen una gran cantidad de tejido glandular (senos densos), por otra parte; de manera inversa, en senos grasos donde los conductos que se encargan de transportar la leche hacia el pezón son muy notorios, éste método para la detección de microcalcificaciones genera muchos falsos positivos; ver [18].

Otro método propuesto es el de **Lee y Tsai**, que de manera similar al anterior utiliza la transformada Top-Hat pero ésta no toma como entrada la mastografía completa sino que utiliza imágenes de cierto tamaño en las que se sabe que existen microcalcificaciones (ROIs). Posteriormente elimina aquellas regiones detectadas que tienen un valor de píxel muy bajo para finalmente eliminar aquellos objetos que midan menos de 3 píxeles de diámetro. En la segunda etapa extraen 4 características (número total de

microcalcificaciones, media, área, distancia media) y finalmente clasifican en maligno o benigno utilizando un algoritmo genético basado en lógica difusa. En este caso la etapa de segmentación es muy simple, parten de una imagen segmentada manualmente en la que se sabe que existen microcalcificaciones, por lo que identificar si realmente funciona en una mastografía real es difícil. Por otra parte eliminan regiones en las que el valor de sus píxeles es bajo, pero hay que recordar que el nivel de gris de una microcalcificación depende de la composición química de ésta y de la cantidad de rayos x que pasan por ella, por lo que puede ser que pasen por alto regiones que si pertenecen a las microcalcificaciones; ver [13, 29].

**Zhao et al.** desarrollaron un método para extraer microcalcificaciones basado en un umbral morfológico adaptable y también utiliza información morfológica del esqueleto de la región. El criterio para seleccionar el umbral es que el esqueleto debe coincidir con la descripción del tamaño de la microcalcificación; ver [59].

**Wu et al.** aplican el operador Laplacian of Gaussian (LoG) para detectar bordes, éste se basa en suavizar la imagen con un filtro Gaussiano antes de encontrar los bordes. Posteriormente utilizan dos umbrales, el primero es para encontrar los contornos obvios y el segundo para decidir si el contorno encontrado debe quedarse o no basándose en el número total de puntos que lo conforman, en este caso no utilizan ningún clasificador. El problema aquí es que en imágenes en que no es obvio el contorno de las microcalcificaciones, es difícil que las detecte; ver [55].

### **3.4. Métodos de agrupamiento y basados en umbrales**

Los métodos de agrupamiento son de los más utilizados en algoritmos de segmentación, los métodos basados en umbralización se consideran como parte de los métodos de agrupamiento donde solo se consideran dos clusters. Debido a que las microcalcificaciones representan frecuencias espaciales altas en la imagen, un enfoque bastante utilizado para su detección es la transformada wavelet.

**Wang y Karayiannis** presentan un enfoque para detectar microcalcificaciones empleando descomposición de imagen en sub-bandas basado en wavelets. Las



microcalcificaciones aparecen en pequeños clusters de pocos píxeles con una intensidad relativamente alta comparada con los píxeles vecinos. Dado que las microcalcificaciones corresponden a componentes de alta frecuencia del espectro de la imagen, descomponen la mastografía en diferentes sub-bandas de frecuencia, eliminando aquellas de baja frecuencia. Finalmente reconstruyen la imagen a partir de todas las sub-bandas que contienen solo frecuencias altas; ver [50].

De igual manera **Flores y González** aplican una transformada wavelet, pero además; utilizan un umbral global y reducen el porcentaje de falsos positivos eliminando regiones que sean más pequeñas que un cierto umbral fijo. Una vez encontradas las ROIs segmentan sus bordes y extraen 11 características (geométricas y de textura), utilizan un clasificador basado en redes neuronales y otro basado en árboles de decisión; ver [16].

Otro método basado en wavelets es el creado por **Mousa et al.** donde primero eliminan el ruido y separan el seno del fondo, seleccionan un umbral global máximo y uno mínimo de manera fija. Posteriormente aplican el proceso de descomposición wavelet. En la etapa de la clasificación introducen las ROIs a un algoritmo de aprendizaje basado en lógica difusa (ANFIS); ver [34].

En general, el problema de utilizar cualquier tipo de wavelets es que genera demasiados falsos positivos debido a que resalta las señales de alta frecuencia pero el ruido también tiene un papel significativo en altas frecuencias.

Otra técnica para detectar microcalcificaciones que está basada en umbrales es la de **Nishikawa et al.**, la cual consta de 3 etapas. La primera resalta las microcalcificaciones para eliminar el fondo de la mastografía mediante un filtro, en la segunda etapa las posibles microcalcificaciones son identificadas por medio de un umbral global el cual está basado en el nivel de gris, erosión morfológica y un umbral local adaptable. En la etapa 3 se reduce el número de falsos positivos mediante el análisis de distribución espacial y la relación entre el tamaño y la forma de las microcalcificaciones; ver [36].

**Davis y Dance** segmentan calcificaciones del tejido del seno mediante el uso de un umbral de área local. La imagen original se divide en sub-imágenes y el histograma del nivel de gris se suaviza mediante un filtro que remueve el nivel máximo y mínimo de éste. El histograma es clasificado como bimodal o unimodal, en caso de ser bimodal el umbral es asignado como el valle del histograma mientras que en caso de ser unimodal, al umbral se

le asigna el valor máximo de gris de esa área. Una vez que todas la sub-imágenes han sido procesadas, cada umbral unimodal es remplazado por un valor calculado mediante la interpolación de los umbrales de las sub-imágenes vecinas; ver [14].

El problema principal de los métodos basados en umbrales es que funcionan adecuadamente para un conjunto pequeño de imágenes, pero elegir un umbral que funcione para un conjunto grande de imágenes es extremadamente difícil.

### **3.5. Métodos de segmentación supervisada**

Los métodos de segmentación supervisada incluyen aquellos métodos que primero entrenan al sistema para reconocer objetos específicos. Posteriormente el sistema es capaz de detectar y clasificar nuevas imágenes dependiendo de la presencia o ausencia del objeto. **L. Zhang et al.** desarrollaron un método para la detección de microcalcificaciones, el cual está formado por dos etapas. En la primera etapa utiliza un conjunto de características de microcalcificaciones como entrada a una red neuronal para reducir la proporción de detección falsa. En la segunda etapa dos características como el tamaño de la región y forma de la región son también utilizadas como entrada de la red neuronal para reducir aún más el número de falsos positivos. Las características que utilizan, las dividieron en dos categorías: espaciales y morfológicas. El primer conjunto de características incluye el nivel promedio de gris del fondo y la desviación estándar del nivel de gris del fondo. El segundo grupo de características incluye que tan compacta es la región y un descriptor Fourier; ver [57].

**Nishikawa et al.** desarrollaron un enfoque basado en Máquinas de Vectores de Soporte (SVM) para la detección de grupos de microcalcificaciones en mastografías digitales. SVM es un método de aprendizaje computacional, el cual ha mostrado un buen desempeño con los datos fuera del conjunto de entrenamiento. Por esta razón decidieron aplicar los SVM para la detección de microcalcificaciones. Utilizaron los SVM para detectar en cada posición de la imagen si una microcalcificación está presente o no; ver [35].

**Choi et al.** seleccionaron 7 características a partir de un conjunto de 29 para probar con 7 clasificadores. El clasificador que mejores resultados obtuvo fue el clasificador lineal,

mientras que las características que se seleccionaron fueron: área del objeto, nivel de gris promedio del objeto, intensidad del gradiente de los píxeles del perímetro del objeto, contraste, RMS (root mean square) de la fluctuación del ruido, fluctuación del ruido del fondo local en un área de 3.5mm x 3.5mm y momento de orden bajo basado en un descriptor de forma; ver [12].

**W. Zhang et al.** desarrollaron una técnica para optimizar los pesos en escalas individuales de wavelets para mejorar el desempeño de la detección de microcalcificaciones basado en un método de aprendizaje supervisado. En el proceso de aprendizaje, se formuló una función de error para representar la diferencia entre la salida deseada y la imagen reconstruida obtenida a partir de los pesos de los coeficientes de la wavelet para una mastografía. Posteriormente la función de error fue minimizada modificando los pesos de los coeficientes por medio de un algoritmo de gradiente conjugado; ver [58].

La mayoría de los algoritmos basados en segmentación supervisada (revisados en la literatura) han reportado buenos resultados para clasificar regiones como microcalcificaciones y regiones que son tejido del seno; ver [11].

Estos métodos hacen uso de algoritmos de aprendizaje y cuentan con ejemplos que pertenecen a microcalcificaciones y ejemplos pertenecientes al tejido del seno. El problema principal de este tipo de algoritmos es que necesitan ejemplos para entrenar al sistema y en ocasiones es difícil conseguir estos ejemplos. Las características ejemplos que utilizan para entrenar pueden influir en el buen funcionamiento del algoritmo, es decir, si en la etapa de entrenamiento utilizan ejemplos donde tienen un cierto rango de nivel de gris y toman en cuenta características basadas en la intensidad de gris, cuando se analiza una nueva mastografía, si tiene un rango de intensidad diferente debido por ejemplo a que fue tomada con otro aparato de rayos X, la clasificación tiene muchas probabilidades de dar un resultado erróneo.

### 3.6. Discusión

Como se vio en las secciones anteriores de este capítulo, **la mayoría de los procesos de detección tienen limitaciones** específicas y generales, a continuación se presenta una tabla en que se mencionan las ventajas y desventajas de cada enfoque.

Tabla 3.1. Tabla comparativa de desventajas y ventajas de métodos de segmentación utilizados para la detección de microcalcificaciones.

| Segmentación   | Métodos                | Ventajas / Desventaja   | Referencias  |
|----------------|------------------------|---|--|
| No-Supervisada | Regiones               | <ul style="list-style-type: none"> <li>- Trabaja bien cuando el criterio de homogeneidad de la región es fácil de definir.</li> <li>- Depende de la selección de la semilla y del criterio de paro.</li> <li>- Costoso en tiempo y espacio.</li> </ul>  | Morrow et al. [33]<br>Shen et al. [42]<br>Wallet et al. [49]<br>Woods et al. [54]<br>Kim y Park [27]     |
|                | Contornos              | <ul style="list-style-type: none"> <li>- Son buenos para lidiar con aspectos geométricos del análisis de imágenes</li> <li>- Los contornos no siempre son fáciles de encontrar sobre todo cuando no hay suficiente contraste con el fondo</li> <li>- No siempre son contornos cerrados, pueden tener espacios.</li> </ul> | Fu et al. [18]<br>Lee y Tsai [29]<br>Zhao et al. [59]<br>Wu et al. [55]                                  |
|                | Agrupamiento/ Umbrales | <ul style="list-style-type: none"> <li>- No necesariamente son regiones conectadas.</li> <li>- Dependen de la selección adecuada de un umbral, por lo que no funcionan para</li> </ul>  | Wang y Karayiannis [50]<br>Flores y Gonzalez [16]<br>Mousa [34]<br>Nishikawa [36]<br>Davies y Dance [14] |

|             |  |   |   |
|-------------|--|---|---|
|             |  | bases de imágenes grandes.<br>- Los métodos basados en wavelets generalmente generan muchos Falsos Positivos (ruido).   |   |
| Supervisada |  | - La selección adecuada del conjunto de características genera buenos resultados.<br>- Requieren de un conjunto grande de ejemplos.<br>- Requieren de información a priori. | L. Zhang et al. [57]<br>Nishikawa et al. [36]<br>Choi et al. [12]<br>W. Zhang et al. [58] |

A pesar de las ventajas y desventajas que tienen los métodos mencionados anteriormente, **un buen detector de microcalcificaciones debería tener las siguientes propiedades:**

1. Debe poder adaptarse al nivel de contraste que lo rodea. Las calcificaciones que son muy brillantes y que se encuentran en un fondo también brillante, deben poder detectarse, como lo es en senos densos donde existe muy poco contraste entre las microcalcificaciones y el tejido que los rodea. De igual manera en zonas donde el nivel de densidad sea muy poco y la intensidad del color de las microcalcificaciones es aún menor, éstas deben poder detectarse.
2. Como el tamaño de las microcalcificaciones está dentro de un rango conocido, pero puede variar, debe poder detectar microcalcificaciones que estén dentro de un rango de tamaño pero no de un solo tamaño en específico.
3. Existen microcalcificaciones de muchas formas, por lo que debe ser capaz de detectar microcalcificaciones de cualquier forma.
4. Debido a las diferentes composiciones químicas de las microcalcificaciones, éstas permiten pasar cantidades diferentes de rayos X, por lo que éstas presentan

diferentes niveles de contraste en una mastografía. Un buen detector debe ser insensible a las grandes variaciones de intensidad de color de las microcalcificaciones; ver [15].

Estas propiedades nos guiaron para construir un método para la detección de microcalcificaciones basado en segmentación supervisada. Utilizando los Discriminantes Lineales de Fisher se propone desarrollar un detector de microcalcificaciones que sea insensible a las diferentes variaciones de intensidad, tamaños y formas que presentan las microcalcificaciones. Para desarrollar un detector que además de lo anterior cumpla con la segunda propiedad deseada, es decir, que sea capaz de identificar microcalcificaciones tanto en senos densos donde existe muy poco contraste y también en senos grasos. En esta tesis también se propone tomar en cuenta la clasificación de la densidad del seno, ya que como se ha analizado anteriormente, no es lo mismo encontrar una microcalcificación en un seno con mucho tejido glandular donde casi no existe contraste que en uno donde el seno tiene muy poco tejido glandular (graso).

Ya que uno de los problemas a resolver en todos los enfoques de detección de microcalcificaciones es la gran proporción de falsos positivos, también se propone agregar un módulo específico para la reducción de falsos positivos basado en algoritmos de aprendizaje computacional.

En el siguiente capítulo se describe la metodología y el esquema general del método propuesto en esta tesis para la detección de microcalcificaciones en mastografías digitales.

## **Capítulo 4**

### **Detección de microcalcificaciones utilizando FLD**

#### **4.1. Introducción**

En el capítulo 3 se analizaron diferentes métodos y algoritmos para la detección de microcalcificaciones, pero hasta el momento ninguno de ellos ha demostrado ser una solución eficaz para este problema.

La solución propuesta en esta tesis, para la detección de microcalcificaciones, cae dentro del grupo de algoritmos basados en segmentación supervisada, es decir, se cuenta con una etapa de entrenamiento y una etapa de clasificación.

Nuestro método permite detectar microcalcificaciones de diferentes formas, tamaños y tonalidades, además es capaz de identificarlas tanto en senos grasos como en senos densos (catalogados como senos difíciles). Para la codificación del método se utilizó Matlab R2008 y Weka 3.4.12; ver [52].

El esquema general del método (vea la figura 4.1) se puede dividir en tres etapas principales: la primera etapa se enfoca a la clasificación de la densidad del seno, la segunda etapa está dedicada a la detección de microcalcificaciones haciendo uso de FLD y la última

etapa tiene el objetivo de reducir el número de falsos positivos mediante la extracción de características descriptivas de las microcalcificaciones y el uso de clasificadores.

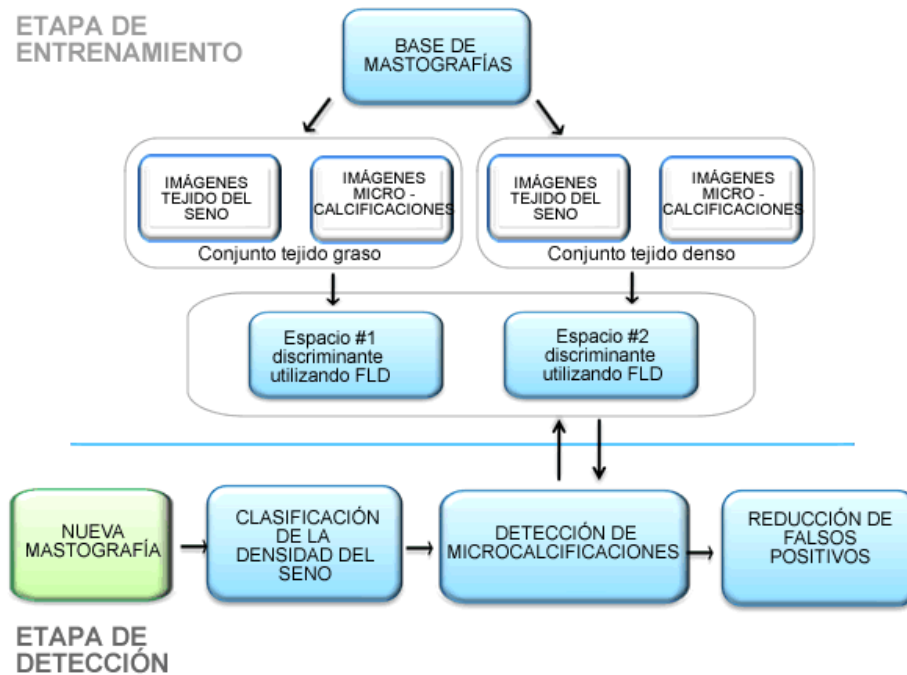


Figura 4.1. Esquema general del método propuesto en esta tesis para la detección de microcalcificaciones en mastografías digitales.

El procedimiento para detectar las microcalcificaciones en una mastografía, comienza con la clasificación del seno de acuerdo al tejido por el que principalmente está formado, ya sea graso o glandular. Si el seno contiene más tejido graso que glandular éste se cataloga como seno graso, de lo contrario se clasifica como seno denso. A partir de esta clasificación, el sistema selecciona uno de los dos espacios discriminantes que permiten diferenciar entre imágenes de microcalcificaciones e imágenes de tejido del seno. Estos espacios fueron generados mediante los discriminantes lineales de Fisher (FLD) y se crearon en la etapa de aprendizaje de nuestro método.

Posteriormente, en la etapa de segmentación, se utiliza una ventana de tamaño fijo para analizar el seno en busca de microcalcificaciones. La región dentro de la ventana se proyecta al espacio discriminante correspondiente, y se identifica si ésta corresponde a una microcalcificación o si se trata de tejido del seno. En caso de que la región se identifique como microcalcificación, se resalta esa zona en la mastografía, con el fin de indicar la



existencia de una posible microcalcificación. Este proceso se repite hasta que la ventana haya recorrido y analizado todo el seno.

En la etapa final, la cantidad de falsos positivos generada en la etapa de segmentación se reduce mediante la extracción de características descriptivas de cada una de las posibles microcalcificaciones y mediante el uso de clasificadores de Weka.

En las siguientes secciones se describen de forma más detallada cada una de las etapas del método propuesto para la detección de microcalcificaciones mediante FLD.

## **4.2. Clasificación de densidad del seno**

El problema de clasificar el seno de acuerdo a su densidad es un tema de gran interés y bastante complejo. Estudios han demostrado que existe una gran relación entre la densidad del seno y el riesgo de desarrollar cáncer de seno; ver [32,44].

Por esta razón varios grupos de investigación están interesados en este problema. Se han propuesto diversas clasificaciones para los senos de acuerdo a su densidad. Entre los más importantes se encuentran las clasificaciones propuestas por Wolfe (4 categorías), ACR (4 categorías) y MIAS (3 categorías); (para más detalle vea la sección 2.1.3).

En nuestro caso, el tipo de tejido por el que mayormente está compuesto el seno nos interesa, debido principalmente a que en senos que están mayormente compuestos por tejido glandular, la diferencia de contraste de las microcalcificaciones y el tejido que las rodea es pequeño; mientras que en senos grasos el nivel de contraste es mucho mayor (vea la figura 4.2). Debido a esto, es fácil notar que existe una diferencia en localizar microcalcificaciones en senos densos que localizarlas en senos grasos.

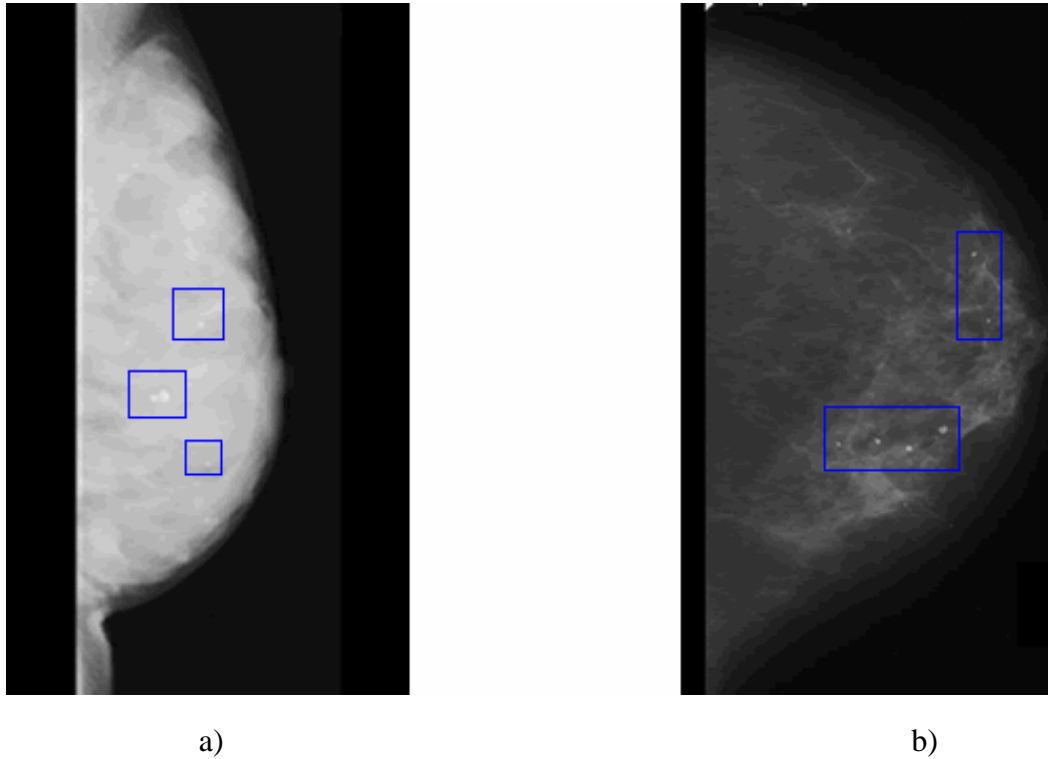


Figura 4.2. Ejemplos de senos que presentan microcalcificaciones. a) Seno Denso y b) Seno Graso.

Para el propósito de esta tesis, solo nos interesa identificar cuál es el tejido que predomina en el seno (denso o graso), por esta razón, solo se utilizaron dos categorías: Seno Denso y Seno Graso.

El esquema general del algoritmo propuesto para la clasificación de la densidad del seno se puede ver en la figura 4.3

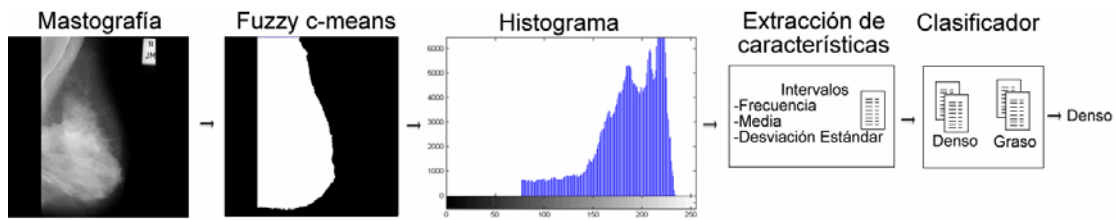


Figura 4.3. Esquema del algoritmo de clasificación de densidad del seno.

En la primera etapa del algoritmo para la clasificación de densidad del seno, se separa el seno del fondo de la mastografía mediante el algoritmo Fuzzy c-means, este algoritmo fue seleccionado ya que mostró obtener buenos resultados para esta tarea; ver [38].

El algoritmo Fuzzy *c-means* genera  $n$  grupos o regiones (en nuestro caso 2 grupos), un grupo está formado por aquellos píxeles que forman el seno y otro grupo está formado por aquellos píxeles que forman parte del fondo de la mastografía. Debido a que algunas mastografías tienen etiquetas con información del paciente, en ocasiones los píxeles pertenecientes a las etiquetas se agrupan como parte del seno. Para evitar que las etiquetas también sean tomadas como parte de la región del seno, la imagen se divide en sub-regiones y solo se conserva aquella región de mayor tamaño (seno), las regiones de menor tamaño (etiquetas) son eliminadas (vea la figura 4.4).

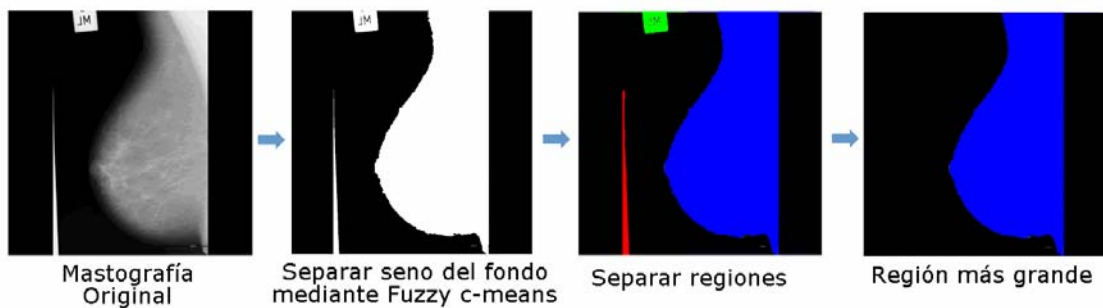


Figura 4.4: Segmentación del seno mediante el uso de Fuzzy c-means.

Una vez que el seno ha sido separado del fondo de la mastografía, se calcula y se analiza el histograma de la imagen del seno. El histograma se divide en 7 intervalos (50-99, 100-150, 151-170, 171-190, 191-210, 211-230, 231-256) y para cada uno de ellos se calculan las siguientes características: media, frecuencia y desviación estándar, todas éstas obtenidas de manera proporcional al tamaño del seno. La media proporciona el nivel promedio de luminosidad del intervalo, la frecuencia nos proporciona el número de píxeles que están dentro de ese intervalo de color y finalmente la desviación estándar proporciona la diferencia de contraste dentro del intervalo.

La selección de los 7 intervalos se realizó de manera experimental, al realizar diferentes pruebas se identificó que dividir el histograma en intervalos y extraer atributos para cada intervalo proporciona información más significativa que utilizar las características extraídas del histograma completo.

A partir de estas características se crea una base de datos para cada una de las bases de mastografías (MIAS e ISSSTEP), las cuales son utilizadas en la etapa de entrenamiento de algunos clasificadores de Weka que permiten determinar la categoría del seno, ya sea denso o graso. Es importante mencionar que la clasificación de las mastografías (densa o grasa) obtenida en esta etapa es utilizada en la etapa de segmentación.

### 4.3. Espacios discriminantes y segmentación de microcalcificaciones en mastografías digitales mediante FLD

La segunda etapa de nuestro método consiste en la segmentación de microcalcificaciones mediante el uso de los discriminantes lineales de Fisher (FLD). El objetivo de utilizar FLD es generar un espacio discriminante de menor dimensión, el cual permita discriminar entre imágenes de microcalcificaciones e imágenes de tejido del seno (vea la figura 4.5).

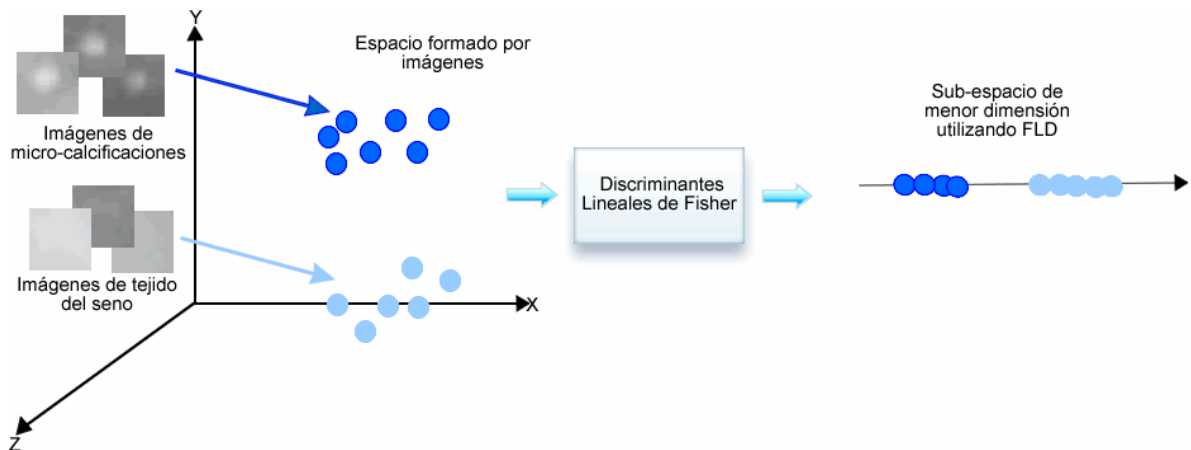


Figura 4.5: Ejemplo para representar como funciona FLD. Con la finalidad de hacer una descripción gráfica, suponemos que las imágenes originalmente están representadas en un espacio 3 dimensional, (aunque de manera real se representan en un espacio n-dimensional) y aplicando FLD se llevan a un espacio de menor dimensión.

#### 4.3.1. Bases de imágenes de microcalcificaciones e imágenes de tejido del seno

En la etapa de aprendizaje del método de detección se crean dos espacios discriminantes, uno para senos densos y otro para senos grasos. Para generar estos espacios se crearon dos bases de imágenes de microcalcificaciones, una base contiene imágenes de microcalcificaciones provenientes de senos grasos y otra base contiene imágenes de microcalcificaciones de senos densos. Cada base cuenta con ejemplos de imágenes de microcalcificaciones y con ejemplos de imágenes de tejido del seno, todas las imágenes son de tamaño de 12 x 12 píxeles, esto debido a que los diversos tamaños que pueden presentar las microcalcificaciones caben dentro de este tamaño de imagen; por otra parte las imágenes tienen formato jpg.

Las imágenes fueron obtenidas de manera manual, solamente las imágenes de microcalcificaciones fueron ajustadas manualmente para que el centroide de la región de la microcalcificación esté aproximadamente en el centro de la imagen, este procedimiento se realizó utilizando un programa de edición de imágenes (Fireworks CS3; ver [1]). Cabe mencionar que se crearon estas dos bases de imágenes de microcalcificaciones para cada una de las bases de mastografías utilizadas (MIAS e ISSSTEP).

#### 4.3.2. Espacios discriminantes para detectar microcalcificaciones

Para generar un espacio discriminante que permita identificar imágenes de tejido del seno e imágenes de microcalcificaciones utilizando FLD, se cuenta con un conjunto de  $M$  imágenes donde  $m_1$  pertenecen a la clase microcalcificación ( $Clase_1$ ) y  $m_2$  pertenecen a la clase tejido del seno ( $Clase_2$ ).

Para aplicar los FLD de manera original, es necesario que cada imagen se represente como vector columna  $X_j$  de tamaño  $N^2$ , en nuestro caso 144 píxeles (vea la figura 4.6).

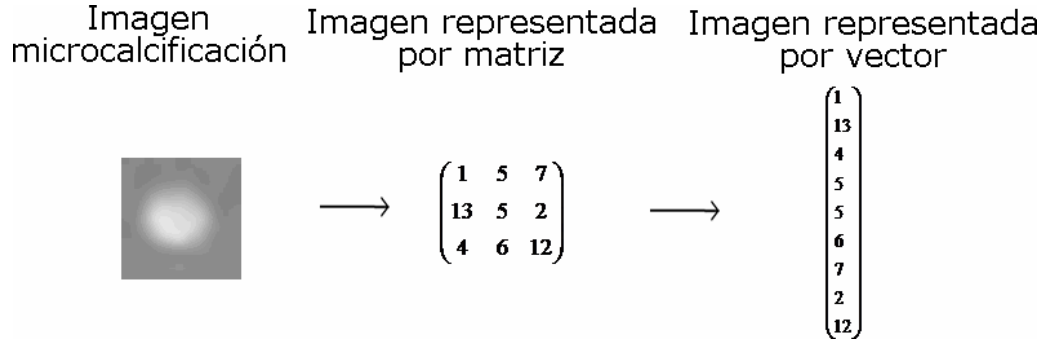


Figura 4.6 Ejemplo de una imagen de tamaño 3 x 3 que puede ser representada mediante una matriz y posteriormente como un vector columna.

Posteriormente se calcula la media general  $\mu$  utilizando todas las imágenes y se calculan las medias  $\mu_1$  y  $\mu_2$  pertenecientes a la clase 1 y a la clase 2 respectivamente. Con estos valores se calcula la matriz de dispersión entre clases  $S_b$  y la matriz de dispersión dentro de las clases  $S_w$ :

$$S_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \quad (4.1)$$

$$S_w = \sum_{i=1}^2 \sum_{X_j \in Clase_i} (X_j - \mu_i)(X_j - \mu_i)^T \quad (4.2)$$

Como ya se mencionó en la sección 2.4.3 de esta tesis, la transformación lineal  $W_{opt}$  que permite mapear los datos (vectores) a un sub-espacio en el que sea fácil discriminar entre las clases existentes (en nuestro caso la clase de microcalcificación y la clase de tejido del seno) está definida por:

$$W_{opt} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} \quad (4.3)$$

Si  $S_w$  es una matriz de rango completo, la matriz de transformación lineal  $W_{opt}$  está formada por los  $m$  eigenvectores con eigenvalores más altos del producto de  $S_w^{-1}$  y  $S_b$ . Pero aplicar directamente FLD a imágenes conlleva un problema debido a que el número de píxeles  $n$  es mucho mayor al número de imágenes  $m$ , la matriz de dispersión entre clases

$S_w$  siempre es una matriz de rango deficiente. Para solucionar este problema, antes de aplicar FLD el espacio es reducido mediante PCA con lo que se asegura que  $S_w$  sea una matriz de rango completo; ver [43].

Debido a lo anterior la transformación lineal óptima  $W$  está definida de la siguiente forma:

$$W_{opt} = \arg \max_W \frac{|W_{FLD}^T W_{PCA}^T S_b W_{PCA} W_{FLD}|}{|W_{FLD}^T W_{PCA}^T S_w W_{PCA} W_{FLD}|} \quad (4.4)$$

Donde la matriz  $W_{opt}$  seleccionada por FLD, está formada por los eigenvectores de  $(W_{PCA}^T S_w W_{PCA})^{-1} (W_{PCA}^T S_b W_{PCA})$ .

Como paso final de la etapa de aprendizaje, todas las imágenes son representadas como vectores columna y son proyectadas en el espacio discriminante encontrado mediante FLD (para más detalles de este paso vea la sección 2.4).

Este procedimiento para crear un espacio discriminante, se realiza para la base de microcalcificaciones provenientes de senos grasos y para la base proveniente de senos densos, generando dos espacios discriminantes. Con estos espacios se detectan las microcalcificaciones presentes en las mastografías.

#### 4.3.3. Segmentación de microcalcificaciones utilizando FLD

Como paso inicial, cuando se desea analizar una mastografía nueva en busca de microcalcificaciones, ésta se clasifica de manera automática en la etapa de identificación de la densidad del seno de nuestro método, ya sea como seno denso o seno graso, esta clasificación de la mastografía es utilizada en la etapa de segmentación de microcalcificaciones.

En la etapa de segmentación se utiliza una ventana de tamaño de 12x12 píxeles para analizar la zona del seno en busca de microcalcificaciones. Este tamaño de ventana se seleccionó ya que todas las microcalcificaciones caben dentro de este tamaño de ventana (vea la figura 4.7).

La región que se encuentra dentro de la ventana se representa como vector columna y se proyecta a uno de los dos espacios discriminantes creados mediante FLD. El espacio al cual se proyecta la imagen se selecciona de acuerdo a la clasificación del seno (graso o denso).

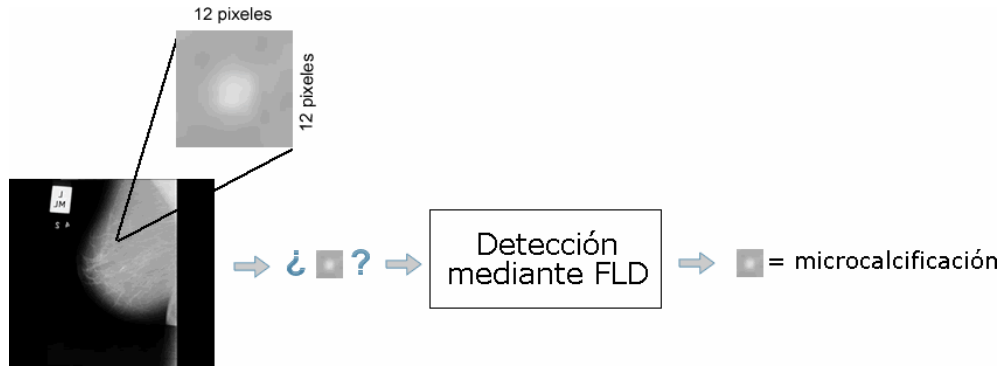


Figura 4.7: Detección de microcalcificaciones mediante FLD.

Una vez proyectado el vector columna que representa la imagen de la región dentro de la ventana al espacio discriminante, mediante el cálculo de la distancia Euclidiana se identifica si la imagen proyectada se asemeja más a la media de la clase de microcalcificaciones o a la media de la clase de tejido del seno. En caso de que se clasifique como parte de la clase de microcalcificaciones, esa zona es marcada en la mastografía con color amarillo, sus coordenadas son almacenadas y la ventana se recorre una posición hacia la derecha; en caso de que esa zona sea clasificada como tejido del seno la ventana simplemente se recorre una posición hacia la derecha. La ventana recorre y analiza el seno de izquierda a derecha de arriba abajo, comenzando en la esquina superior izquierda. Este proceso se repite hasta que la ventana haya analizado todo el seno y se han encontrado todas las posibles microcalcificaciones de la imagen de entrada.

#### 4.4. Reducción de Falsos Positivos

La etapa final de nuestro método se enfoca en reducir la cantidad de falsos positivos. A partir de las regiones que se segmentaron como posibles microcalcificaciones, se extraen atributos basados en las tonalidades de la imagen como son los niveles mínimo, máximo y medio de la intensidad de gris. Además de los atributos anteriores, se extraen características morfológicas, como son: área, perímetro, diámetro equivalente, excentricidad, área rellena, extensión, área convexa, orientación, longitud del eje menor, longitud del eje mayor y



solidez (vea capítulo 2). Es importante notar que debido a que se utiliza una ventana de tamaño fijo en la etapa de segmentación, las regiones identificadas como microcalcificación contienen un poco del tejido del seno que las rodea, por tal motivo, para calcular los valores de las características morfológicas de cada región, es necesario separar la microcalcificación del tejido del seno.

Para segmentar las posibles microcalcificaciones de su fondo se utilizó el método de Otsu. Este método selecciona un umbral global de manera automática basándose en la varianza de los niveles de gris de los píxeles, el umbral seleccionado minimiza la varianza de las tonalidades dentro de cada región y maximiza la varianza de las tonalidades de regiones diferentes; ver [37].

Aplicando este método se genera una imagen binaria a partir de la cual se calculan las características morfológicas de la región correspondiente a la posible microcalcificación.

Para cada una de las regiones detectadas como posibles microcalcificaciones se calculan las características mencionadas anteriormente, éstas se almacenan en un archivo con formato .arff, siendo éste el formato que se utiliza para las bases de datos de Weka.

Por último, se utilizaron las bases de datos de microcalcificaciones como datos de entrenamiento para crear diferentes clasificadores con Weka (evaluando con validación cruzada) y utilizarlos para confirmar (para una nueva imagen) si las regiones segmentadas como microcalcificaciones en realidad lo son o si son parte del tejido del seno que se confundió con una microcalcificación y así se reduce la cantidad de falsos positivos generada en la etapa de segmentación.



## Capítulo 5

### Experimentos y Resultados

En este capítulo se describen los experimentos realizados para probar cada etapa del método propuesto para la detección de microcalcificaciones, así mismo, se describen las bases de imágenes y de datos utilizadas. La métrica para evaluar la etapa de segmentación de microcalcificaciones de nuestro método es el análisis FROC (proporción de marcas de verdaderos positivos vs falsos positivos por imagen). Para la etapa de clasificación de la densidad del seno y la etapa de reducción de falsos positivos se utilizó el análisis ROC (proporción de verdaderos positivos vs proporción de falsos positivos) además de la matriz de confusión.

#### 5.1. Bases de Mastografías

Para evaluar nuestro método se utilizaron dos bases de mastografías: La base de imágenes MIAS y la base de imágenes ISSSTEP. La Base *Mammographic Image Analysis Society* (MIAS), es una base pública de imágenes de mastografías, ésta contiene 322 imágenes con un tamaño de 1024 x 1024 píxeles y su formato original es el Portable Gray Map, (pero por conveniencia para esta tesis, las imágenes se transformaron al formato jpg).

Cabe mencionar que de las 322 mastografías de esta base de imágenes, solo 20 tienen microcalcificaciones. El ángulo con el que se hicieron las tomas de las mastografías es el medio lateral, no están orientadas hacia un mismo lado y vienen agrupadas en pares, donde las mastografías nombradas con números pares son mastografías izquierdas y las impares son mastografías derechas (vea la figura 5.1). La base MIAS cuenta con un documento de información, en el cual se describen las siguientes características para cada mastografía: densidad del seno, la anormalidad que presenta la mastografía, malignidad de la anormalidad, coordenadas del centro y el radio aproximado en píxeles de un círculo que encierra la anormalidad.

- La densidad del seno puede tener uno de los siguientes valores: Graso, Graso-Denso o Denso.
- Los diversos tipos de anormalidades que se pueden presentar son: calcificaciones, masas claramente definidas, masas con márgenes puntiagudos, otro tipo de masas, distorsión en la arquitectura del seno, asimetría y seno normal.
- La malignidad de la anormalidad está clasificada en benigna y maligna.
- El sistema de coordenadas comienza en la parte inferior izquierda de la imagen.

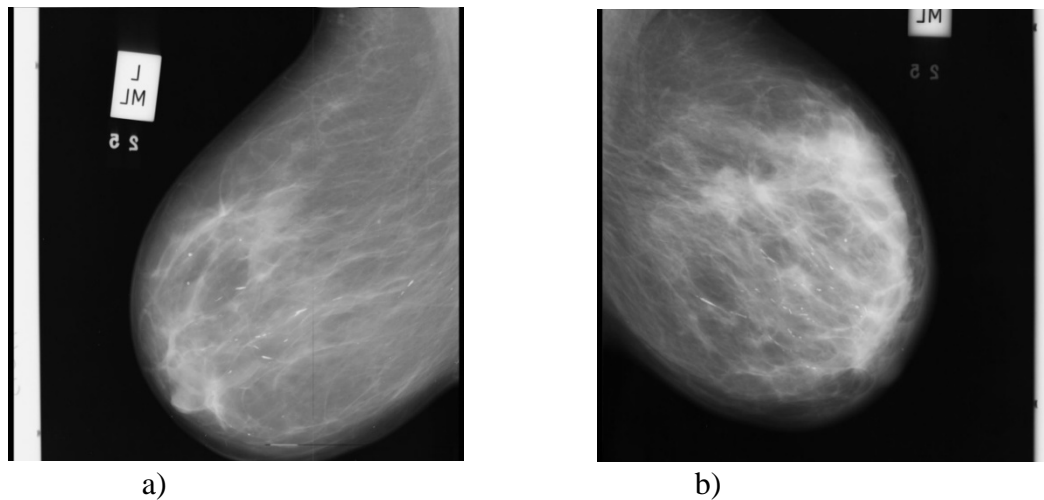


Figura 5.1. Ejemplo de un par de mastografías medio lateral de un paciente de la base MIAS, las dos mastografías presentan microcalcificaciones a) seno izquierdo b) seno derecho.

La otra base de mastografías utilizada para los experimentos es la ISSSTEP, esta base está formada por mastografías del laboratorio de rayos X del Instituto de Seguridad y

Servicios Sociales de los Trabajadores al Servicio de los Poderes del Estado de Puebla (ISSSTEP). Esta base de datos consta de 108 casos, y cada caso se compone de 4 imágenes de mastografías, 2 mastografías (derecha e izquierda) de ángulo medio lateral y 2 mastografías cráneo-caudales (vea la figura 5.2). De estas imágenes solamente 50 mastografías presentan microcalcificaciones.

Las mastografías fueron digitalizadas con un scanner especial para negativos Epson Expression 1680 Profesional Firewire, con una resolución de 1600 dpi. Los tamaños de las imágenes no son uniformes entre ellas, por lo que se adaptaron las imágenes a un tamaño de 1024 x 1024 píxeles con un formato jpg.

Con la ayuda de un Radiólogo, se analizaron las mastografías, se clasificaron las anomalías encontradas y se identificó la zona del seno en que se encontraban. Con esta información se creó una base de mastografías con las lesiones o anomalías marcadas.

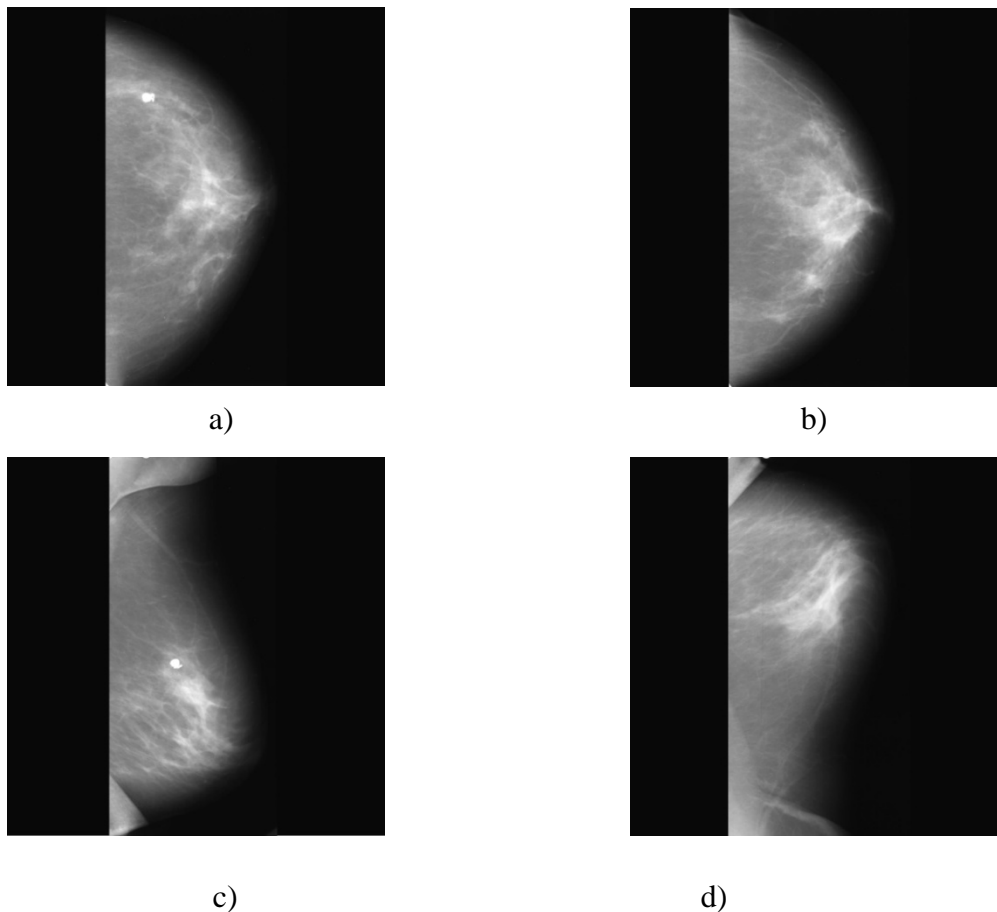


Figura 5.2. Ejemplos de mastografías pertenecientes a un caso de la base de imágenes ISSSTEP, a) y b) Mastografías medio lateral c) y d) Mastografías cráneo-caudal.

## 5.2. Experimentos de la etapa de clasificación de la densidad del seno

Para evaluar el desempeño de la etapa de clasificación de la densidad del seno propuesta en nuestro método, se utilizaron la base MIAS y la base ISSSTEP (descritas en la sección 5.1). Como entrenamiento se utilizaron 150 instancias de la base MIAS, de las cuales 101 corresponden a imágenes de senos densos y 49 a senos grasos; mientras que para la base ISSSTEP se utilizaron 145 instancias de entrenamiento, de las cuales 104 pertenecen a la clase denso y 41 a la clase graso. Además cabe mencionar que todas las instancias del conjunto de entrenamiento son mastografías que no contienen microcalcificaciones. Utilizando cada base de datos (MIAS e ISSSTEP) se realizaron tres experimentos.

El primer experimento se realizó utilizando validación cruzada de 10 pasos (*10 Fold Cross –Validation*). Esta técnica utiliza 90% de los ejemplos de la base de datos como ejemplos de entrenamiento y el 10% restante de los datos para prueba, esto se repite 10 veces tomando diferentes conjuntos de entrenamiento y de prueba (vea la figura 5.3).

Experimento 1: Utilizando validación cruzada

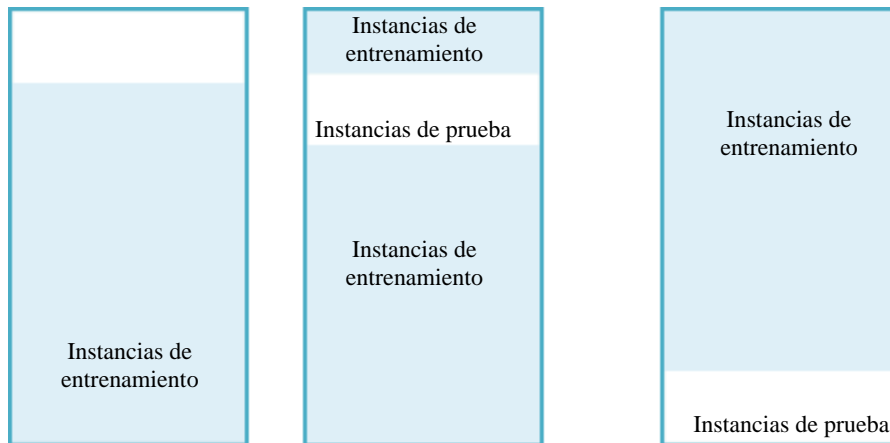


Figura 5.3. Esquema del primer experimento mediante la validación cruzada para evaluar la etapa de clasificación de la densidad del seno.

El segundo experimento se realizó utilizando como conjunto de prueba mastografías que no contienen microcalcificaciones, mientras que para el tercer experimento se utilizó como conjunto de prueba, instancias pertenecientes a las mastografías que presentan microcalcificaciones en cada base de mastografías (MIAS e ISSSTEP). Cabe mencionar que ninguna de las instancias de prueba forman parte de las instancias de entrenamiento (vea la figura 5.4).

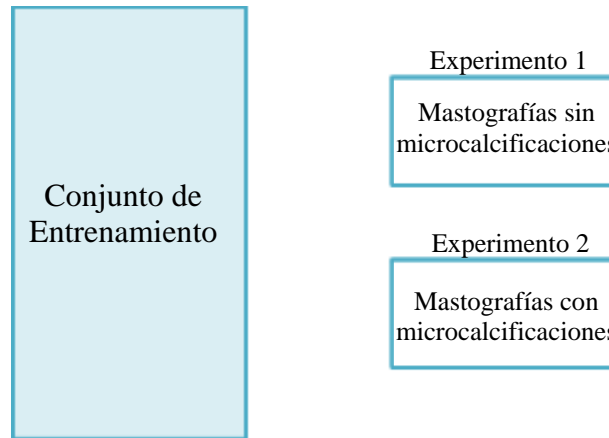


Figura 5.4. Esquema de los experimentos realizados para la etapa de clasificación de la densidad del seno utilizando conjuntos de prueba diferentes.

Los clasificadores que se utilizaron para realizar las pruebas en Weka son: un clasificador basado en reglas (PART), una red neuronal (MultiLayer Perceptron) y un clasificador basado en instancias (IB1) (para más detalles de cómo funcionan estos clasificadores vea la sección 2.2). Como medidas de evaluación de esta etapa se utilizaron los valores de sensibilidad el cual especifica la proporción de verdaderos positivos (TPR) y la especificidad que se refiere a la proporción de verdaderos negativos (TNR) (vea la sección 2.2.1). En nuestro caso, un caso de la clase positivo corresponde a la clase graso mientras que la clase negativo corresponde a la clase denso, por lo que el valor de sensibilidad representa la proporción de instancias correctamente clasificadas como clase graso mientras que el valor de especificidad indica la proporción de instancias clasificadas de manera correcta como parte de la clase denso.

Cabe mencionar que las mastografías catalogadas como graso-denso en la base MIAS, para propósito de esta tesis nosotros las catalogamos como clase denso ya que el tejido que continua predominando es el tejido glandular.

### 5.2.1. Experimentos utilizando la base MIAS

Las matrices de confusión correspondientes a la etapa de clasificación del seno, obtenidas a partir del primer experimento utilizando validación cruzada de 10 pasos con la base MIAS, se pueden observar en la figura 5.5.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 42                | 7     |
|                   | Denso                 | 2                 | 99    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | PART  | Graso             | Denso |
|                   | Graso | 40                | 9     |
|                   | Denso | 6                 | 95    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 42                | 7     |
|                   | Denso | 4                 | 97    |

Figura 5.5. Matrices de confusión obtenidas mediante validación cruzada de 10 pasos utilizando la base de datos para la clasificación de densidad del seno a partir de la base de imágenes MIAS a) red neuronal, b) clasificador basado en reglas y c) clasificador basado en instancias.

De la figura 5.5 se puede ver, que de manera global, los tres clasificadores obtuvieron un desempeño similar, clasificando de manera correcta aproximadamente 141 de las 150 instancias disponibles.



Analizando cada clase por separado, se observa que al clasificar la clase denso se cometieron menos errores que con la clase graso, pero en general se obtuvieron resultados aceptables para ambas clases.

En la tabla 5.1 se pueden observar los resultados de sensibilidad y especificidad obtenidos en la etapa de identificación de la densidad del seno, expresados como la proporción de verdaderos positivos (TPR) y verdaderos negativos (TNR) respectivamente.

Tabla 5.1. Clasificación obtenida con 3 algoritmos de aprendizaje utilizando la base MIAS utilizando la técnica de validación cruzada de 10 pasos.

| Experimento 1: Validación Cruzada Base MIAS |           |                       |                        |
|---|-----------|-----------------------|------------------------|
| Clasificadores                              | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer<br>Perceptron                    | 94%       | 0.857                 | 0.98                   |
| PART  | 90%       | 0.816                 | 0.941                  |
| IB1   | 92.67%    | 0.857                 | 0.96                   |

El porcentaje de exactitud (*accuracy*) que se obtuvo utilizando la base MIAS está entre el 90% y 94%.

De acuerdo a los resultados de la tabla 5.1, se puede observar que los tres clasificadores obtuvieron una proporción de verdaderos positivos (sensibilidad) mayor a 0.81 y una proporción de verdaderos negativos (especificidad) superior a 0.94, esto quiere decir que se cometieron más errores con la clase positivo (clase graso) que con la clase negativo (clase denso).

Para el segundo experimento de la etapa de clasificación del seno se utilizaron como conjunto de prueba 50 mastografías sin microcalcificaciones pertenecientes a la base MIAS (20 mastografías de la clase graso y 30 mastografías de la clase denso), se pueden observar en la figura 5.6.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 17                | 3     |
|                   | Denso                 | 2                 | 28    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | PART  | Graso             | Denso |
|                   | Graso | 20                | 0     |
|                   | Denso | 2                 | 28    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 18                | 2     |
|                   | Denso | 3                 | 27    |

Figura 5.6. Matrices de confusión obtenidas para el primer experimento utilizando como conjunto de prueba mastografías sin microcalcificaciones de la base MIAS. los clasificadores utilizados son: a) red neuronal, b) clasificador basado en reglas y c) clasificador basado en instancias.

De la figura 5.6 se puede ver, que de manera global, los tres clasificadores obtuvieron un desempeño similar, clasificando de manera correcta aproximadamente 45 de las 50 instancias disponibles.

Analizando cada clase por separado, se observa que al clasificar la clase denso se cometieron aproximadamente la misma cantidad de errores que con la clase graso, pero en general se obtuvieron resultados aceptables para ambas clases.

En la tabla 5.2 se pueden observar los resultados de sensibilidad y especificidad obtenidos en la etapa de identificación de la densidad del seno, expresados como la proporción de verdaderos positivos (TPR) y verdaderos negativos (TNR) respectivamente.

Tabla 5.2. Clasificación de la densidad del seno obtenida con 3 algoritmos de aprendizaje utilizando mastografías sin microcalcificaciones de la base MIAS.

| Experimento 1: Mastografías sin microcalcificaciones |           |                       |                        |
|--|-----------|-----------------------|------------------------|
| Base MIAS  |           |                       |                        |
| Clasificadores                                       | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer<br>Perceptron                             | 90%       | 0.85                  | 0.933                  |
| PART   | 96%       | 1                     | 0.933                  |
| IB1  | 90%       | 0.9                   | 0.9                    |

El porcentaje de exactitud (*accuracy*) que se obtuvo utilizando la base MIAS está entre el 90% y 92%.

De acuerdo a los resultados de la tabla 5.2, se puede observar que los tres clasificadores obtuvieron una proporción de verdaderos positivos (sensibilidad) mayor a 0.85 y una proporción de verdaderos negativos (especificidad) superior a 0.9, esto quiere decir que se cometieron más errores con la clase positivo (clase graso) que con la clase negativo (clase denso).

Para realizar el tercer experimento con la base MIAS, se utilizó un total de 20 instancias de prueba, de las cuales 5 pertenecen a la clase graso y 15 a la clase denso. Estas imágenes del conjunto de prueba no forman parte del conjunto de las 150 instancias de entrenamiento de la base MIAS. En la figura 5.7 se pueden ver las matrices de confusión de los resultados del segundo experimento utilizando la base MIAS.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 2                 | 3     |
|                   | Denso                 | 1                 | 14    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | NNge  | Graso             | Denso |
|                   | Graso | 5                 | 0     |
|                   | Denso | 0                 | 15    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 4                 | 1     |
|                   | Denso | 1                 | 14    |

Figura 5.7. Matrices de confusión utilizando como conjunto de prueba solo las imágenes que contienen microcalcificaciones, utilizando como conjunto de entrenamiento la base MIAS a) matriz obtenida utilizando red neuronal, b) matriz obtenida utilizando clasificador basado en reglas y c) matriz obtenida utilizando clasificador basado en instancias.

La exactitud de clasificación obtenida en este experimento para la clase graso fue muy variada, con la red neuronal se clasificó correctamente menos de la mitad de las instancias, mientras que con el clasificador basado en reglas se clasificó de manera correcta el 100% de las instancias. En cuanto a la clase denso, la cantidad de instancias clasificadas de manera correcta fue similar para los tres clasificadores, uno de ellos alcanzando el 100% de exactitud.

En la tabla 5.3 se pueden visualizar los resultados obtenidos para este experimento en términos de sensibilidad y especificidad.

Tabla 5.3. Clasificación obtenida con 3 algoritmos de aprendizaje implementados en Weka utilizando la base de datos MIAS. El conjunto de prueba consta únicamente de las imágenes que contienen microcalcificaciones.

| Experimento 2: Mastografías con microcalcificaciones |           |                       |                        |
|--|-----------|-----------------------|------------------------|
| Base MIAS  |           |                       |                        |
| Clasificadores                                       | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer Perceptron                                | 80%       | 0.4                   | 0.933                  |
| PART   | 100%      | 1                     | 1                      |
| IB1  | 90%       | 0.8                   | 0.933                  |

El porcentaje de exactitud (*accuracy*) que se alcanzó en este experimento está entre el 80% y 100%. Debido a que las instancias de prueba correspondientes a la clase graso (clase positiva) solamente son cinco, el cometer uno o dos errores afecta de manera dramática en el valor de sensibilidad (proporción de verdaderos positivos), como es el caso del resultado obtenido por la red neuronal cuyo valor de sensibilidad no rebasa el 0.5, de manera opuesta, el clasificador basado en reglas obtuvo una clasificación perfecta obteniendo un valor de sensibilidad de 1. De acuerdo a esto se puede ver que en este experimento los valores de sensibilidad no tienden a seguir un patrón, principalmente debido al hecho de que son muy pocos ejemplos, pero el objetivo de este experimento en particular a diferencia del primero que utiliza validación cruzada, no es evaluar el desempeño general de la clasificación sino identificar como se clasifican las mastografías que se utilizan en la etapa de segmentación de microcalcificaciones de nuestro método.

Respecto al valor obtenido de especificidad que representa la proporción de instancias clasificadas correctamente como clase denso (clase negativa) están por encima de 0.9.

### 5.2.2. Experimentos utilizando la base de datos ISSSTEP

En el primer experimento utilizando la base ISSSTEP, se evaluó el desempeño de los clasificadores utilizando validación cruzada de 10 pasos. El conjunto evaluado consta de 145 instancias, de las cuales 104 pertenecen a la clase denso y 41 a la clase graso.

En la figura 5.8 se pueden analizar las matrices de confusión de los resultados que se obtuvieron para este experimento.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 38                | 3     |
|                   | Denso                 | 4                 | 100   |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | PART  | Graso             | Denso |
|                   | Graso | 38                | 3     |
|                   | Denso | 3                 | 101   |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 37                | 4     |
|                   | Denso | 3                 | 101   |

Figura 5.8. Matrices de confusión de la base de datos Hist\_ISSSTEP utilizando validación cruzada de 10 pasos a) matriz utilizando red neuronal, b) matriz utilizando clasificador basado en reglas y c) matriz utilizando clasificador basado en instancias.

En términos generales, los resultados obtenidos utilizando la base ISSSTEP, fueron mejores que los obtenidos con la base de datos MIAS. A diferencia de los resultados de la base MIAS en que se clasificaron de mejor manera las instancias pertenecientes a la clase denso, utilizando la base ISSSTEP, todos los clasificadores evaluados mostraron un buen desempeño para ambas clases. De manera global, para la clase graso se clasificaron de manera correcta entre 37 y 38 instancias de las 41 disponibles mientras que para la clase denso se clasificaron correctamente entre 100 y 101 instancias de las 104 correspondientes a esta clase.

En la tabla 5.4 se pueden visualizar las medidas de sensibilidad y especificidad generadas por los tres clasificadores que se probaron.

Tabla 5.4. Clasificación obtenida a partir de 3 algoritmos de aprendizaje utilizando la base de datos ISSSTEP, mediante la técnica de validación cruzada de 10 pasos.

| Experimento 1: Validación Cruzada Base ISSSTEP |           |                       |                        |
|--|-----------|-----------------------|------------------------|
| Clasificadores                                 | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer Perceptron                          | 95.17%    | 0.927                 | 0.962                  |
| PART   | 95.86%    | 0.927                 | 0.971                  |
| IB1  | 95.17%    | 0.902                 | 0.971                  |

En la tabla 5.4, se puede ver que el porcentaje de precisión de la clasificación obtenida para la base de datos ISSSTEP, se encuentra por arriba del 95%. Específicamente el valor de sensibilidad es mayor a 0.9 en los tres clasificadores. En cuanto al valor de especificidad, es decir la proporción de instancias clasificadas correctamente como pertenecientes a la clase denso, se obtuvo un valor por encima de 0.96.

La proporción de clasificaciones correctas para la clase densa utilizando la base ISSSTEP es muy similar a la obtenida con la base MIAS, pero la clasificación correcta para la clase grasa tiende a ser mejor con la base ISSSTEP.

En el segundo experimento utilizando la base ISSSTEP, se evaluó el desempeño de los clasificadores utilizando como conjunto de prueba mastografías sin microcalcificaciones. El conjunto de prueba para el segundo experimento cuenta con 36 instancias de las cuales 12 pertenecen a la clase grasa y 24 pertenecen a la clase denso.

En la figura 5.9 se pueden analizar las matrices de confusión de los resultados que se obtuvieron para este experimento.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 10                | 2     |
|                   | Denso                 | 1                 | 23    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | PART  | Graso             | Denso |
|                   | Graso | 10                | 2     |
|                   | Denso | 2                 | 22    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 11                | 1     |
|                   | Denso | 1                 | 23    |

Figura 5.9. Matrices de confusión de la base de datos ISSSTEP utilizando como conjunto de prueba mastografías sin microcalcificaciones a) matriz utilizando red neuronal, b) matriz utilizando clasificador basado en reglas y c) matriz utilizando clasificador basado en instancias.

En términos generales, los resultados obtenidos utilizando la base ISSSTEP, fueron mejores que los obtenidos con la base de datos MIAS. A diferencia de los resultados de la base MIAS en que se clasificaron de mejor manera las instancias pertenecientes a la clase denso, utilizando la base ISSSTEP, todos los clasificadores evaluados mostraron un buen desempeño para ambas clases. De manera global, para la clase graso se clasificaron de manera correcta entre 10 y 11 instancias de las 12 disponibles mientras que para la clase denso se clasificaron correctamente 23 instancias de las 24 correspondientes a esta clase. En la tabla 5.5 se pueden visualizar las medidas de sensibilidad y especificidad generadas por los tres clasificadores que se probaron.



Tabla 5.5. Clasificación obtenida a partir de 3 algoritmos de aprendizaje utilizando la base ISSSTEP, considerando solamente mastografías con microcalcificaciones.

| Experimento 1: Mastografías sin microcalcificaciones |           |                       |                        |
|--|-----------|-----------------------|------------------------|
| Base ISSSTEP   |           |                       |                        |
| Clasificadores                                       | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer Perceptron                                | 91.67%    | 0.833                 | 0.958                  |
| PART   | 88.88%    | 0.833                 | 0.917                  |
| IB1  | 94.44%    | 0.917                 | 0.958                  |

En la tabla 5.5, se puede ver que el porcentaje de precisión de la clasificación obtenida para la base ISSSTEP, se encuentra por arriba del 91%. Específicamente el valor de sensibilidad es mayor a 0.83 en los tres clasificadores. En cuanto al valor de especificidad, es decir la proporción de instancias clasificadas correctamente como pertenecientes a la clase denso, se obtuvo un valor por encima de 0.91.

La proporción de clasificaciones correctas para la clase denso utilizando la base ISSSTEP es muy similar a la obtenida con la base MIAS, pero la clasificación correcta para la clase graso tiende a ser mejor con la base ISSSTEP. Esto último debido principalmente a que algunas mastografías de la base MIAS, muestran parte del musculo pectoral por lo que el clasificador confunde los senos grasos como si fueran senos densos.

Para el tercer experimento, el conjunto de prueba está formado por las 50 instancias correspondientes a las imágenes que presentan microcalcificaciones en la base ISSSTEP, las cuales no fueron contempladas en el conjunto de entrenamiento. De estas 50 instancias 24 corresponden a la clase graso y 26 a la clase denso. En la figura 5.10 se muestran las matrices de confusión que se obtuvieron al realizar este experimento.

|                   |                       | Clasificados como |       |
|-------------------|-----------------------|-------------------|-------|
| a) Valores Reales | Multilayer Perceptron | Graso             | Denso |
|                   | Graso                 | 22                | 2     |
|                   | Denso                 | 4                 | 22    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| b) Valores Reales | PART  | Graso             | Denso |
|                   | Graso | 21                | 3     |
|                   | Denso | 7                 | 19    |

|                   |       | Clasificados como |       |
|-------------------|-------|-------------------|-------|
| c) Valores Reales | IB1   | Graso             | Denso |
|                   | Graso | 22                | 2     |
|                   | Denso | 3                 | 23    |

Figura 5.10. Matrices de confusión utilizando como conjunto de prueba solo mastografías que contienen microcalcificaciones, utilizando como conjunto de entrenamiento la base de datos ISSSTEP a) utilizando red neuronal, b) utilizando clasificador basado en reglas y c) utilizando clasificador basado en instancias.

Las clasificaciones obtenidas por los tres algoritmos de aprendizaje fueron muy cercanas, de las 24 instancias de prueba de la clase graso, entre 22 y 23 fueron clasificadas correctamente. Para la clase denso, aunque los resultados fueron más variados, la red neuronal y el IB1 clasificaron de manera correcta entre 22 y 23 de las 26 instancias de prueba. De manera inversa a los resultados del mismo experimento pero realizado con la base MIAS, aquí se clasifica mejor la clase graso que la clase denso. En la tabla 5.6 se muestran los valores de sensibilidad y especificidad correspondientes a este experimento.

Tabla 5.6. Clasificación obtenida a partir de 3 algoritmos de aprendizaje de Weka, utilizando la base de datos ISSSTEP, utilizando como conjunto de prueba solamente mastografías que contienen microcalcificaciones.

| Experimento 2: Mastografías con microcalcificaciones |           |                       |                        |
|--|-----------|-----------------------|------------------------|
| Base ISSSTEP   |           |                       |                        |
| Clasificadores                                       | Exactitud | Sensibilidad<br>(TPR) | Especificidad<br>(TNR) |
| MultiLayer Perceptron                                | 88%       | 0.91                  | 0.84                   |
| PART   | 80%       | 0.87                  | 0.73                   |
| IB1  | 90%       | 0.91                  | 0.88                   |

Los porcentajes de la exactitud general de clasificación están entre el 80% y el 90%, los resultados obtenidos por los tres clasificadores son aceptables, algunos de ellos con mejor desempeño, llegando al 90% de exactitud de clasificación.

### 5.3. Experimentos de la etapa de segmentación de microcalcificaciones

Para evaluar la etapa de segmentación de nuestro método se realizaron tres pruebas para cada base de mastografías (MIAS e ISSSTEP). En el primer experimento se probó el método propuesto en esta tesis de manera original, es decir, se tomó en cuenta la clasificación de la densidad de los senos obtenida a partir de los resultados de la etapa de la identificación de la densidad del seno de nuestro método; además se utilizaron los dos espacios discriminantes, uno para los senos grasos y otro para los senos densos.

El segundo experimento tiene el objetivo de mostrar que tanto afecta a nuestro método, el tomar en cuenta la densidad del seno. En este experimento no se tomó en cuenta la clasificación de la densidad del seno obtenido en la primera etapa de nuestro método, solo

se utilizó un espacio proveniente de imágenes de microcalcificaciones tanto de senos grasos como de senos densos.

El objetivo del tercer experimento es realizar una comparación con algún otro método para la detección de microcalcificaciones. Como ya se mencionó en el capítulo 3, dedicado al estado del arte, una de las técnicas más utilizadas para la detección de microcalcificaciones es el uso de wavelets y umbrales. Estos métodos parten de la idea de que las microcalcificaciones están en un rango de frecuencias altas y mediante el uso de wavelets y umbrales son capaces de detectarlas. Debido a estas características y a la dificultad para adquirir y reproducir algún otro método, se seleccionó el método desarrollado por Flores y González cuya segmentación está basada en wavelets y umbrales (locales y globales); ver [16].

Para la evaluación de los tres experimentos, se utilizó el análisis FROC, el cual permite evaluar si ciertas regiones fueron detectadas de manera correcta en una imagen. En nuestro caso permite evaluar si la detección de microcalcificaciones en la mastografía fue correcta. Las medidas de evaluación utilizadas en el análisis FROC, son la proporción de marcas de verdaderos positivos (TPR) y la proporción de marcas de falsos positivos por imagen (FPI). Un valor de TPR igual a 1 significa que todas las microcalcificaciones en todas las mastografías fueron detectadas de manera correcta y un valor de FPI de 0 significa que ninguna región que no corresponde a una microcalcificación fue clasificada como tal. Por lo tanto, entre más cercano esté el valor de TPR a 1 significa que la cantidad de microcalcificaciones detectadas es mayor, (para más detalles del análisis FROC vea el capítulo 2).

En cada mastografía se calcula cuantas marcas de microcalcificaciones se detectaron y al final se calcula el promedio de todas las mastografías, esto nos indica la proporción de marcas de verdaderos positivos. Para el cálculo de los FPI se cuenta el número total de falsos positivos en todas las mastografías y se calcula el promedio.

Por otra parte, tomando en cuenta que al evaluar las mastografías que no presentan microcalcificaciones solamente se obtienen falsos positivos en la etapa de segmentación, los resultados obtenidos para este conjunto de mastografías no se presentan en esta sección, sino que se presentan en los resultados obtenidos en la etapa de reducción de falsos positivos.

### 5.3.1. Experimentos utilizando la base MIAS

Para realizar los experimentos correspondientes a esta etapa se utilizaron las 20 mastografías que presentan microcalcificaciones de la base MIAS, de las cuales 5 son senos grasos y 15 senos densos.

En la tabla 5.7 se pueden observar los resultados obtenidos con nuestro método, la parte superior corresponde a los resultados alcanzados tomando en cuenta la densidad del seno y en la parte inferior de la tabla se muestran los valores obtenidos sin tomar en cuenta la densidad del seno.

Tabla 5.7. Resultados de los experimentos de segmentación de microcalcificaciones utilizando la base de imágenes MIAS.

| Detección de microcalcificaciones mediante FLD<br>Tomando en cuenta la densidad del seno<br>Base MIAS   |              |            |
|---|--------------|------------|
| Densidad  | TPR%         | FPI        |
| Seno Graso  | 91.86%       | 1          |
| Seno Denso  | 90.94%       | 4          |
| Promedio  | <b>91.4%</b> | <b>2.5</b> |
| Detección de microcalcificaciones mediante FLD<br>Sin tomar en cuenta la densidad del seno<br>Base MIAS |              |            |
| Seno Graso  | 88.99%       | 18         |
| Seno Denso  | 70.07%       | 21.64      |
| Promedio  | 79.53%       | 19.82      |

En el primer experimento, en el que se toma en cuenta la clasificación de la densidad del seno, la proporción promedio de las microcalcificaciones detectadas utilizando la base MIAS fue de 91.4%. En senos grasos se detectaron de manera correcta el 91.86% mientras que en senos densos (senos considerados como difíciles) se detectó correctamente el 90.94%. El valor promedio de la proporción de marcas de regiones que fueron identificadas como microcalcificación siendo que eran tejido del seno fue de 2.5 por imagen.

Los resultados obtenidos para el segundo experimento, para los que no se toma en cuenta la clasificación del seno, fueron menos competitivos que los obtenidos tomando en cuenta la clasificación de la densidad del seno. La proporción de microcalcificaciones detectadas fue de 79.53%. En los senos grasos se detectaron de manera correcta 88.90 % de las microcalcificaciones, lo cual es muy cercano a los resultados obtenidos tomando en cuenta la densidad del seno. Sin embargo, para la proporción de microcalcificaciones detectadas en los senos densos, disminuyó en más del 20%, logrando una tasa de 70.07% de marcas de verdaderos positivos. Además, al no tomar en cuenta el tipo de seno (graso o denso), se aumentan significativamente el número de falsos positivos por imagen.

En la figura 5.11 se muestra un ejemplo de la detección de microcalcificaciones tomando en cuenta la densidad del seno y la detección cuando esta densidad no se considera.

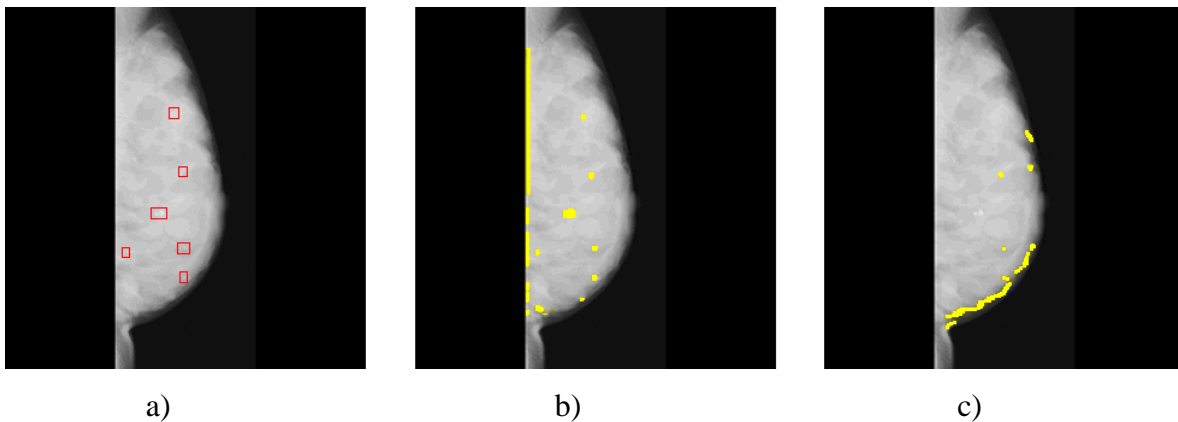


Figura 5.11. a) Mastografía original, los rectángulos de color rojo se agregaron solo para mostrar la ubicación de las microcalcificaciones b) detección de microcalcificaciones tomando en cuenta la clasificación de densidad del seno y c) muestra detección de microcalcificaciones sin tomar en cuenta la clasificación del seno.

Con el fin de tener un punto de comparación para evaluar nuestros resultados obtenidos con el método de detección basado en FLD, se realizó una comparación con el método desarrollado por Flores y González, el cual está basado en wavelets y umbralización; ver [16]. Para esta prueba se utilizaron las mismas 20 mastografías pertenecientes a la base MIAS que presentan microcalcificaciones. En la tabla 5.8 se presentan los resultados obtenidos por los dos algoritmos.

Tabla 5.8. Resultados obtenidos con el método basado en FLD propuesto en esta tesis y el método propuesto por Flores y González utilizando 20 imágenes de mastografías que contienen microcalcificaciones provenientes de la base MIAS.

| Comparación de resultados de la base MIAS |                         |     |  |     |
|---|-------------------------|-----|--|-----|
|   | Algoritmo basado en FLD |     | Algoritmo basado en wavelets y umbralización |     |
| Densidad                                  | TPR%                    | FPI | TPR%   | FPI |
| Senos Densos                              | 90.94%                  | 4   | 74.15%                                       | 20  |
| Senos Grasos                              | 91.86%                  | 1   | 81.25%                                       | 16  |
| Promedio                                  | <b>91.4%</b>            | 2.5 | <b>77.7%</b>                                 | 18  |

De acuerdo a los resultados obtenidos, se puede ver que nuestro método basado en FLD, obtuvo un mejor desempeño en la etapa de segmentación que el obtenido por el método basado en wavelets y umbralización. Analizando solamente los resultados de los senos densos, el porcentaje de detección de microcalcificaciones que se consiguió con el método basado en umbralización fue de 74%, menos del 15% del obtenido con el método basado en FLD. Para los senos grasos los resultados generados con el método basado en FLD superan en un 10% a los obtenidos con el método de umbralización.

De manera global, los resultados alcanzados con el método basado en umbralización, fueron similares a los obtenidos cuando la densidad del seno no es considerada en nuestro

método, para los cuales el porcentaje de las microcalcificaciones detectadas disminuye considerablemente y el número de falsos positivos se incrementa de manera importante comparado con el método basado en FLD que toma en cuenta la densidad del seno.

En la figura 5.12 se pueden observar dos ejemplos de los resultados obtenidos con mastografías de la base MIAS utilizando el método basado en wavelets y el método basado en FLD.

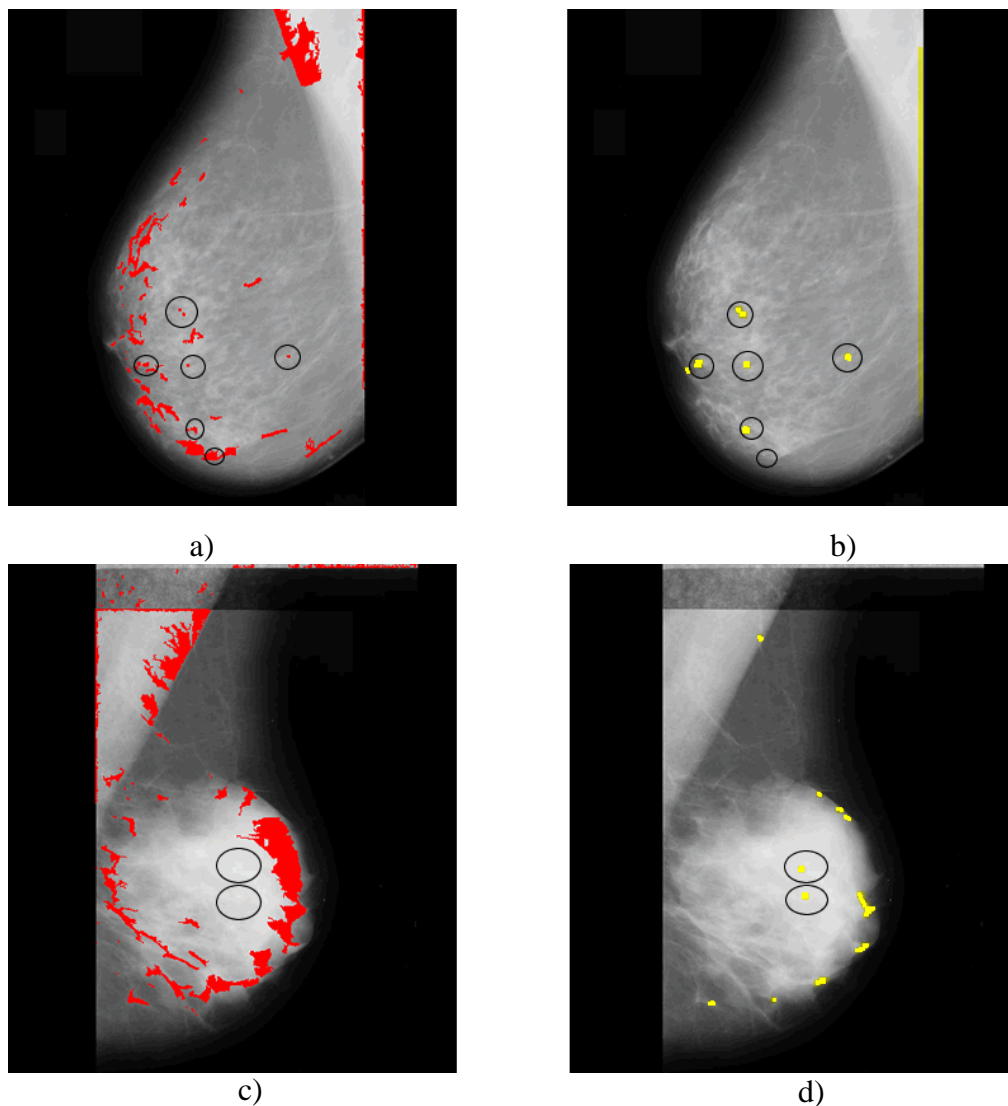


Figura 5.12. Ejemplo de dos mastografías analizadas en busca de microcalcificaciones, las ubicación correcta de las microcalcificaciones se indica con los círculos negros, las posibles microcalcificaciones detectadas por el método basado en umbrales (a, c) se segmentan en color rojo, mientras que las posibles microcalcificaciones detectadas por el método basado en FLD (b, d) son segmentadas en color amarillo.



En el inciso a) de la Figura 5.12 se muestra el resultado obtenido para la primer mastografía, utilizando el método basado en umbrales, donde todas la microcalcificaciones en la mastografía fueron detectadas, pero el número de falsos positivos es elevado. Para este mismo ejemplo de mastografía, el resultado obtenido utilizando el método propuesto en este trabajo, el cual está basado en FLD, se muestra en el inciso b) en el cual únicamente una microcalcificación de las seis existentes no se detectó, con lo que no se afecta el resultado de manera dramática, además utilizando al método basado en FLD el desempeño de detección es mejor ya que hablando específicamente de este ejemplo el número de falsos positivos es casi nulo.

Para el segundo ejemplo de mastografía de la figura 5.12 incisos c) y d), aplicando el método basado en umbrales (c), ninguna de las dos microcalcificaciones se detectó y el número de falsos positivos es un problema, mientras que aplicando el método basado en FLD (d) las dos microcalcificaciones se detectaron exitosamente y el número de falsos positivos sigue siendo bajo.

### **5.3.2. Experimentos utilizando la base ISSSTEP**

Para probar nuestro método con la base de imágenes ISSSTEP, se utilizaron las 50 mastografías que contienen microcalcificaciones, de las cuales 24 son senos grasos y 26 son senos densos. Los resultados que se obtuvieron mediante el método basado en FLD se muestran en la tabla 5.9, en la parte superior de la tabla se hace referencia a los resultados obtenidos tomando en cuenta la densidad del seno y la parte inferior corresponde a los resultados generados sin tomar en cuenta la densidad.

Tomando en cuenta la clasificación del seno, la proporción promedio de las microcalcificaciones detectadas fue de 93.17%, en senos difíciles la tasa de microcalcificaciones detectadas de manera correcta fue de 92.03% y para los senos grasos fue de 94.31%. La cantidad de marcas de falsos positivos por imagen en promedio fue de 6.78.

La proporción de microcalcificaciones detectadas correctamente utilizando la base ISSSTEP aumento en un 2 % aproximadamente comparado con la base MIAS, lo cual es bueno, pero la cantidad de falsos positivos por imagen también se incrementó un poco.

Tabla 5.9. Resultados de experimentos de segmentación de microcalcificaciones utilizando la base ISSSTEP.

| Detección de microcalcificaciones mediante FLD<br>Tomando en cuenta densidad del seno<br>Base ISSSTEP      |               |             |
|--|---------------|-------------|
| Densidad   | TPR%          | FPI         |
| Seno Graso   | 94.31%        | 6.09        |
| Seno Denso   | 92.03%        | 7.48        |
| Promedio   | <b>93.17%</b> | <b>6.78</b> |
| Detección de microcalcificaciones mediante FLD<br>Sin tomar en cuenta la densidad del seno<br>Base ISSSTEP |               |             |
| Seno Graso   | 65.14%        | 2.31        |
| Seno Denso   | 83.70%        | 5.37        |
| Promedio   | 74.42%        | 3.84        |

Cuando la densidad del seno no se toma en cuenta, la proporción de las microcalcificaciones detectadas fue de 74.42% que es casi 20% menor que cuando se utiliza la clasificación de densidad del seno.

En la figura 5.13 se visualiza el ejemplo de un resultado de una mastografía tomando en cuenta la densidad del seno y no tomándola en cuenta.

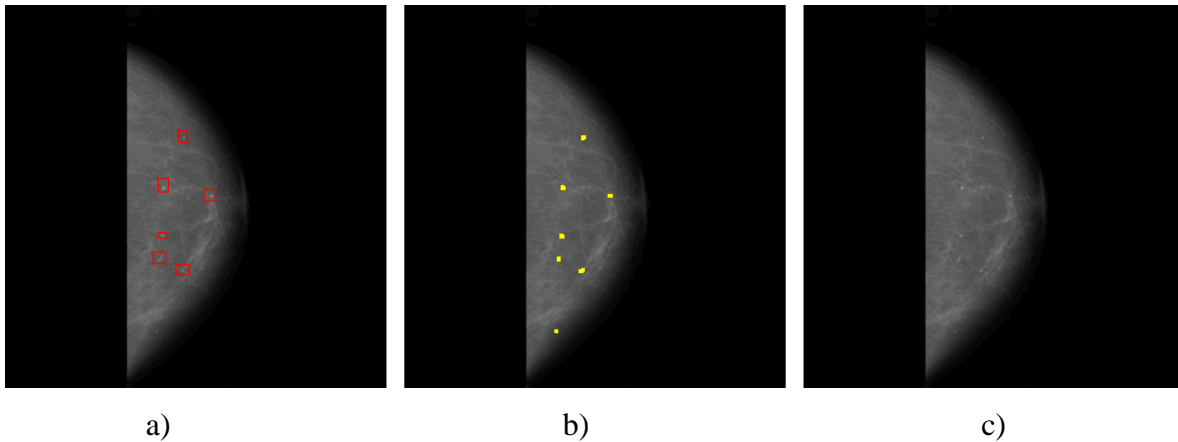


Figura 5.13. a) Mastografía original, los rectángulos de color rojo se agregaron solo para mostrar la ubicación de las microcalcificaciones b) tomando en cuenta la densidad del seno se detectaron todas las microcalcificaciones y c) no tomando en cuenta la densidad del seno, no se logro detectar ninguna microcalcificación.

Para realizar la comparación de nuestro método con el método basado en umbrales y wavelets, se utilizaron las mismas 50 mastografías de la base ISSSTEP que contienen microcalcificaciones. Los resultados obtenidos se pueden observar en la tabla 5.10.

Tabla 5.10. Resultados obtenidos con el método basado en FLD propuesto en esta tesis y el método propuesto por Flores y González, utilizando 50 imágenes de mastografías de la base ISSSTEP, las cuales contienen microcalcificaciones.

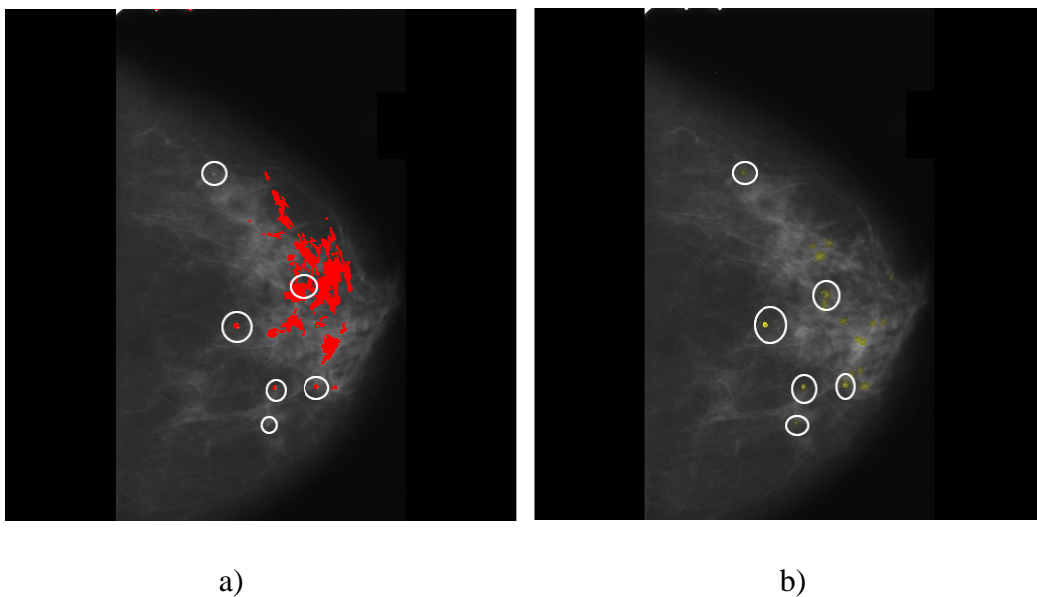
| Comparación de resultados base ISSSTEP |                         |      |  |       |
|--|-------------------------|------|--|-------|
|  | Algoritmo basado en FLD |      | Algoritmo basado en umbralización y wavelets |       |
| Densidad                               | TPR %                   | FPI  | TPR%   | FPI   |
| Senos Densos                           | 92.03%                  | 7.48 | 75.76%                                       | 21.03 |
| Senos Grasos                           | 94.31%                  | 6.09 | 85.73%                                       | 19.86 |
| Promedio                               | <b>93.4%</b>            | 6.78 | <b>80.75%</b>                                | 20.45 |

Al igual que con los resultados obtenidos a partir de la base MIAS, utilizando la base ISSSTEP, se puede ver que el método basado en FLD superó en más del 10 %, de manera global, al método basado en umbrales y wavelets.

Analizando únicamente los resultados obtenidos con los senos grasos, el método basado en wavelets obtuvo menor porcentaje exactitud de detección que nuestro método, pero este porcentaje de detección sigue siendo aceptable según el experto (radiólogo). En contraste con esto, en senos densos (difíciles), los resultados del método basado en wavelets fueron pobres, alcanzando solamente un 75.76% y aún peor; generando demasiados falsos positivos mientras que el método basado en FLD alcanzó un 92.03% y solamente un valor de 7.48 de falsos positivos por imagen.

En la figura 5.14 se presentan dos ejemplos de los resultados obtenidos con ambos métodos, en el inciso a) utilizando el método basado en umbrales y wavelets se logran identificar solamente 4 de las 6 microcalcificaciones existentes en esa mastografía y el número de falsos positivos es elevado, mientras que el resultado obtenido por el método basado en FLD se muestra en el inciso b) para el que las 6 microcalcificaciones son detectadas y el número de falsos positivos es casi nulo.

En el segundo ejemplo, el método basado en umbrales y wavelets no logra detectar la microcalcificación existente (c), mientras que el método basado en FLD la detecta de manera correcta con un número de falsos positivos pequeño (d).



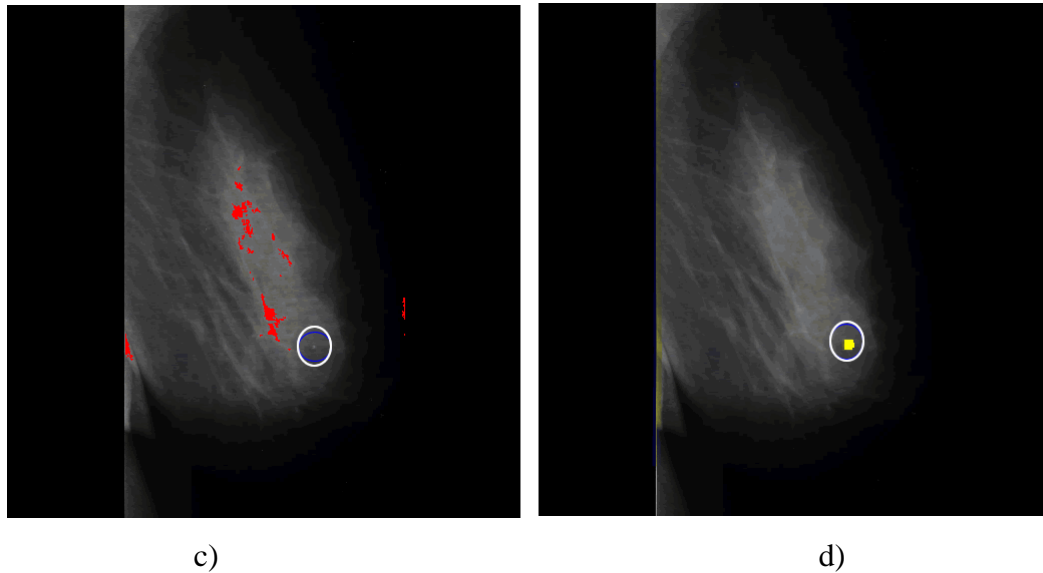


Figura 5.14. Ejemplo de detección de microcalcificaciones para dos mastografías pertenecientes a la base de imágenes ISSTEP, las posibles microcalcificaciones detectadas por el método basado en umbralización (a, c) se segmentan en color rojo, mientras que las posibles microcalcificaciones detectadas por el método basado en FLD (b, d) son segmentadas en color amarillo.

#### 5.4. Experimentos de la etapa de reducción de falsos positivos

Con el fin de disminuir el número de falsos positivos generados en la etapa de detección de microcalcificaciones descrita en la sección anterior, para cada una de las regiones que fueron identificadas como posibles microcalcificaciones se calcularon las siguientes características: área, perímetro, diámetro equivalente, excentricidad, área rellena, extensión, área convexa, orientación, longitud del eje menor, longitud del eje mayor, solidez, intensidad mínima, media y máxima del nivel de gris.

Para realizar el experimento de esta etapa, estas características son utilizadas como entradas para diferentes algoritmos de aprendizaje de Weka, se seleccionó un clasificador basado en instancias, otro en reglas, un algoritmo basado en árboles y una red neuronal. Además para contar con un punto de evaluación de los resultados obtenidos en la fase de reducción de falsos positivos, se hizo una comparación con los resultados correspondientes a esta misma etapa pero obtenidos con el método basado en wavelets y umbrales; ver [16].

Las características que ellos utilizan son: área, diámetro, densidad, convexidad, radio interno de la región, radio externo de la región, orientación, radio, circularidad, energía, factor de estructura, distancia, redondez, longitud del contorno, excentricidad.

Las regiones a partir de las cuales se extraen las características por ambos métodos provienen del mismo conjunto de mastografías utilizadas en los experimentos realizados para la fase de segmentación descrita en la sección anterior.

Para las mastografías que presentan microcalcificaciones la técnica que se utilizó para probar el desempeño de la fase de reducción de falsos positivos de ambos métodos es mediante la validación cruzada a 10 pasos (*10 Fold Cross-Validation*). Mientras que para probar las mastografías sin microcalcificaciones se utilizó como conjunto de entrenamiento la base generada utilizando las mastografías con microcalcificaciones.

Los resultados obtenidos por los dos métodos utilizando mastografías con microcalcificaciones son expresados en términos de la proporción de verdaderos positivos TPR y la proporción de verdaderos negativos TNR, así como el porcentaje general de exactitud de clasificación. Para las mastografías sin microcalcificaciones solamente se presentan los resultados de TNR ya que no se cuenta con instancias positivas. En nuestro caso la TPR nos indica la proporción de instancias correspondientes a la clase microcalcificación que fueron clasificadas correctamente, mientras que la TNR nos indica la cantidad de instancias que fueron clasificadas como tejido siendo que realmente son tejido del seno.

#### **5.4.1. Experimento utilizando la base MIAS**

Los resultados obtenidos con nuestro método en la etapa de reducción de falsos positivos utilizando mastografías con microcalcificaciones y mastografías sin microcalcificaciones se presentan en la tabla 5.11.

Tabla 5.11. Resultados obtenidos por nuestro método en la etapa de reducción de falsos positivos utilizando mastografías con y sin microcalcificaciones.

| Fase de Reducción de Falsos Positivos<br>base MIAS |  |      |      |  |
|--|--|------|------|--|
|  | Mastografías con<br>microcalcificaciones |      |      | Mastografías sin<br>microcalcificaciones |
| Clasificador                                       | Exactitud %                              | TPR  | TNR  | TNR                                      |
| MultiLayer<br>Perceptron                           | 95.69%                                   | 0.85 | 0.99 | 0.99                                     |
| IB1  | 95.69%                                   | 0.87 | 0.98 | 0.95                                     |
| J48  | 97.50%                                   | 0.95 | 0.98 | 0.95                                     |
| PART   | 97.50%                                   | 0.94 | 0.98 | 0.87                                     |

De la tabla 5.11 se puede observar que el número de falsos positivos generados en la etapa anterior se reduce en gran medida alcanzando una proporción de verdaderos negativos mayor al .95 en la mayoría de los casos.

Los resultados del experimento de comparación entre nuestro método y el basado en wavelets y umbrales de la etapa de reducción de falsos positivos utilizando la base de mastografías MIAS, se muestran en la tabla 5.12.

En términos globales, considerando el desempeño de los cuatro clasificadores y las características seleccionadas, los porcentajes de exactitud de clasificación entre ambos métodos son satisfactorios, alcanzando valores alrededor del 95%.

Tabla 5.12. Resultados de la proporción de clasificación correcta de microcalcificaciones (TPR) y la proporción de verdaderos negativos (TNR) para reducir falsos positivos de la base de imágenes MIAS utilizando diferentes tipos de algoritmos de aprendizaje. El lado izquierdo de la tabla muestra resultados obtenidos con el método basado en FLD propuesto en esta tesis, mientras que del lado derecho se muestran los obtenidos con el método basado en wavelets y umbrales.

| Fase de Reducción de Falsos Positivos |                      |      |      |                                      |      |      |
|---------------------------------------|----------------------|------|------|--------------------------------------|------|------|
| Validación Cruzada de 10 pasos        |                      |      |      |                                      |      |      |
| base MIAS                             |                      |      |      |                                      |      |      |
| Clasificador                          | Método basado en FLD |      |      | Método basado en wavelets y umbrales |      |      |
|                                       | Exactitud %          | TPR  | TNR  | Exactitud %                          | TPR  | TNR  |
| MultiLayer Perceptron                 | 95.69%               | 0.85 | 0.99 | 94.30%                               | 0.95 | 0.93 |
| IB1                                   | 95.69%               | 0.87 | 0.98 | 93.22%                               | 0.91 | 0.94 |
| J48                                   | 97.50%               | 0.95 | 0.98 | 99.72%                               | 0.99 | 0.99 |
| PART                                  | 97.50%               | 0.94 | 0.98 | 99.72%                               | 0.99 | 0.99 |

Analizando únicamente los valores de la proporción de verdaderos positivos TPR, se observa que el método basado en wavelets obtiene resultados un poco más cercanos al 1 que nuestro método, pero aún así nuestros resultados siguen siendo bastante aceptables.

En cuanto al valor de TNR todos los valores están por encima de 0.93, por lo que la proporción de falsos positivos generada tanto por el método basado en FLD como por el basado en wavelets y umbrales es mínima, obteniendo resultados bastante competitivos por ambos métodos.

#### 5.4.2. Experimentos utilizando la base de imágenes ISSSTEP

Los valores de la proporción de verdaderos positivos y verdaderos negativos obtenidos en el experimento para evaluar la etapa de reducción de falsos positivos con nuestro método



utilizando mastografías con y sin microcalcificaciones de la base ISSSTEP, se presentan en la tabla 5.13.

Tabla 5.13. Resultados obtenidos por nuestro método en la etapa de reducción de falsos positivos utilizando mastografías con y sin microcalcificaciones.

| Fase de Reducción de Falsos Positivos<br>base ISSSTEP |  |      |      |  |
|---|--|------|------|--|
|   | Mastografías con<br>microcalcificaciones |      |      | Mastografías sin<br>microcalcificaciones |
| Clasificador  | Exactitud %                              | TPR  | TNR  | TNR                                      |
| MultiLayer<br>Perceptron                              | 99.23%                                   | 1    | 0.99 | 0.92                                     |
| IB1   | 96.74%                                   | 0.90 | 0.98 | 0.92                                     |
| J48   | 95.60%                                   | 0.90 | 0.97 | 0.87                                     |
| PART  | 97.13%                                   | 0.90 | 0.99 | 0.88                                     |

En los resultados de la tabla 5.13, se puede ver que utilizando la base ISSSTEP en general se obtienen mejores resultados que con la base MIAS, obteniendo una proporción de verdaderos positivos por arriba de .90 y una proporción de verdaderos negativos mayor a .92 en la mayoría de los casos.

Los resultados obtenidos en el experimento para comparar nuestro método utilizando la base de mastografías ISSSTEP se pueden observar en la tabla 5.14, el lado izquierdo de la tabla muestra los resultados obtenidos por nuestro método mientras que el lado derecho muestra los resultados obtenidos por el método basado en wavelets y umbrales; ver [16].

Tabla 5.14. Proporción de clasificación correcta de microcalcificaciones (TPR) y la proporción de verdaderos negativos (TNR) para reducir falsos positivos de la base ISSSTEP utilizando diferentes tipos de algoritmos de aprendizaje. El lado izquierdo de la tabla muestra los resultados obtenidos con el método basado en FLD propuesto en esta tesis, mientras que del lado derecho se muestran los obtenidos con el método basado en wavelets y umbrales.

| Fase de Reducción de Falsos Positivos |                      |      |      |                                      |      |      |
|---------------------------------------|----------------------|------|------|--------------------------------------|------|------|
| Validación cruzada de 10 pasos        |                      |      |      |                                      |      |      |
| base ISSSTEP                          |                      |      |      |                                      |      |      |
| Clasificador                          | Método basado en FLD |      |      | Método basado en wavelets y umbrales |      |      |
|                                       | Exactitud %          | TPR  | TNR  | Exactitud %                          | TPR  | TNR  |
| MultiLayer Perceptron                 | 99.23%               | 1    | 0.99 | 87.81%                               | 0.82 | 0.90 |
| IB1                                   | 96.74%               | 0.90 | 0.98 | 83.52%                               | 0.77 | 0.86 |
| J48                                   | 95.60%               | 0.90 | 0.97 | 94.35%                               | 0.92 | 0.95 |
| PART                                  | 97.13%               | 0.90 | 0.99 | 94.35%                               | 0.91 | 0.95 |

Los resultados generados por los dos métodos utilizando la base ISSSTEP, muestran una diferencia notoria en comparación con los obtenidos con la base MIAS. En esta etapa para reducir los falsos positivos, nuestro método obtiene un desempeño de clasificación global por encima del 95% en todos los casos mientras que el método basado en wavelets obtiene porcentajes de desempeño entre el 83 % y 94%, siendo la mayoría de estos resultados aceptables.

Analizando la proporción de verdaderos positivos (TPR), se observa que en todos los casos nuestro método alcanza más de 0.9, mientras que para el método basado en wavelets, el valor de TPR difiere entre un clasificador y otro, oscilando entre 0.7 y 0.9.

En cuanto al número de verdaderos negativos los dos métodos logran reducir en gran medida este valor, alcanzando como máximo un valor para TNR de 0.1

## 5.5. Resultados del desempeño global del método propuesto

Para obtener el desempeño global de nuestro método utilizando el conjunto de mastografías con microcalcificaciones se realizaron dos experimentos. El primer experimento toma el valor de microcalcificaciones detectadas de cada mastografía y se realiza un promedio de los valores de detección obtenidos por cada mastografía; mientras que para el segundo experimento en lugar de tomar en cuenta cada mastografía, se toman en cuenta el total de microcalcificaciones detectadas del conjunto de mastografías, con lo que se obtiene los valores TPR y TNR globales.

Los resultados del desempeño global se muestran tomando en cuenta los resultados obtenidos por el clasificador basado en instancias (IB1), ya que en todas las etapas de nuestro método este clasificador generó buenos resultados. Los resultados obtenidos para los experimentos para medir el desempeño global de nuestro método se observan en la tabla 5.15.

Tabla 5.15. Resultados del desempeño global, utilizando mastografías con microcalcificaciones.

| Desempeño Global utilizando Mastografías con Microcalcificaciones |   |       |  |      |
|---|---|-------|--|------|
|   | Promedio de microcalcificaciones detectadas por mastografía |       | Total de microcalcificaciones detectadas de la base de matografías |      |
|   | TPR   | TNR   | TPR  | TNR  |
| <b>Base MIAS</b>  | 0.885   | 0.907 | 0.9  | 0.86 |
| <b>Base ISSSTEP</b>   | 0.9   | 0.908 | 0.896  | 0.88 |

El promedio de la proporción de microcalcificaciones detectadas por mastografía es aproximadamente de 0.9 para ambas bases de mastografías, mientras que el promedio de la proporción de falsos positivos generados por mastografías es de 0.1.

Si en lugar de analizar cuantas microcalcificaciones fueron detectadas en cada mastografía, obtenemos la proporción total de microcalcificaciones detectadas en todo el conjunto de mastografías, observamos que este valor está por arriba de 0.89 y la cantidad total de verdaderos negativos es aproximadamente de 0.87.

El desempeño global obtenido con las mastografías sin microcalcificaciones son los resultados obtenidos en la etapa de reducción de falsos positivos, ya que al no presentar casos positivos lo que se evalúa es la cantidad de falsos positivos generados.

## **5.6. Discusión**

En esta tesis se presentó un método para la detección de microcalcificaciones que toma en cuenta la densidad del seno. Para la fase de detección se utilizan los discriminantes lineales de Fisher con el fin de encontrar un sub-espacio que permita representar, discriminar e identificar imágenes de microcalcificaciones y de tejido del seno. Además, se agregó una etapa en la cual se extraen diversas características de las regiones detectadas como posibles microcalcificaciones y mediante el uso de algoritmos de aprendizaje se reduce la cantidad de falsos positivos generados en la etapa de detección.

Para la etapa de detección se utilizó como medida de evaluación el análisis FROC (marcas de verdaderos positivos TPR y falsos positivos por imagen FPI), mientras que para la etapa de reducción de falsos positivos en la cual se utilizan clasificadores, se evaluaron los resultados mediante el análisis ROC (proporción de verdaderos positivos TPR y proporción de falsos positivos FPR).

Conforme a la literatura revisada en esta tesis, se identificó que algunos de los algoritmos propuestos para la detección de microcalcificaciones solo toman en cuenta senos catalogados como fáciles (grasos) donde existe una diferencia de contraste evidente entre las microcalcificaciones y el tejido que las rodea.

Otros más, además de tomar en cuenta senos grasos también consideran senos densos (difíciles) en los que no existe gran diferencia entre la microcalcificación y el tejido que la rodea, pero estos algoritmos cuando logran identificar microcalcificaciones en senos densos generan un gran número de falsos positivos.

De acuerdo a los experimentos realizados y a los resultados obtenidos, se puede ver que nuestro método, al tomar en cuenta la densidad del seno, es capaz de detectar microcalcificaciones no solo en senos catalogados como fáciles (grasos), sino que también identifica microcalcificaciones en senos completamente densos (difíciles) y con la ventaja de generar un número menor de falsos positivos aun con mastografías que no presentan microcalcificaciones.

Al realizar los experimentos para la etapa de detección de microcalcificaciones, se identificó que probablemente debido a que los FLD no se basan en la variación total para

encontrar el espacio discriminante y tomando en cuenta que la iluminación de una imagen representa una de las mayores variaciones [5], nuestro método identifica microcalcificaciones de todo tipo de nivel de intensidad, aún incluso en mastografías opacas, a diferencia del método basado en wavelets el cual debido a que se basa en umbrales locales y globales, no logra poseer esta característica.

Nuestro método generó resultados prometedores, estando éstos por encima de los obtenidos con el método basado en wavelets y umbrales en aproximadamente un 10% para la etapa de detección de microcalcificaciones y reduciendo en gran medida la cantidad de falsos positivos de esta etapa.

Hablando del desempeño global de nuestro método, se alcanzo una proporción promedio de microcalcificaciones detectadas aproximadamente de 0.9, generando menos de 0.1 de proporción de falsos positivos por mastografías

El porcentaje global de los métodos de detección de microcalcificaciones que constan de 2 etapas (segmentación y reducción de falsos positivos) se obtiene con los porcentajes obtenidos en cada etapa. Es fácil ver que el porcentaje de detección de la etapa de reducción de falsos positivos depende del porcentaje de detección obtenido en la etapa de segmentación, esto debido a que la cantidad de microcalcificaciones que pueden ser detectadas en la etapa de reducción de falsos positivos será a lo más la cantidad de microcalcificaciones que se detectaron en la etapa de segmentación.

Tomando en cuenta lo anterior, aunque en la etapa de reducción de falsos positivos el método propuesto para la detección de microcalcificaciones basado en FLD y el método basado en wavelets y umbrales obtienen un desempeño muy parecido, el método basado en FLD obtiene un mejor desempeño global ya que obtiene un mayor porcentaje de microcalcificaciones detectadas en la etapa de segmentación.

Después de realizar los experimentos y analizar los resultados identificamos que nuestro método logra detectar microcalcificaciones de diversos tamaños, niveles de intensidad de gris y formas, no solo aquellas que estén presentes en senos grasos sino también en senos densos.



## **Capítulo 6**

### **Conclusiones**

#### **6.1. Sumario**

El presente trabajo de tesis consistió en el diseño de un método para la detección de microcalcificaciones presentes en mastografías. El método propuesto está compuesto por tres etapas principales: clasificación de la densidad del seno, detección de microcalcificaciones mediante el uso de los Discriminantes Lineales de Fisher y reducción de falsos positivos.

La etapa de clasificación de la densidad del seno permite identificar cuál es el tejido que predomina en el mismo (denso o graso). Para lograr esto, se extrae la parte correspondiente a la imagen del seno del fondo de la mastografía mediante el uso del algoritmo Fuzzy c-means. Posteriormente se calcula el histograma del seno y a partir de éste se extrae un conjunto de características como son la media, desviación estándar y frecuencia, las cuales permiten identificar si se trata de un seno denso o de un seno graso.

En la fase de segmentación de microcalcificaciones, para la cual se utiliza un conjunto de ejemplos de imágenes de microcalcificaciones e imágenes de tejido del seno, se aplicaron los Discriminantes Lineales de Fisher con el fin de encontrar un espacio de menor

dimensión que resalte las diferencias entre estos dos tipos de imágenes, facilitando la detección de microcalcificaciones. Este conjunto de ejemplos de imágenes son representados como vectores y posteriormente son proyectados al espacio encontrado.

Se crea un espacio para cada tipo de densidad del seno (denso o graso), un espacio con ejemplos de imágenes de microcalcificaciones y tejido del seno provenientes de senos grasos y otro con imágenes provenientes de senos densos.

Cuando se analiza una nueva mastografía en busca de microcalcificaciones, ésta primero se clasifica en seno graso o seno denso, para posteriormente recorrer la imagen del seno con una ventana de tamaño fijo. La sub-imagen dentro de la ventana se representa con un vector y se proyecta en el espacio discriminante que le corresponde dependiendo de si el seno es graso o denso. Una vez proyectada la imagen se identifica a que clase se asemeja más (microcalcificación o tejido del seno). En caso de que se clasifique como microcalcificación, se marcan en la mastografía con color amarillo las coordenadas donde se encuentra la ventana para señalar que en esa posición existe una posible microcalcificación.

Como paso final se calculan 15 características de forma y color para cada región identificada como posible microcalcificación y mediante el uso de diversos algoritmos de aprendizaje automático (tomando como entrada las características extraídas de dichas regiones), se reduce la cantidad de falsos positivos.

Para cada una de las tres etapas de nuestro método se realizaron experimentos con el fin de evaluar su desempeño con las bases de mastografías MIAS e ISSSTEP. Para la etapa de clasificación de densidad del seno se utilizó como medida de evaluación el análisis ROC, tomando en cuenta la tasa de verdaderos positivos TPR y la tasa de falsos positivos FPR. Se seleccionaron tres algoritmos de aprendizaje implementados en Weka, entre ellos una red neuronal (MultiLayer Perceptron), un algoritmo basado en instancias (IB1) y uno basado en reglas (NNge). Para medir el desempeño se utilizó validación cruzada de 10 pasos, los resultados obtenidos en estos experimentos fueron satisfactorios, alcanzando valores de precisión general por encima del 92%.

En la etapa de segmentación de microcalcificaciones se utilizó la métrica de evaluación de análisis FROC, con la que se contempla la tasa de marcas de verdaderos positivos TPR y las marcas de falsos positivos por imagen FPI. Además de los experimentos enfocados en



evaluar nuestro método se hizo una comparación con el método basado en wavelets y umbrales propuesto por Flores y González.

Los resultados obtenidos para la segmentación de microcalcificaciones de nuestro método demostraron ser mejores y bastante satisfactorios en todos los experimentos realizados, alcanzando valores entre 91% y 93% de detección, y aún mejor con una cantidad de falsos positivos baja, superando el desempeño del método basado en wavelets.

Para la evaluación de la etapa de reducción de falsos positivos se utilizó el análisis ROC. La evaluación se realizó mediante validación cruzada de 10 pasos, se experimentó con cuatro clasificadores implementados en Weka, una red neuronal (MultiLayerPerceptron), uno basado en instancias (IB1), un árbol de decisión (J48) y uno basado en reglas (PART), obteniendo valores de TPR mayores a 0.9. Además, se realizó una comparación con la etapa de reducción de falsos positivos del método propuesto por Flores y González cuyos resultados mostraron ser bastantes buenos, pero aún así nuestro método igualó y mejoró estos resultados.

De acuerdo a los experimentos realizados, nuestro método mostró ser capaz de identificar microcalcificaciones de diversas formas, tonalidades y tamaños. Además, tomando en cuenta la información de la clasificación de la densidad del seno en la etapa de detección, nuestro método mostró resultados bastante satisfactorios al identificar microcalcificaciones en senos en que la diferencia de contraste del tejido del seno con la calcificación es pequeña (senos densos).

## **6.2. Aportación**

La principal contribución de esta tesis es:

- Se propuso un nuevo método para la detección de microcalcificaciones que toma en cuenta la densidad del seno, el cual permite identificar microcalcificaciones de diversas formas, tamaños y tonalidades no solo en senos grasos sino también en senos densos (difíciles), además de disminuir el número de falsos positivos detectados.

### **6.3. Conclusiones**

El método para la detección de microcalcificaciones basado en Discriminantes Lineales de Fisher propuesto en esta tesis, el cual toma en cuenta la densidad del seno, demostró obtener un mejor desempeño obteniendo resultados bastante satisfactorios en todos los experimentos realizados en cada una de las etapas que lo componen.

De acuerdo a los experimentos realizados, se identificó que debido a que los Discriminantes Lineales de Fisher no se basan en la variación total de los datos y la iluminación de una imagen es una de la principales variaciones, nuestro método es capaz de identificar tanto microcalcificaciones con tonalidades brillantes como microcalcificaciones opacas, permitiendo aún en mastografías muy oscuras detectar las microcalcificaciones correctamente, sin realizar un pre-procesamiento a la imagen.

También se observó que al utilizar la clasificación de la densidad del seno en nuestro método se incrementó la detección de microcalcificaciones en más del 10 %, permitiendo identificar microcalcificaciones en senos densos sin incrementar el número de falsos positivos.

El hecho de segmentar el seno del fondo permite reducir aún más el número de falsos positivos ya que los defectos de la mastografía fuera del seno como son las etiquetas y artificios en el fondo, no son tomados en cuenta al momento de analizar la mastografía.

### **6.4. Trabajo futuro**

Como trabajo futuro, además de detectar las microcalcificaciones en la mastografía, se podría agregar un módulo de diagnóstico que indique la probabilidad de malignidad de la microcalcificación utilizando la forma y el número de éstas, incluso agregar información de la ubicación de las microcalcificaciones en el seno (conductos mamarios, lobulillos, tejido graso) para identificar el grado de malignidad de éstas.

Una idea para incrementar aun más el desempeño del método propuesto en esta tesis, es mejorar la etapa dedicada a la clasificación de la densidad del seno, proponer un clasificador que permita identificar exactamente que regiones en el seno corresponden a

tejido graso y que regiones pertenecen a tejido denso. Con esto se podría utilizar diferentes espacios discriminantes de manera dinámica en el análisis de una misma mastografía.

Un camino diferente sería fusionar la información de las posiciones de las microcalcificaciones de la mastografía medio-lateral y de la cráneo caudal con el fin de eliminar falsos positivos.

Otra dirección de trabajo futuro sería aplicar el método para la detección de tumores u otras lesiones en una mastografía. Incluso transportar y adaptar a un nivel más genérico el método propuesto en esta tesis, donde se puedan detectar diversos objetos de interés, aún con diferencias de iluminación y rodeados por diferentes tipos de texturas.

Una distinta dirección orientada a la práctica, podría ser desarrollar e implantar el método propuesto en un Hospital con la finalidad de llevarlo más allá de lo teórico, proporcionando una herramienta de ayuda a radiólogos en la tarea de la detección de microcalcificaciones.



## Bibliografía

- [1] Adobe Systems Incorporated. <http://www.adobe.com/es/products/fireworks> (último acceso: Junio, 2008).
- [2] Anderson, James. Redes Neuronales. AlfaOmega, 2007.
- [3] Anton, Howard. Anton Introducción al algebra lineal. Limusa Wiley, 2003.
- [4] Bejar, Javier. «Inteligencia Artificial.» <http://www.lsi.upc.es/~bejar/ia/material/teoria/aprendizaje.pdf> (último acceso: Abril, 2009).
- [5] Belhumeur, Peter, Joao Hespanha, y David Kriegman. «Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection.» IEEE Transaction on Pattern Analysis and Machine Intelligence 19, nº 7 (1997): 711-720.
- [6] Berry, Michael, y Gordon Linoff. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Wiley Computer Publishing, 2004.
- [7] Bewick, Viv, Liz Cheek, y Jonathan Ball. «Statistics review 13: Receiver Operating Characteristic Curves.» (Critical Care) 8, nº 6 (2004).
- [8] Bovik, Al, y Jerry Gibson. Handbook of Image and Video Processing (Communications, Networking and Multimedia). Academic Press, 2000.
- [9] Cevikal, H, M Neamtu, M Wilkes, y A Barkana. «Discriminative Common Vectors for Face Recognition.» IEEE Transactions on Pattern Analysis and Machine Intelligence 27, nº 1 (Enero,2005): 4-13.
- [10] Chakraborty, DP, y LH Winter. «Free-Response Methodology: Alternate Analysis and a New Observer- Performance Experiment.» Radiology 174 (1990): 873-881.
- [11] Cheng, HD, Xiaopeng Cai, Xiaowei Chen, Living Hu, y Xueling Lou. «Computer Aided Detection and Classification of Microcalcifications in Mammograms: a survey.» The Journal of Pattern Recognition 36 (2003): 2967-2991.
- [12] Choi, J, y SH Nam. «A Methods of Image Enhancement and Fractal Dimension for Detection of Microcalcifications in Mammogram.» Proceedings of the 20th annual international conference of the IEEE Engineering in Medicine and Biology Society 20 (1998): 1009-1012.

- [13] Cura, Jorge, Marco Cura, Alejandro Cura, y Carlos Menke. Manual de microcalcificaciones mamarias, factor predictivo de lesiones mamarias. JMA-Cura.
- [14] Davies, H, y D Dance. «Automatic Computer Detection of Clustered Calcifications in Digital Mammograms.» *Physics in Medicine and Biology* 35 (1990): 1111-1118 .
- [15] Dengler, Joachin, Sabina Behrens, y Joahnn Desaga. «Segmentation of Microcalcifications in Mammograms.» *IEEE Transactions on Medical Imaging* 12, nº 4 (1993): 634 - 642.
- [16] Flores, Beatriz, y Jesús González. «Data Mining with Decision Trees and Neuronal Networks for Calcifications Detection in Mammograms.» *LNSC* 2972/2004 (2004): 232-241.
- [17] Frigas, Antonys. Microcalcifications resource site. <http://infoacademy.gr/microcalc> (último acceso: Abril, 2009).
- [18] Fu, JC, SK Lee, S Wong, JY Yeh, AH Wang, y HK Wu. «Image Segmentation Feature Selection and Pattern Classification for Mammographic Microcalcifications.» *Computer Medical Imaging and Graphics* 29 (2005): 419-429 .
- [19] Fu, K, y J Mui. «A Survey on Image Segmentation.» *Pattern Recognition* 13, nº 1 (1981): 3-16.
- [20] García, M, y otros. «Global Cancer Facts & Figures 2007.» American Cancer Society.[http://www.cancer.org/downloads/STT/Global\\_Facts\\_and\\_Figures\\_2007\\_rev2.pdf](http://www.cancer.org/downloads/STT/Global_Facts_and_Figures_2007_rev2.pdf) (último acceso: Noviembre, 2008).
- [21] Gonen, Mithat. «Receiver Operating Characteristic ROC, Statistic and Data Analysis.» <http://www2.sas.com/proceedings/sugi31/210-31.pdf> (último acceso: Febrero, 2009).
- [22] González, Rafael, Richard Woods, y Steven Eddins. *Digital Image Processing using Matlab*. Prentice Hall, 2004.
- [23] Gutierrez Osuma, Ricardo. «Introduction to Pattern Analysis.» [http://research.cs.tamu.edu/prism/lectures/pr/pr\\_11.pdf](http://research.cs.tamu.edu/prism/lectures/pr/pr_11.pdf) (último acceso: Noviembre, 2008).
- [24] Hand, David, Heikki Mannila, y Padhraic Smyth. *Principles of data mining*. MIT Press, 2001.
- [25] Hernández Orallo, José, José Ramírez Quintana, y César Ferri Ramírez. *Introducción a la minería de datos*. Pearson, 2004.

- [26] Jonathon, Shlens. «A tutorial on Principal Components Analysis.» Vers. 2. <http://www.sn1.salk.edu/~shlens/pub/notes/pca.pdf> (último acceso: Octubre, 2008).
- [27] Kim, J, y H Park. «Surrounding Region Dependence Method for Detection of Clustered Microcalcifications on Mammograms.» ICIP'97 3 (1997): 535.
- [28] Kopans, D.B. «Breast Imaging.» (Annals of Oncology Springer) 9, nº 12 (1998).
- [29] Lee, Yougbum, y Du Tsai. «Computerized Classification of Microcalcifications on Mammograms using Fuzzy Logic and Genetic Algorithm.» Proceedings of SPIE Medical Imaging 5370 (2004): 952-957.
- [30] McAndrew, Alasdair. Introduction to Digital Image Processing with Matlab. Thomson, 2004.
- [31] Mitchell, Tom M. Machine Learning. McGraw-Hill, 1997.
- [32] Morishita, Mariko, Akira Ohtsuru, Ichiro Isomoto, y Suchini Yamashita. «Categorization of Mammographic Density for Breast Cancer: Clinical Significance.» En Cancer Imaging: Lung and breast carcinoma, de M.A. Hayat. Academic Press, 2007.
- [33] Morrow, W, R.B. Paranjape, R.M. Rangayyan, y J.E. Desautels. «Region Based Contrast Enhancement of Mammograms.» IEEE Transactions Medical Imaging 11, nº 3 (1992): 392-406.
- [34] Mousa, Rafayah, Qutaishat Munib, y Abadía Moussa. «Breast Cancer Diagnosis System based on Wavelet Analysis and Fuzzy Neural.» Elsevier Expert Systems with Applications 28, nº 4 (2005): 713-723 .
- [35] Nishikawa, Robert, Issam El- Naqa, Yongyi Yang, Miles Wernick, y Nicolás Galatsanos. «Support Vector Machine learning for Detection of Microcalcifications in Mammograms.» IEEE Transactions on Medical Imaging 21, nº 12 (2002): 201 - 204.
- [36] Nishikawa, Robert, Yulei Jiang, y Maryellen Giger. «Computer Detection of Clustered Microcalcifications.» Proceedings of the IEEE International Conference on Systems 2, nº 18-21 (1992): 1375- 1378.
- [37] Otsu, N. «A threshold selection method from gray-level histograms.» IEEE Transactions on Systems, 9 nº 1 (1979). 62-66.
- [38] Oliver, Arnau. <http://eia.udg.es/~aoliver/> (último acceso: Junio, 2008).

- [39] Pajares, G, y J De la Cruz. Visión por computador. Imágenes digitales y aplicaciones. Ra-Ma, 2001.
- [40] Prieto, Alejandro. «CENETEC.» Mastografía.  
[http://www.cenetec.gob.mx/interior/gaceta5/inf\\_pacientes1.html](http://www.cenetec.gob.mx/interior/gaceta5/inf_pacientes1.html) (último acceso: Marzo, 2009).
- [41] Secretaria de Salud México. « Estadísticas de Mortalidad en México.» (Salud Publica de México) 44, n° 6 (2002).
- [42] Shen, L, R Rangayyan, y J.E. Desautles. «Detection and Classification of Mammographic Calcifications.» International Journal of Pattern Recognition and Artificial Intelligence 7 (1993): 1403-1416 .
- [43] Smith, Linsay.  
[http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf)  
(último acceso: Marzo, 2009).
- [44] Suckling, J., J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis. I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal and J. Savage. «The Mammographic Image Analysis Society digital mammogram database.» Exerpta Medica. 1994. 375-378.
- [45] Teknomo, Kaédi. «Linear Discriminant Analysis - Tutorial.» <http://people.revoledu.com/Kardi/tutorial/LDA/LDA.html> (último acceso: Octubre, 2008).
- [46] Thangavel, K, M Karnan, R Sivakumar, y K Mohideen. «Automatic Detection of Microcalcifications in Mammograms.» Journal of Graphics Vision and Image Processing 5 (2005): 31-61.
- [47] Turk, Matthew, y Alex Pentland. «Face Recognition using Eigenfaces.» IEEE Computer Vision and Pattern recognition, 1991: 586-591.
- [48] Vernon, David. Machine Vision, Automated Visual Inspection and Root Vision. Prentice Hall, 1991.
- [49] Wallet, BC, JL Solka, y CE Pruebe. «A Method for Detecting Microcalcifications in Digital Mammograms.» Journal of digital imaging 10 (1997): 136-139.
- [50] Wang, Ted, y Nicolaos Karayiannis. «Detection of Microcalcifications in Digital Mammograms using Wavelets.» IEEE Transactions on Medical Imaging 17, n° 4 (1998): 498 - 509.
- [51] Welling, Max. <http://www.ics.uci.edu/~welling/classnotes/classnotes.html> (último acceso: Noviembre, 2008).



- [52] Witten, Ian H., y Eibe Frank. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 2005.
- [53] Witten, Ian H. y Eibe Frank. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann, 1999.
- [54] Woods, KS, CC Doss, KW Bowyer, JL Solka, CE Pruebe, y WP Kegelmeyer. «Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography.» International Journal of Pattern Recognition and Artificial Intelligence 7 (1993): 1417-1436 .
- [55] Wu, Yuan, Qiang Huang, YongHong Peng, y Wuchao Situ. «Detection of Microcalcifications Based on Dual Threshold.» LNCS 4046/2006 (2006): 347-354.
- [56] Yaffe, Martin, Jeffrey Byng, y Norman Boyd. «Quantitative Image Analysis for Estimation of Breast Cancer Risk.» Handbook of medical imaging, 323-340. Academic Press, 2000.
- [57] Zhang, L, W Qian, R Sankar, D Song, y R Clark. «A New False Positive Reduction Method for MCCs Detection in Digital Mammography.» IEEE International Conference on Acoustic, Speech in Signal Processing 2 (2001): 1033-1036.
- [58] Zhang, Wei, Hiroyuki Yoshida, Robert Nishikawa, y Kunio Doi. «Optimally Weighted Wavelet Transform Based on Supervised Training for Detection of Microcalcifications in Digital Mammograms.» Medical Physics 25, nº 6 (1998): 949-956.
- [59] Zhao, D, M Shridhar, y DG Daul. «Morphology on Detection of Calcifications in Mammograms.» IEEE International Conference on Acoustics, Speech and Signal Processing 3 (1992): 129-132.