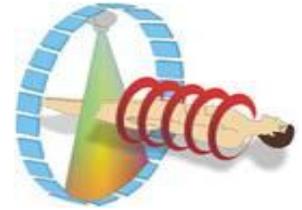




**UNIVERSIDAD DE LOS ANDES  
FACULTAD DE INGENIERIA  
DECANATO DE ESTUDIOS DE POSTGRADO  
MAESTRIA EN INGENIERIA BIOMEDICA**



# **HERRAMIENTA DE SOFTWARE PARA LA DETECCIÓN DE LA ESTEATOSIS HEPÁTICA EN IMÁGENES MÉDICAS**

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

Tesis de Maestría

Presentada ante la Universidad de Los Andes

por

**Elymar Coromoto Rivas Newman**

Como requisito parcial para optar al Grado de

**Magister Scientiae en Ingeniería Biomédica**

Realizada bajo la Tutoría del Profesor

**Dr. Rubén de Jesús Medina**

Diciembre 2015

## AGRADECIMIENTOS

A Dios y a la Santísima Virgen, por llenarme de salud y fuerzas para cumplir esta meta y seguir siempre adelante.

A mis padres Marnolio y Eliana, que siempre han sido un gran apoyo espiritual e incondicional para que cada día de mi vida sea lleno de mucha alegría y entusiasmo para lograr todos mis sueños.

A Daniel, por ayudarme, darme ánimo y apoyo para lograr otra de mis metas.

A mis hermanos Marely, Magnolio y Julio, que me han brindado siempre esas palabras de ayuda para continuar luchando cada día sin decaer en el camino.

A mis sobrinos y cuñados por brindarme su amistad y energía.

A mi amiga Crissel con la que compartí mi estancia en Mérida y me enseñó que con perseverancia se logran las metas. Gracias por tu amistad.

A mis profesores del postgrado en especial al Profesor Rubén Medina por enseñarme y tener la paciencia para poder culminar este trabajo.

A todos los compañeros que de alguna u otra forma estuvieron presentes en este camino y que con su amistad me brindaron su apoyo.

Al los Dres. José Rivas, Javic González y Gustavo Rondón por facilitarme las imágenes para poder realizar este trabajo.

A mi familia que siempre estuvo pendiente de darme palabras de entusiasmo para seguir adelante en este camino.

*A todos... Muchas Gracias y que Dios los bendiga!*

## Resumen

La esteatosis hepática (EH) es una de las enfermedades difusas del hígado caracterizada por un desequilibrio en la síntesis hepatocítica de triglicéridos a partir de ácidos grasos. El objetivo de este trabajo es detectar esta patología aplicando técnicas de procesamiento digital de imágenes mediante el uso de imágenes tomográficas compuesta por sus fases: simple, arterial y venosa; y por imágenes de ultrasonido. La técnica de procesamiento de imágenes aplicada es la extracción de características de textura. Para ello se desarrolló un algoritmo en Matlab que consiste en seleccionar de la imagen médica un polígono que representa el hígado. Dicha región se subdivide en una cuadrícula con ventanas de tamaño 4x4. Posteriormente se calcula la matriz de co-ocurrencias y algunos parámetros de textura. Los atributos extraídos de imágenes de pacientes con EH se comparan con los atributos extraídos de imágenes de pacientes sanos mediante las técnicas de Máquinas de Soporte Vectorial de mínimos cuadrados (LSSVM) y Regresión Logística Binaria (RLB). La LSSVM fue entrenada con 12 pacientes sanos y 24 enfermos a través de los atributos obtenidos de imágenes provenientes de ultrasonido. Para la validación se generaron 7 muestras aleatorias conformadas por 5 pacientes sanos y 6 enfermos. En las imágenes tomográficas se consideró sólo la fase venosa, por presentar los mejores resultados de clasificación. La LSSVM fue entrenada con 4 pacientes sanos y 4 pacientes enfermos, y validados los resultados con 7 muestras aleatorias constituidas por 4 pacientes con esteatosis hepática y 2 pacientes sanos. Las mismas muestras fueron usadas para clasificar a los pacientes usando el modelo de RLB. Como resultado se obtuvo que la sensibilidad en ambos métodos fue mayor a 90% y la especificidad mayor a 66.67% en ambas modalidades de imágenes. Y al comparar las dos técnicas se puede concluir que la RLB detecta mejor la EH y ambas técnicas coinciden con que la clasificación para las imágenes de US es superior que para la TC.

**Palabras Claves:** *Esteatosis Hepática, Tomografía Computarizada, Ultrasonido, Matriz de Co-ocurrencias, Máquinas de Soporte Vectorial, Regresión Logística Binaria.*

# ÍNDICE GENERAL

<b>1. Introducción</b>	<b>13</b>
1.1 Motivación y Planteamiento del Problema .....	14
1.2 Antecedentes .....	15
1.3 Objetivos .....	17
1.3.1 Objetivo General .....	17
1.3.2 Objetivos Específicos .....	17
1.4 Metodología .....	18
1.4.1 Criterios de Inclusión .....	18
1.4.2 Criterios de Exclusión .....	18
1.5 Organización del Trabajo .....	18
<b>2. Anatomía del Hígado e Imagenología Médica</b>	<b>20</b>
2.1 Generalidades .....	20
2.1.1 Morfología General .....	20
2.1.2 Situación .....	21
2.1.3 Proyección Esquelética .....	21
2.2 Caras del Hígado .....	22
2.2.1 Cara Superior o Diafragmática .....	22
2.2.2 Cara Inferior .....	22
2.2.3 Cara Posterior .....	23
2.3 Divisiones Morfológicas .....	23
2.4 Segmentación Hepática .....	24
2.5 Histología Hepática .....	26
2.6 Esteatosis Hepática .....	27
2.6.1 Esteatosis Macrovesicular y Microvesicular .....	28

2.6.2 Diagnóstico de la Esteatosis Hepática .....	28
2.6.3 Tratamiento de la Esteatosis Hepática .....	30
2.7 Imagenología Médica .....	30
2.7.1 Rayos X .....	30
2.7.2 Tomografía Computarizada .....	31
2.7.2.1 Recuerdo Histórico .....	31
2.7.2.2 Vóxel, Pixel y Matriz .....	32
2.7.2.3 Fundamentos de la Tomografía Computarizada .....	33
2.7.2.4 Densidad y Escala de Grises .....	34
2.7.2.5 Selección de Ventana y Nivel de Ventana .....	35
2.7.2.6 Unidades de Tomografía Computarizada .....	36
2.7.2.7 Técnicas de Exploración .....	38
2.7.3 Ecografía Hepática .....	40
2.7.4 Resonancia Magnética .....	42
2.7.4.1 Resonancia Magnética del Hígado .....	43
2.8. Conclusiones .....	44
<b>3. Matriz de Co-ocurrencias y Análisis Multivariante</b> .....	<b>46</b>
3.1 Matriz de Co-ocurrencias .....	46
3.1.1 Parámetros de Textura .....	47
3.2 Análisis de Varianza Multivariante .....	48
3.2.1 Modelo MANOVA con un tratamiento .....	48
3.2.1.1 Relación entre diversos contrastes .....	50
3.2.2 Modelo MANOVA con dos o más tratamientos .....	51
3.3 Distribución $T^2$ Hotelling .....	52
3.4 Metodología Aplicada .....	53
3.5 Resultados Obtenidos .....	55
3.5.1 Fase Simple .....	56
3.5.2 Fase Arterial .....	59

3.5.3 Fase Venosa .....	62
3.6. Conclusiones .....	65
<b>4. Máquinas de Soporte Vectorial y Regresión Logística Binaria</b> .....	<b>66</b>
4.1 Conceptos Fundamentales Máquinas de Soporte Vectorial .....	66
4.1.1 Clasificación con Vectores de Soporte .....	66
4.2 Máquinas de Soporte Vectorial de Mínimos Cuadrados LSSVM ....	67
4.2.1 Software de LSSVM .....	71
4.3 Regresión Logística .....	72
4.3.1 Regresión Logística Binaria .....	74
4.3.1.1 Estimación de los coeficientes del Modelo .....	76
4.3.1.2 Estadístico de Wald .....	76
4.3.1.3 Estadístico G de Razón de Verosimilitud .....	77
4.4 Metodología Aplicada .....	77
4.4.1 Máquinas de Soporte Vectorial de Mínimos Cuadrados .....	77
4.4.2 Análisis de Regresión Logística Binaria .....	78
4.4.3 Curvas ROC .....	78
4.4.4 Prueba T de Wilcoxon .....	79
4.4.5 Estadístico de Contraste .....	79
4.4.6 Prueba U de Mann – Whitney .....	80
4.4.7 Estadístico de Contraste .....	80
4.5 Análisis de la Base de Datos .....	81
4.5.1 Entrenamiento del Modelo .....	82
4.5.2 Validación del Modelo .....	82
4.5.3 Clasificación de los Pacientes usando el Modelo de RLB ....	82
4.5.4 Modelo Estadístico para Imágenes de Tomografía .....	83
4.5.5 Modelo Estadístico para Imágenes de Ultrasonido .....	83
4.5.6 Validación del Modelo .....	84
4.5.7 Análisis de Resultados .....	86

4.6 Clasificación de los Pacientes usando MSV .....	87
4.6.1 Etapa de Entrenamiento para Imágenes de TC .....	88
4.6.2 Etapa de Validación para Imágenes de TC .....	88
4.6.3 Etapa de Entrenamiento para Imágenes de US .....	89
4.6.4 Etapa de Validación para Imágenes de US .....	89
4.6.5 Análisis de Resultados .....	89
4.7 Conclusiones .....	94
<b>5. Conclusiones Generales y Recomendaciones</b> .....	<b>95</b>
5.1 Conclusiones Generales .....	95
5.2 Recomendaciones .....	97

## **Bibliografía**

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

## ÍNDICE DE FIGURAS

2.1. <i>Órganos del Abdomen</i> .....	21
2.2. <i>Proyección Esquelética</i> .....	22
2.3. <i>Cara Antero Posterior del Hígado</i> .....	22
2.4. <i>Cara Postero Inferior del Hígado</i> .....	23
2.5. <i>Esquemas de la Segmentación Hepática</i> .....	26
2.6. <i>Imágenes Histológicas representativas de la Esteatosis Hepática</i> .....	28
2.7. <i>Cortes: Sagital (A), Coronal (B), Axial (C)</i> .....	31
2.8. <i>Un Vóxel corresponde al área de un pixel por espesor de corte</i> .....	32
2.9. <i>Vóxel, Pixel y Matriz</i> .....	33
2.10. <i>Fundamento de un Equipo de Tomografía Computarizada</i> .....	34
2.11. <i>Escalas de números TC del cuerpo humano</i> .....	35
2.12. <i>Selección de Ventana y nivel de ventana</i> .....	36
2.13. <i>Unidad de TC de Tercera Generación</i> .....	37
2.14. <i>Fase Simple de la TC</i> .....	38
2.15. <i>Fase Arterial o Vascular de la TC</i> .....	39
2.16. <i>Fase Venosa o Tardía de la TC</i> .....	40
2.17. <i>Hígado muy hiperecogénico en comparación con la corteza renal en un paciente con EH</i> .....	41
2.18. <i>Hígado Hiperecogénico</i> .....	41
2.19. <i>Esteatosis Hepática con conservación focal de la grasa en el segmento medial del lóbulo izquierdo y alrededor de la vesícula biliar</i> .....	44
3.1. <i>Imagen Tomográfica, hígado</i> .....	53
4.1. <i>Transformación del espacio de entrada al espacio característico</i> .....	67
4.2. <i>Esquema de dos conjuntos linealmente mediante un hiperplano (<math>P_0</math>) de separación óptimo <math>w^T \cdot x + b</math> que maximiza el margen <math>m</math></i> .....	68
4.3. <i>Transformación del espacio de entrada al espacio característico</i> .....	70
4.4. <i>Curva ROC para clasificación de matrices 4x4 en imágenes de TC</i> .....	85

4.5. Curva ROC para clasificación de matrices 4x4 en imágenes de US .....	85
4.6. Curva ROC de TC .....	91
4.7. Curva ROC de US .....	92

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

## ÍNDICE DE TABLAS

3.1. <i>Análisis de Varianza para un modelo con dos tratamientos e interacción ...</i>	51
3.2. <i>Descripción de los Datos .....</i>	55
3.3. <i>Estadísticas Básicas de Pacientes Sanos en Fase Simple de la TC .....</i>	56
3.4. <i>Estadísticas Básicas de Pacientes con Esteatosis Hepática en Fase Simple de la TC .....</i>	56
3.5. <i>ANOVA para Contraste, Correlación, Energía, Homogeneidad en Fase Simple de la TC.....</i>	57
3.6. <i>Análisis Multivariante para Contraste, Correlación, Energía, Homogeneidad en pacientes con EH y pacientes sanos en Fase Simple de la TC.....</i>	58
3.7. <i>MANOVA modelo Contraste, Correlación, Energía, Homogeneidad = Diagnóstico en Fase Simple de la TC .....</i>	58
3.8. <i>Estadísticas Básicas de Pacientes Sanos en Fase Arterial de la TC .....</i>	59
3.9. <i>Estadísticas Básicas de Pacientes con Esteatosis Hepática en Fase Arterial de la TC .....</i>	59
3.10. <i>ANOVA para Contraste, Correlación, Energía, Homogeneidad en Fase Arterial de la TC.....</i>	60
3.11. <i>Análisis Multivariante para Contraste, Correlación, Energía, Homogeneidad en pacientes con EH y pacientes sanos en Fase Arterial de la TC .....</i>	61
3.12. <i>MANOVA modelo Contraste, Correlación, Energía, Homogeneidad = Diagnóstico en Fase Arterial de la TC .....</i>	61
3.13. <i>Estadísticas Básicas de Pacientes Sanos en Fase Venosa de la TC .....</i>	62
3.14. <i>Estadísticas Básicas de Pacientes con Esteatosis Hepática en Fase Venosa de la TC .....</i>	62
3.15. <i>ANOVA para Contraste, Correlación, Energía, Homogeneidad en Fase Venosa de la TC.....</i>	63
3.16. <i>Análisis Multivariante para Contraste, Correlación, Energía,</i>	

<i>Homogeneidad en pacientes con EH y pacientes sanos en Fase Venosa de la TC.....</i>	64
<i>3.17. MANOVA modelo Contraste, Correlación, Energía, Homogeneidad = Diagnóstico en Fase Venosa de la TC .....</i>	64
<i>4.1. Puntos de corte óptimos para clasificar a los pacientes con esteatosis hepática, en base al modelo de RLB.....</i>	86
<i>4.2. Resultados de clasificación de pacientes con esteatosis hepática. La sensibilidad (<math>S_e</math>) y la especificidad (<math>S_p</math>) son calculados por cada modalidad de imagen.....</i>	86
<i>4.3. Resultados de la clasificación de pacientes con esteatosis hepática de imágenes de TC usando máquinas de soporte vectorial .....</i>	90
<i>4.4. Resultados de la clasificación de pacientes con esteatosis hepática de imágenes de ultrasonido usando máquinas de soporte vectorial .....</i>	90
<i>4.5. Resultado del Área bajo la Curva para TC.....</i>	91
<i>4.6. Tabla de Contingencia para TC.....</i>	91
<i>4.7. Resultado del Área bajo la Curva para US.....</i>	92
<i>4.8. Tabla de Contingencia para US.....</i>	93

## Capítulo I

### Introducción

La obesidad es un factor de riesgo común y bien conocido para esteatosis hepática. Tanto un excesivo índice de masa corporal (IMC) como la obesidad son reconocidos como posibles desencadenantes. En pacientes con obesidad severa quienes van a cirugía bariátrica, la prevalencia de Esteatosis hepática puede ser mayor al 90% y 5% de ellos pueden tener cirrosis no diagnosticada [1].

La obesidad y la diabetes están aumentando a un ritmo alarmante en los Estados Unidos y alrededor del mundo. Recientemente las estimaciones sugieren que la prevalencia de la obesidad ha aumentado el 33,8% en general y la prevalencia de la diabetes en adultos entre 45-64 años de edad se ha incrementado hasta el 10,6% en los Estados Unidos. La prevalencia del síndrome metabólico, que es una compilación de 5 factores clínicos que abarcan la obesidad y diabetes, también está aumentando. La Esteatosis Hepática (patología hepática de síndrome metabólico), es ahora la principal causa de la elevación anormal de las enzimas hepáticas y se ha asociado con un mayor riesgo de desarrollar diabetes y arterioesclerosis. La prevalencia de Esteatosis hepática ha sido estimada en un rango que oscila entre el 2,8 y el 46% alrededor del mundo y varía de acuerdo a la población estudiada y a las herramientas diagnósticas utilizadas para su detección [2].

El Gold Estándar para el diagnóstico y la clasificación de la Esteatosis hepática es la biopsia hepática, sin embargo, los métodos de diagnóstico por imágenes se hacen necesarios para su estudio, en especial para cierto tipo de casos, como los grupos etarios extremos (niños, ancianos) [3], en donde no se recomiendan los procedimientos invasivos por los riesgos relacionados con este tipo de intervencionismo (por ejemplo comorbilidades).

Las técnicas de diagnóstico por imagen desempeñan un papel cada vez más importante en el diagnóstico de las enfermedades hepáticas difusas puesto que en gran número de casos aclaran la coincidencia que a menudo existe entre las manifestaciones clínicas y analíticas de la mayor parte de los procesos hepáticos difusos. Así se elimina en algunos casos, la necesidad de recurrir a la biopsia. Existen diferentes métodos de imágenes para el diagnóstico de esta patología hepática como son ultrasonido, tomografía computarizada y resonancia magnética, las cuales permiten una adecuada caracterización de la arquitectura del parénquima hepático y de sus cambios estructurales [4].

### **1.1. Motivación y Planteamiento del Problema**

La enfermedad hepática grasa no alcohólica (EHGNA) se define como una entidad clínico-patológica que engloba un amplio espectro de lesiones en el hígado, que se producen en ausencia de un consumo crónico de alcohol, en el que se incluyen la esteatosis simple, la esteatosis asociada a cambios necroinflamatorios con grados variables de fibrosis denominada esteatohepatitis, y finalmente la cirrosis. Actualmente se considera que la EHGNA es el componente hepático del denominado síndrome metabólico, que se puede definir como la agrupación de diferentes factores de riesgo vascular y metabólico como la obesidad visceral, la hiperglucemia secundaria a resistencia a la insulina, la dislipidemia y la hipertensión arterial. Debido a la creciente incidencia del síndrome metabólico en los países desarrollados, la EHGNA está emergiendo como una de las enfermedades hepáticas más frecuentes en nuestro medio. En los últimos años se han producido importantes avances científicos en el conocimiento de los mecanismos implicados en la patogenia de esta enfermedad crónica del hígado, así como en el diagnóstico no invasivo y en el tratamiento farmacológico de la EHGNA [5].

Aunque la prevalencia de la EHGNA en la población general no se conoce con exactitud, fundamentalmente debido a que no se dispone de marcadores diagnósticos

no invasivos que sean sensibles y específicos de esta enfermedad hepática, los datos provenientes de estudios poblacionales estiman su prevalencia entre un 20-30% de la población en los países occidentales y un 15% en los países asiáticos. La información disponible sobre la incidencia de la EHGNA es escasa, aunque en un estudio longitudinal prospectivo realizado en una población italiana se comprobó que la incidencia de hígado graso era de 18,5 casos por 1.000 personas y año. Referente a la esteatohepatitis, los datos sobre la prevalencia/incidencia son muy variables y de difícil interpretación, ya que el diagnóstico de certeza sólo se puede realizar por biopsia hepática, estimándose su prevalencia en un 2-3% de la población general [5].

La Esteatosis Hepática representa un importante problema mundial de salud pública, constituyendo una pandemia; afecta a países ricos y pobres por igual. El trasplante de hígado es un manejo apropiado cuando hay insuficiencia hepática. Entre 30-40% de los pacientes con cirrosis relacionada con Esteatosis Hepática requieren un trasplante de hígado [6].

Afortunadamente, el hígado graso puede ser una enfermedad bastante benigna cuando se diagnostica y trata a tiempo. Actualmente la detección de esta patología se hace a través de la visualización de las imágenes médicas por parte del especialista en diagnóstico por imágenes, por esta razón se propone en este trabajo el desarrollo de una herramienta que realice la detección a través de imágenes médicas de esta enfermedad hepática para así brindar un diagnóstico preciso y oportuno para su posterior tratamiento terapéutico.

## **1.2. Antecedentes**

Qayyum et al [7], en su trabajo propusieron registrar la medición de la esteatosis hepática por medio de imágenes tomográficas, mediante el cálculo de densidad de la escala de Hounsfield para esto tomaron tres regiones de interés ubicadas en el hígado, la primera región en el lóbulo derecho por encima de la vena porta, la segunda región en el lóbulo derecho por debajo de la vena porta y la tercera región en el lóbulo izquierdo. A su vez se tomaron tres regiones circulares de interés en el bazo con el fin

de obtener la densidad del bazo para luego ser comparada con las tomadas en el hígado. La densidad media del hígado y el bazo se calculó a partir de las regiones de interés. Como resultado se restaron la densidad promedio del bazo y la densidad promedio del hígado logrando correlación con respecto a cortes histológicos.

Davidson et al [8], desarrollaron un protocolo para la medición de grasa en el hígado mediante tomografía computarizada (TC), se llevó a cabo un estudio preliminar con 118 hombres y 76 mujeres para determinar un punto de referencia vertebral fácilmente identificables en la que la imagen de TC muestra tanto el hígado como el bazo. En esta muestra, también se evaluó la región característica de atenuación del tejido del hígado y del bazo a nivel de T12-L1, mediante la subdivisión de cada imagen en cuadrantes de anterior a posterior, cada uno de los cuales se dividieron en regiones medias y laterales. La adquisición de una imagen de TC a nivel de T12-L1 es un método práctico y confiable para la medición rutinaria de la grasa en el hígado en la investigación y en el ámbito clínico.

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

Lupsor et al [9], establece que la esteatosis hepática es una enfermedad frecuente en la práctica médica y tiene una gran importancia debido a la posible evolución hacia la cirrosis. La evaluación clínica y de laboratorio tiene el valor predictivo positivo muy reducido. Por otra parte, una serie de técnicas de imágenes se pueden utilizar para el diagnóstico y cuantificación de la esteatosis: resonancia magnética (RM), tomografía computarizada (TC), y no menos importante, la ecografía. Los métodos por imágenes descritos representan una alternativa válida para el diagnóstico no invasivo de la esteatosis hepática con un grado razonable de sensibilidad y especificidad. De estos la RM puede ser el método cuantitativo más seguro, sin embargo es caro y no está fácilmente disponible. La TC permite un diagnóstico significativo para la detección de esteatosis hepática. El estudio por ultrasonido puede ser muy preciso, aunque tiene algunos límites ya que va a depender de las técnicas de procesamiento de imágenes que se le aplique. Sin embargo, ninguno de los métodos

de imagen descritos puede distinguir entre la esteatosis simple y la esteatohepatitis. La biopsia hepática sigue siendo el único método de diagnóstico preciso.

Chun Yeh et al [10], clasificó la esteatosis hepática usando ultrasonido de alta frecuencia, para esto utilizaron características de imágenes derivadas de la concurrencia de niveles de gris, Transformada Wavelet y máquina de soporte vectorial de tipo clasificación. Los resultados de la clasificación se correlacionaron con muestras histológicas indicando que la esteatosis hepática puede ser caracterizada con más precisión usando ultrasonido de alta frecuencia.

Gómez et al [11], en su trabajo investigaron el comportamiento de 22 estadísticas de co-ocurrencia combinadas con 6 niveles de cuantificación de escala de grises para clasificar lesiones de mamas en imágenes de ultrasonido. La base de datos estuvo compuesta por 436 ultrasonidos mamarios conformada por 217 carcinomas y 219 lesiones benignas. Delimitaron la región de interés con un rectángulo y calcularon la matriz de co-ocurrencia con diferentes niveles de cuantificación, ángulos y distancias. Como mejor resultado obtuvieron un ROC de 0.87 con 6 niveles de cuantificación. Los descriptores de textura que contribuyeron notablemente para distinguir lesiones mamarias fue el contraste y la correlación calculados a partir de GLCM con un ángulo de  $90^\circ$  y distancia de más de 5 píxeles.

### **1.3. Objetivos:**

#### **1.3.1. Objetivo General:**

- Detectar la esteatosis hepática por medio de imágenes médicas.

#### **1.3.2. Objetivos Específicos:**

- Aplicar técnicas de procesamiento de imágenes para la detección de esteatosis hepática mediante el uso de imágenes tomográficas e imágenes de ultrasonido.

- Aplicar técnicas de análisis estadístico y máquinas de soporte vectorial para la clasificación de pacientes con esteatosis hepática mediante el uso de imágenes tomográficas e imágenes de ultrasonido.
- Validar clínicamente la herramienta computarizada a desarrollar.

## **1.4. Metodología**

### **1.4.1. Criterios de Inclusión**

Está conformado por pacientes de ambos sexos, divididos en dos grupos. En el primer grupo se incluirán todos aquellos pacientes que posean diagnóstico imagenológico (ultrasonido ó tomografía computarizada) que indiquen la presencia de esteatosis hepática y un segundo grupo integrado por pacientes que no presenten ninguna patología hepática asociada, pacientes que se consideren sanos.

### **1.4.2. Criterios de Exclusión**

Se excluirán todos aquellos pacientes que tengan otra patología hepática asociada.

www.bdigital.ula.ve

## **1.5. Organización del Trabajo**

En el capítulo 2 se describe las bases teóricas referentes a la anatomía y fisiología del hígado seguido por la imagenología médica que abarca la tomografía computarizada, ultrasonido y resonancia magnética. En el capítulo 3, se encontrará las bases teóricas y resultados obtenidos de la técnica de análisis de textura como es la matriz de co-ocurrencias y análisis estadístico. En el capítulo 4 se utilizan las Máquinas de Soporte Vectorial para la clasificación de pacientes con esteatosis hepática y pacientes sanos y se compara con el modelo de regresión logística binaria. Por último en el Capítulo 5 se destacan las conclusiones fundamentales del estudio y las recomendaciones para futuras mejoras de las técnicas de detección propuestas.

Este trabajo dio lugar a las siguientes publicaciones:

Rivas, E., Ortiz, J., Medina, R., Moreno, F. *La Matriz de Co-ocurrencia como Herramienta para la Detección de la Esteatosis Hepática en Imágenes Tomográficas.*

Congreso Internacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas – Cimenics 2012. Margarita, Venezuela.

Rivas, E., Benitez, A., Medina, R., Moreno, F. *Detección de la Esteatosis Hepática en Imágenes Médicas por medio de la Matriz de Co-ocurrencia*. IV Congreso Venezolano de Bioingeniería – BIOVEN 2012. San Cristóbal, Venezuela.

Rivas, E., Moreno, F., Benitez, A., Willie, M., Vanegas, P., Medina, R. *Hepatic Steatosis Detection Using the Co-occurrence Matrix in Tomography and Ultrasound Images*. 20th Symposium on Signal Processing, Images and Computer Vision - STSIVA 2015. Bogotá, Colombia. Publicado en la IEEE 978-1-4673-9461-1/15.

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

## Capítulo II

### Anatomía del Hígado e Imagenología Médica

En este capítulo se describen las bases teóricas de la anatomía del hígado entre estas la morfología, divisiones, segmentación e histología; a su vez se estudia la patología hepática de interés como es la Esteatosis Hepática, diagnóstico y tratamiento. Además se encuentran los fundamentos de la imagenología médica y las técnicas de diagnóstico por imagen las cuales permiten una adecuada caracterización de la arquitectura del parénquima hepático y de sus cambios estructurales; la tomografía computarizada con sus tres fases: simple, arterial y venosa; el ultrasonido y la resonancia magnética.

#### 2.1. Generalidades

##### 2.1.1. Morfología General

La forma general del hígado, clásicamente, se compara con la mitad superior de un ovoide horizontal, de gran extremo derecho, alargado transversalmente debajo del diafragma, por delante del plano posterior. De coloración rojo pardusca, de consistencia firme, está constituido por un parénquima friable rodeado por una delgada cápsula fibrosa, la cápsula de Glisson (túnica fibrosa), prolongación de las vainas fibrosas que rodean a los vasos portales o vainas periportales.

Sus dimensiones en el adulto son de alrededor de 28 cm de longitud por 15 cm en su sentido anteroposterior y 8 cm de espesor a nivel de la parte derecha. Su peso de 1.400 a 1.500 g en el cadáver, es en realidad, más elevado y alcanza entre 2.300 y 2.400 g en el ser vivo, al estar ingurgitado de sangre. Sus dimensiones son proporcionalmente más importantes en el recién nacido y el niño pequeño que en el adulto (*figura 2.1*) [12].

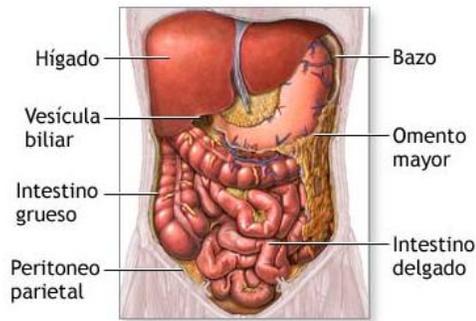


Figura 2.1: *Órganos del Abdomen, tomado de [13].*

### 2.1.2. Situación

El hígado está situado en el piso supramesocólico del abdomen, donde ocupa la casi totalidad del hipocondrio derecho; se amolda sobre la cara inferior de la cúpula diafragmática derecha, se adosa por detrás al plano posterior y a la vena cava inferior y recubre así la región piloroduodenal y el extremo derecho del páncreas. Su extremo izquierdo, más o menos aguzado, rebasa la línea media y cruza la cara anterior del esófago en contacto con la cara inferior del diafragma, llegando a veces a tomar contacto con el bazo [12].

### 2.1.3. Proyección Esquelética

El hígado es un órgano tanto torácico como abdominal. Hacia arriba, en efecto, su proyección describe una línea cóncava hacia abajo que alcanza el 5° espacio intercostal derecho sobre la línea mamilar.

Hacia abajo, el borde anterior del hígado sigue oblicuamente hacia arriba y adentro el reborde costal, al cual normalmente no sobrepasa y debajo del cual sólo es perceptible a la palpación en la inspiración profunda; luego cruza la fosa epigástrica siguiendo una línea que va desde el 9° cartílago costal derecho al 8° cartílago costal izquierdo (*figura 2.2*) [12].

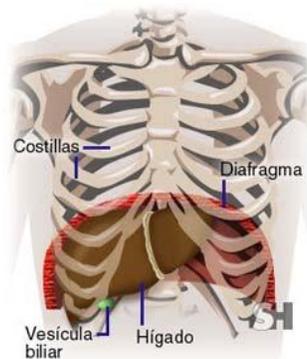


Figura 2.2: *Proyección Esquelética, tomado de [13].*

## 2.2. Caras del Hígado

### 2.2.1 Cara Superior o Diafragmática

Es regularmente lisa y muy convexa en sentido anteroposterior. Ancha en su parte derecha, se aguza progresivamente hacia la izquierda. Está revestida en su totalidad por peritoneo visceral y presenta la unión de sus 2/3 derechos y de su 1/3 izquierdo la inserción del ligamento falciforme o ligamento suspensorio, repliegue peritoneal sagital de 2 hojas que hacen reflexión lateralmente, por una parte sobre la cara superior del hígado y por la otra sobre la cara inferior del diafragma (*figura 2.3*) [12].

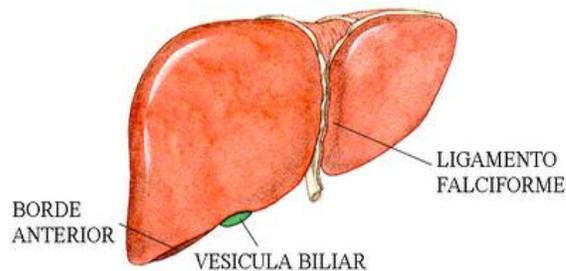


Figura 2.3: *Cara antero superior del hígado, tomado de [14].*

### 2.2.2 Cara Inferior

Es oblicua hacia abajo y adelante. Tiene, en conjunto, una forma trapezoidal de base amplia derecha. Se caracteriza por la presencia en su parte media del hilio

hepático (porta hepatis) donde se dividen los elementos del pedículo infrahepático; vena porta, arteria hepática y vías biliares. Esta cara inferior está marcada, sobre todo, por un surco en forma de “H” constituido por 1 surco trasversal que corresponde al hilio y 2 surcos anteroposteriores, uno derecho y el otro izquierdo (*figura 2.4*) [12].

### 2.2.3 Cara Posterior

Es la más estrecha. De forma triangular con base hacia la derecha y vértice hacia la izquierda, es en conjunto cóncava hacia atrás y se amolda sobre la vena cava y sobre la convexidad de la columna vertebral. Se caracteriza por la presencia de 2 surcos, el surco vertical derecho y el surco vertical izquierdo (*figura 2.4*) [12].

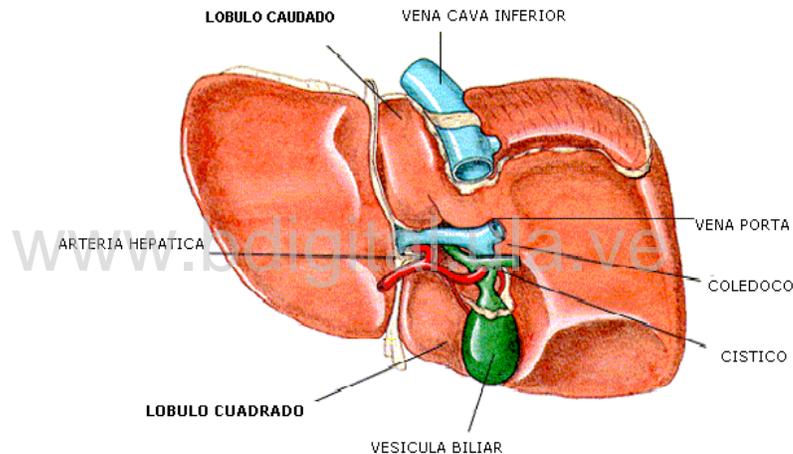


Figura 2.4: *Cara postero inferior del hígado, tomado de [14].*

### 2.3. Divisiones Morfológicas

Es clásico dividir al hígado en 4 lóbulos:

1. El lóbulo izquierdo, situado a la izquierda del ligamento falciforme y del conducto de Arancio.
2. El lóbulo cuadrado, visible solamente en la cara inferior del hígado y limitado por el borde anterior por delante, el surco umbilical a la izquierda, el lecho vesicular a la derecha y el hilio por detrás.

3. El lóbulo caudado o lóbulo de Spigel, situado entre el borde posterior del hilio por delante, la vena cava por detrás, el conducto de Arancio a la izquierda y el surco de la vena cava a la derecha.
4. El lóbulo derecho, situado a la derecha del ligamento falciforme.

En realidad los límites de estos lóbulos son extremadamente imprecisos y, si bien es posible distinguir 4 lóbulos sobre la cara inferior, aparecen solamente 2 sobre la convexidad del hígado. Por otra parte, estos lóbulos no corresponden absolutamente a una distribución real de los vasos intrahepáticos y no reflejan una realidad funcional. Veremos que actualmente, es preferible sustituir esta división morfológica por una segmentación basada en la distribución vascular intraparenquimatosa de los diferentes pedículos del hígado [12].

#### **2.4. Segmentación Hepática**

La segmentación hepática portal se basa en la distribución del pedículo hepático en el interior del órgano, envuelto por su cápsula fibrosa perivascular. La distribución de la vena porta hepática es, obviamente, el elemento rector. Las ramas derecha e izquierda de la vena porta hepática irrigan territorios que se denominan porción hepática derecha (hígado derecho) y porción hepática izquierda (hígado izquierdo). Están separadas por un plano denominado fisura portal principal. En la cara visceral del hígado, esta fisura sigue el fondo de la fosa de la vesícula biliar, corta el porta hepático por su parte media y termina en el borde izquierdo de la vena cava inferior. Así pues, puede hablarse de pedículo hepático derecho e izquierdo.

La rama derecha de la vena porta hepática da origen a dos ramas principales que permiten distinguir dos divisiones derechas: una medial, que está próxima a la fisura portal principal y la otra lateral que es periférica. Su límite está marcado por la fisura portal derecha. Cada una de estas divisiones a su vez se subdivide en dos segmentos uno anterior y otro posterior.

La rama izquierda de la vena porta hepática se divide, en el extremo izquierdo del hilio, en dos ramas: la rama más pequeña irriga la división lateral izquierda; la rama más gruesa, que es umbilical, irriga la división medial izquierda. Las dos divisiones están separadas por la fisura umbilical. El segmento medial de la división medial izquierda corresponde al lóbulo cuadrado, que se extiende hasta la cara diafragmática del órgano. Posteriormente al porta hepático se encuentra el lóbulo caudado, que forma el segmento posterior [15].

Así, la distribución del pedículo hepático comprende cinco divisiones: una división corresponde al lóbulo caudado, que está a caballo sobre las dos porciones hepáticas derecha e izquierda; las otras cuatro, dos mediales y dos laterales, forman las porciones hepáticas derecha e izquierda. Estas divisiones se subdividen en segmentos, que están numerados, en la cara visceral del hígado y alrededor del porta hepático, en sentido inverso al de las agujas del reloj; el octavo segmento sólo aparece en la cara diafragmática. La circulación venosa hepática de retorno presenta una segmentación idéntica a la disposición portal. Las venas del lóbulo caudado drenan el segmento posterior, que se corresponde con bastante exactitud con el sector portal homólogo. Estas venas drenan directamente en la vena cava inferior (Couinaud). Dejando esto aparte, los tres sectores hepáticos (izquierdo, medio y derecho) corresponden a las venas hepáticas izquierda, intermedia y derecha.

La vena hepática izquierda está situada en el plano de la fisura umbilical, que separa las divisiones lateral y medial izquierdas. Drena el sector hepático izquierdo, que corresponde al lóbulo izquierdo clásico. La vena hepática intermedia está situada en el plano de la fisura portal principal, entre las porciones hepáticas derecha e izquierda. Asegura la circulación de retorno del sector hepático medio.

La vena hepática derecha se halla en el plano de la fisura portal derecha; drena el sector hepático derecho. Los troncos principales de la circulación aferente, que son superiores al pedículo hepático, también están situados en las fisuras hepáticas. El

pedículo medial izquierdo se encuentra en la fisura umbilical y el pedículo medial derecho está situado en la fisura portal derecha (Couinaud) (figura 2.5) [15].

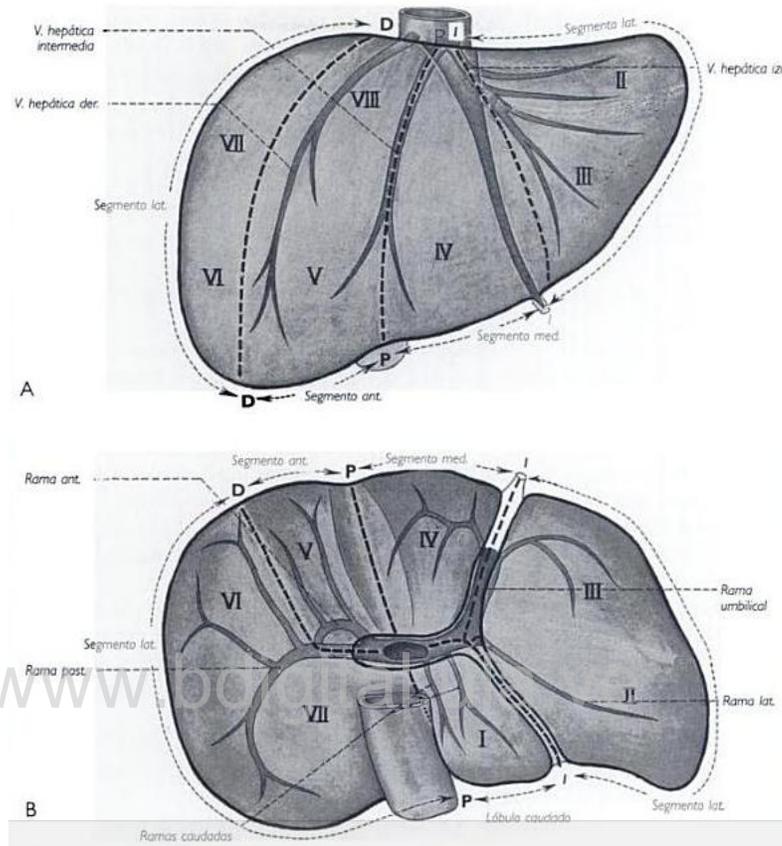


Figura 2.5: Esquemas de la segmentación hepática (Couinaud), tomado de [15].

## 2.5. Histología Hepática

Los componentes estructurales del hígado son las láminas de células hepáticas llamadas hepatocitos que están separadas por amplios canales vasculares conocidos como sinusoides. La sangre llega a los sinusoides desde las ramas terminales de la vena porta y la arteria hepática, aportando respectivamente, sangre rica en elementos nutritivos procedente del aparato gastrointestinal y sangre rica en oxígeno. Las ramas mayores de estos dos vasos siguen trayectos lado a lado en los tabiques fibrosos conocidos como espacio porta, donde también se encuentran los conductos biliares que llevan la bilis en dirección opuesta, desde el hígado al duodeno.

La sangre de la arteria hepática y de la vena porta pasa por los sinusoides, donde establece un íntimo contacto con los hepatocitos con el fin de intercambiar los elementos nutritivos y los productos metabólicos. Después, alcanza las ramas de la vena hepática desde donde llega a la vena cava inferior [16].

## **2.6. Esteatosis Hepática**

La esteatosis hepática es la acumulación de lípidos en el citoplasma de los hepatocitos. El hallazgo de hígado graso es muy frecuente, y de acuerdo con su etiología se clasifica en esteatosis alcohólicas y esteatosis no alcohólicas (NASH, non alcoholic steato hepatitis), relacionando estas últimas fundamentalmente con la diabetes de tipo 2, dislipemias (hipercolesterolemia, hipertrigliceridemia) y con la obesidad.

Desde el punto de vista histológico, se distingue la esteatosis en macrovesicular y microvesicular, si bien con frecuencia se combinan ambas morfologías. La esteatosis macrovesicular se caracteriza por una gran vacuola citoplasmática que desplaza al núcleo, mientras que la microvesicular se describe por la presencia de microvacuolas grasas en el citoplasma del hepatocito, generalmente de menor tamaño que el núcleo y sin desplazarlo.

Cuantitativamente, según el porcentaje de hepatocitos afectados, se clasifica como leve si es menor del 30%, moderada entre el 30% y 60% y grave si están afectados más del 60% de los hepatocitos [17].

Desde el punto de vista clínico, algunos pacientes con esteatosis hepática no experimentan síntomas, mientras que otros pueden manifestar dolor en el cuadrante superior derecho, cansancio y malestar general. La hepatomegalia es un hallazgo frecuente y pueden aumentar los valores de las pruebas de funcionamiento hepático, en especial las aminotransferasas. La esteatohepatitis alcohólica y no alcohólica puede progresar hacia la cirrosis [17].

### 2.6.1. Esteatosis Macrovesicular y Microvesicular

La esteatosis macrovesicular es la forma más frecuente y se debe a una combinación de aumento del aporte, oxigenación inadecuada y disminución de la secreción de lípidos en el hígado. La obesidad, el alcoholismo, la resistencia a la insulina y la diabetes, la caquexia, los fármacos (como los corticoides y el tamoxifeno), los trastornos metabólicos y la hepatitis C producen esteatosis macrovesicular.

La esteatosis microvesicular secundaria a la oxidación defectuosa de los ácidos grasos libres se suele asociar con disfunción hepática grave. El hígado graso agudo del embarazo, el síndrome de Reye, los fármacos, el alcoholismo crónico, los trastornos del ciclo de la urea y las citopatías mitocondriales son algunas de las causas de la esteatosis microvesicular (*figura 2.6*) [17].

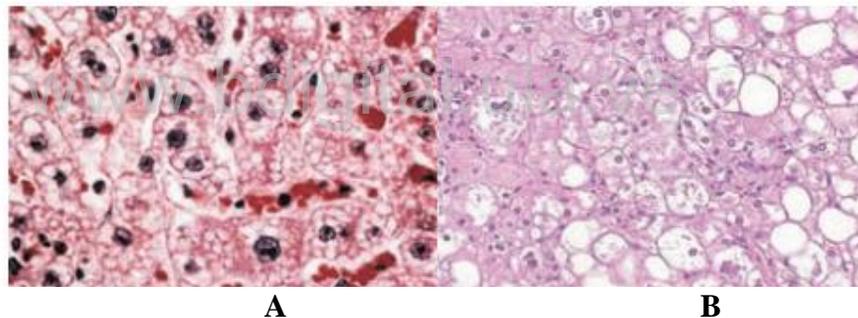


Figura 2.6: *Imágenes histológicas representativas de esteatosis hepática (A) Microvesicular y (B) Macrovesicular, tomado de [18].*

### 2.6.2. Diagnóstico de la Esteatosis Hepática

El diagnóstico de NASH requiere una anamnesis detallada del consumo de alcohol y de medicamentos potencialmente hepatotóxicos (incluyendo hierbas medicinales y medicamentos comprados sin receta médica). Asimismo, deben excluirse otros factores conocidos como causas de la Esteatohepatitis, tales como derivación yeyunoileal, la nutrición parenteral total y ciertas alteraciones genéticas o metabólicas. La desnutrición grave puede provocar también esteatohepatitis no alcohólica, que se manifiesta como atrofia marrón del hígado. En los lactantes el

kwashiorkor (desnutrición proteínico-energética) puede estar asociado con la esteatosis hepática [19].

La esteatosis hepática y la esteatohepatitis se descubren muchas veces durante un chequeo médico de rutina. La mayoría de los pacientes no presentan síntomas, aunque algunos refieren cansancio, malestar general o sensación de plenitud o dolor a la palpación en el cuadrante superior derecho del abdomen. La exploración física puede revelar una constitución androide, con obesidad central o en el tronco y aumento del cociente entre la cintura y las caderas. Los niños son generalmente obesos y pueden tener acantosis pigmentaria, que es un indicador de resistencia a la insulina que puede estar presente en un tercio de los niños y adolescentes. A veces, el paciente presenta síntomas indicativos de hepatopatía en fase avanzada, tales como ascitis, varices hemorrágicas y encefalopatía hepática.

El análisis de sangre revela muchas veces una elevación de leve a moderada de los niveles séricos de aminotransferasa aspartato (ASAT) y aminotransferasa alanina (ALAT). El cociente ASAT/ALAT que se obtiene puede ser indicativo de cirrosis, pero este cociente suele ser más bajo en los pacientes con cirrosis que en los pacientes con hepatopatía alcohólica. Los niveles séricos de glucosa y triglicéridos en ayunas pueden estar elevados. Además, en hasta un 50% de los pacientes puede observarse una elevación de los niveles séricos de ferritina (una proteína de la fase aguda). En aproximadamente un 10% de los casos se observa aumento de la saturación sérica de transferrina-hierro, si bien el almacenamiento de hierro hepático normalmente no aumenta, a no ser que exista una mutación en el gen HFE. La ecografía puede mostrar una situación denominada hígado brillante. La TC generalmente revela la presencia de un hígado más oscuro que el bazo. La RM puede mostrar un aumento de la atenuación en las imágenes de señalización sopesadas en T1, lo cual es indicativo de la presencia de grasa [19].

### **2.6.3. Tratamiento de la Esteatosis Hepática**

El tratamiento consiste en la pérdida de peso mediante la dieta y ejercicio físico. Una disminución de sólo el 10% en el peso corporal puede producir una mejoría considerable en la resistencia a la insulina. Algunos autores aconsejan reducir el contenido de hidratos de carbono en la dieta. Se ha investigado con muchos fármacos, y algunos de ellos parecen ser prometedores, aunque la Food and Drug Administration (FDA) de Estados Unidos no ha autorizado hasta la fecha ningún medicamento para el tratamiento del NASH. Algunos de los fármacos actualmente en investigación para el tratamiento de esta enfermedad son el ácido ursodeoxicólico, la metformina, la vitamina E y las tiazolinedionas de nueva generación [19].

## **2.7. IMAGENOLOGÍA MÉDICA**

### **2.7.1. Rayos X**

El descubrimiento de los rayos X fue realizado por Wilhelm Conrad Röntgen, en diciembre de 1895, en el Instituto de Física de la Universidad de Würzburg. Este científico, al trabajar con tubos al vacío que emitían un nuevo tipo de rayos llamados catódicos, notó que esa irradiación atravesaba diversos materiales e incluso el cuerpo humano y la llamó rayos X, porque desconocía su naturaleza. Además, observó que su imagen se reflejaba en una placa fotográfica. Hoy sabemos que los rayos X son ondas electromagnéticas que se forman siempre que chocan electrones dotados de alta energía contra un metal, al ser acelerados por un gran voltaje. El descubrimiento de estos rayos X dio lugar al nacimiento de una nueva especialidad médica, la radiología. Las radiografías se basan en que los rayos X atraviesan los tejidos del cuerpo humano e impresionan la placa fotográfica, pero los huesos y los objetos metálicos no los dejan pasar, con lo que su sombra queda proyectada. Con los años se crearon diferentes técnicas y sistemas de radiología como las tomografías computarizadas y la resonancia magnética. Röntgen fue el primer investigador en recibir el premio Nobel de Física [20].

### 2.7.2. Tomografía Computarizada

La tomografía computarizada es una técnica de diagnóstico por imagen que permite la visualización de cortes del organismo a partir de múltiples determinaciones de absorción de rayos X. La formación de la imagen se limita a cortes transversales (axiales) perpendiculares al eje longitudinal del cuerpo, aunque la reconstrucción final de la misma puede obtenerse en cualquier plano. La denominación actual de la técnica en castellano, tomografía computarizada (TC), sustituye a la denominación tomografía axial computarizada (TAC) empleada anteriormente. La abreviatura TC se corresponde con la utilizada en el mundo anglosajón CT (computed tomography) (*figura 2.7*) [21].

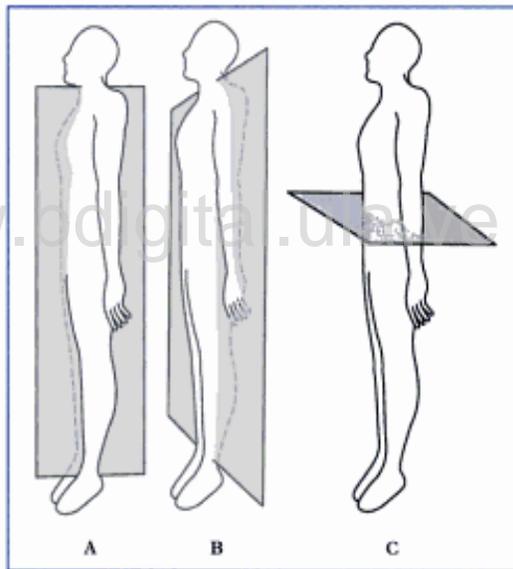


Figura 2.7: *Cortes sagital (A), coronal (B) y axial (C), tomado de [21].*

#### 2.7.2.1. Recuerdo Histórico

La TC ha sido la contribución más valiosa al campo del diagnóstico por imagen desde el descubrimiento de los rayos X por Roentgen. En los fundamentos de la técnica trabajaron independientemente el físico norteamericano A. M. Cormack y el ingeniero inglés G. M. Hounsfield. Este último dirigía la sección médica del laboratorio central de investigación de la compañía Electrical Musical Instruments

(EMI), una empresa dedicada fundamentalmente al campo de la producción discográfica. Hounsfield planteó los principios básicos y el diseño de la primera unidad. Las experiencias clínicas comenzaron en 1972, aunque la primera imagen craneal se obtuvo un año antes. La primera TC de cuerpo entero se consiguió en 1974. En 1979, Hounsfield y Cormack recibieron el Premio Nobel de Medicina [21].

### 2.7.2.2. Vóxel, Píxel y Matriz

La imagen de TC es un conjunto bidimensional de valores cuantificados en una escala de grises (píxeles). Estos valores representan el coeficiente de atenuación de cada elemento de volumen o vóxel. Así, el término vóxel aporta idea de volumen y corresponde a un área pequeña de la imagen por el espesor de corte. La longitud o altura del vóxel (1-10 mm) está determinada por la anchura del haz de rayos X (corte tomográfico). Un píxel es la representación gráfica en una matriz plana de la información obtenida en un vóxel. El término píxel ofrece, por tanto, una idea de superficie. La matriz es la representación de todos los datos obtenidos en la realización del corte, es decir, una serie de píxeles dispuestos en filas y columnas. Las modernas unidades de TC permiten utilizar matrices de 256 x 256, 512 x 512 y 1.024 x 1.024 (figuras 2.8 y 2.9) [21].

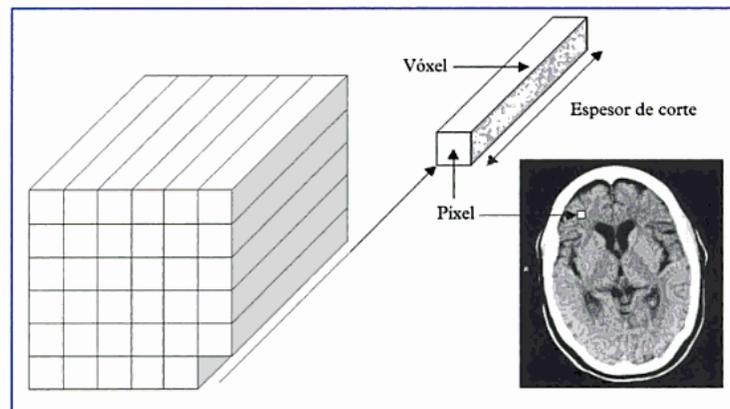


Figura 2.8: Un vóxel corresponde al área de un píxel por el espesor de corte, tomado de [21].

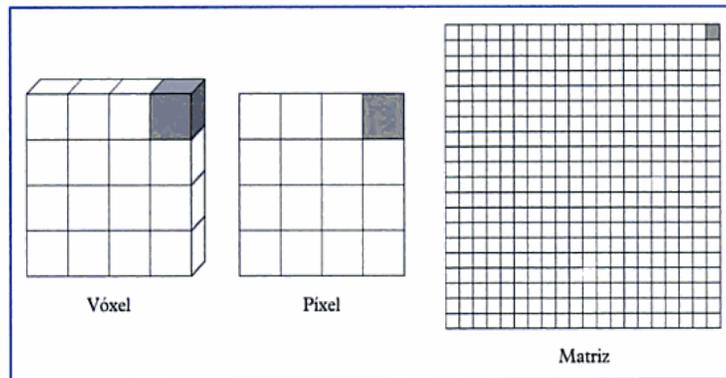


Figura 2.9: *Vóxel, píxel y matriz, tomado de [21].*

El tamaño del píxel está relacionado con el campo de visión y el tamaño de la matriz de imagen. El campo de visión corresponde al diámetro de la imagen reconstruida. Cuando aumenta para una dimensión matricial fija, el tamaño de píxel se incrementa de forma directamente proporcional. Sin embargo, al aumentar el tamaño de la matriz, se reduce el del píxel según la expresión:

$$\text{Tamaño del píxel} = \frac{\text{Campo de visión}}{\text{Matriz de la imagen}} \quad (2.1)$$

### 2.7.2.3. Fundamentos de la Tomografía Computarizada

La TC se basa en la ecuación que nos indica la atenuación que sufre el haz de radiación al atravesar un absorbente determinado.

$$I_x = I_0 \cdot e^{-\mu x} \quad (2.2)$$

Dónde  $I_x$  es el valor de la intensidad de la radiación X después de atravesar el espesor  $x$  de un objeto, e  $I_0$  es el valor de la intensidad de la radiación que incide sobre la superficie del mismo, siendo  $e$  el número base del sistema logarítmico natural. El coeficiente lineal de atenuación  $\mu$  depende del número atómico, de la densidad del medio y de la longitud de la onda (espectro de energía) del haz de radiación incidente. Aunque la atenuación total sea similar en los dos casos, el perfil de atenuación a través del espesor atravesado es diferente. Al no haber diferencias en la transmisión del haz de radiación que atraviesa cada una de ellas, ambas zonas

mostrarían el mismo ennegrecimiento en una radiografía. Es decir, la radiografía convencional sólo ofrece el valor medio de  $\mu$  del trayecto recorrido por el haz de rayos X. Sin embargo, la TC proporciona la atenuación promedio de cada uno de los vóxeles que componen el corte corporal, es decir, proporciona los valores  $\mu_1 \mu_2 \mu_3 \mu_4$  (figura 2.10). Ahora bien, con un solo haz no es posible obtener esta información, sino que es preciso disponer de diversos haces que incidan sobre cada vóxel desde distintos ángulos (proyecciones). Es decir, sólo es posible la reconstrucción de la estructura de un objeto a partir de múltiples proyecciones del mismo [21].

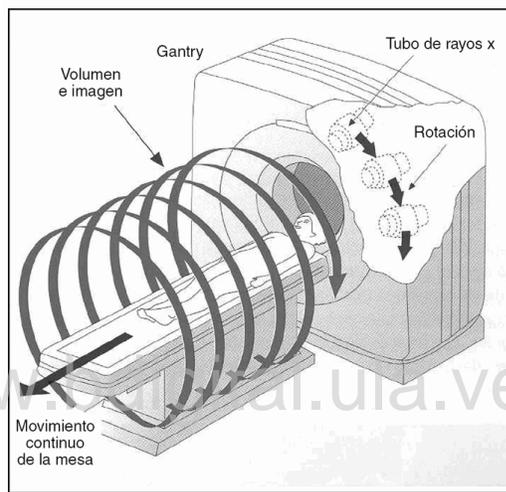


Figura 2.10: *Fundamentos de un Equipo de Tomografía Computarizada, tomado de[22].*

#### 2.7.2.4. Densidad y Escala de Grises

Los valores de densidad o atenuación correspondientes a cada uno de los píxeles se normalizaron respecto a un material de referencia, el agua, y se representaron dentro de una escala de grises en la que el blanco correspondía al valor de máxima densidad y el negro al valor de mínima densidad. En realidad, son valores de píxel que representan el coeficiente de atenuación lineal de un elemento de volumen o vóxel. En esta escala se estableció como unidad de absorción el Hounsfield, en honor al ingeniero que planteó los fundamentos de la técnica. Así, las unidades Hounsfield (UH) son 0 para el agua, -1000 para el aire y +1000 para la absorción del hueso compacto. Entre estos valores se encuentran los de los distintos

tejidos (grasa, sangre, músculo, etc). El valor en UH que corresponde a un tejido concreto será:

$$UH = \frac{\mu_{tejido} - \mu_{agua}}{\mu_{agua}} * 1.000 \quad (2.3)$$

Dónde  $\mu_{tejido}$  equivale al coeficiente de atenuación del píxel (expresión bidimensional de un vóxel del tejido); y  $\mu_{agua}$ , al coeficiente de atenuación del agua. La figura 5 ilustra la escala de valores TC con las densidades principales del cuerpo humano, desde el límite inferior correspondiente a la densidad del aire (negro) hasta el límite superior que corresponde al hueso compacto (blanco) (figura 2.11) [21].

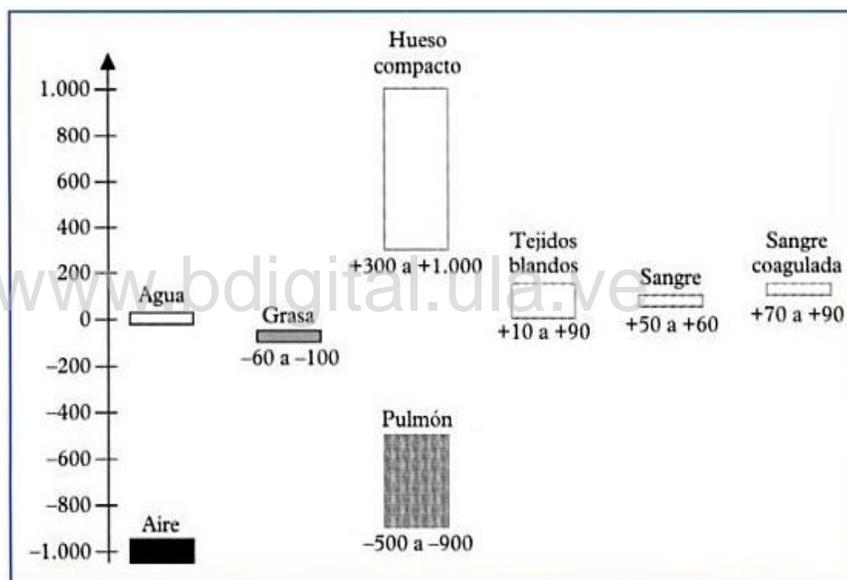


Figura 2.11: Escalas de números TC del cuerpo humano, tomado de [21].

### 2.7.2.5. Selección de Ventana y Nivel de Ventana

La gama de valores anterior supone 2.000 tonos de grises (entre -1000 y +1000 UH). Sin embargo, como el ojo humano no es capaz de distinguir más de 30 tonalidades, representar en una imagen toda la gama de valores conlleva no poder visualizar una gran cantidad de información. Por lo tanto, sólo se representa mediante escala de grises un sector parcial de los valores TC previamente seleccionado por el radiólogo (selección de ventana). La utilización de ventanas permite en definitiva

extraer la información que posee el ordenador mostrando sólo una parte de la misma, aquella que interesa en cada momento (figura 2.12) [21].

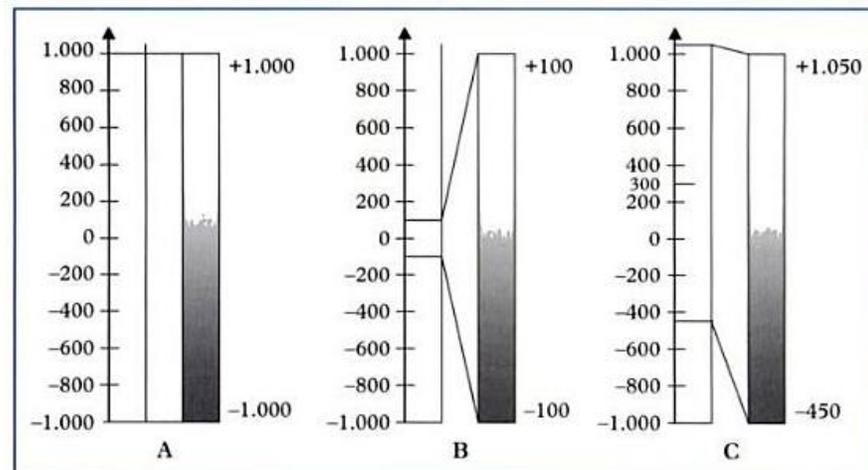


Figura 2.12: Selección de ventana y nivel de ventana, tomado de [21].

### 2.7.2.6. Unidades de Tomografía Computarizada

La figura 2.13 muestra una unidad de TC de tercera generación. El gantry o cuerpo vertical de la unidad presenta un orificio central en el que se introduce la mesa de exploración con el paciente. En su interior se encuentra el sistema de adquisición de datos compuesto básicamente por el tubo de rayos X, el sistema de colimación, el generador de rayos X, los detectores y una serie de dispositivos electrónicos asociados. En los primeros equipos, el generador no estaba en el gantry y los cables a través de los cuales suministraba la potencia al tubo de rayos X se desenrollaban durante la rotación del conjunto tubo-detectores para obtener un corte. Para conseguir el corte siguiente, la rotación se producía en la dirección contraria y los cables se volvían a enrollar. La introducción de la tecnología de anillos deslizantes evitó que la transferencia de electricidad y datos tuviera que hacerse a través de grandes cables y permitió la rotación continua del tubo, del generador y del sistema de detección que hizo posible el desarrollo de la TC helicoidal.



Figura 2.13: *Unidad de TC de tercera generación, tomado de [13].*

El tubo de rayos X emite un haz de radiación de alta energía (se utilizan tensiones que oscilan entre 80 y 140kV), muy homogéneo, que se hace aún más monoenergético mediante la colocación a la salida de filtros de aluminio y cobre. La mayoría de los tubos empleados en la actualidad generan los rayos X a ráfagas (radiación pulsada), con pulsos de 2 o 3 ms de duración, produciendo irradiaciones de menos de un segundo por corte. Los detectores transforman la energía de la radiación X que incide sobre ellos en una corriente eléctrica que es cuantificada por un sistema electrónico. Los tipos de detectores utilizados van desde los de centello de las primeras generaciones a las cámaras de ionización de gas xenón de unidades de tercera generación o los detectores de estado sólido (centelladores de policristilamina acoplados a fotodiodos) de los TC convencionales más avanzados y de las unidades de TC helicoidal [21].

El haz de radiación es colimado a la salida del tubo y antes de entrar en los detectores. Este último sistema de colimación es el responsable de regular el espesor del corte tomográfico. La mayoría de los equipos obtienen secciones de 0.5 a 10mm. Un computador de gran potencia realiza los cálculos necesarios para el tratamiento de datos y la reconstrucción de la imagen.

### 2.7.2.7. Técnicas de Exploración con Tomografía Computarizada

#### ▪ TC sin contraste “Fase Simple”

La TC sin contraste puede ser utilizada como complemento antes de la TC realizada con contraste o como única técnica de TC. En general, se recurre a esta última opción cuando el paciente es alérgico a los contrastes yodados, si se encuentra en insuficiencia renal o si no es el hígado el órgano sospechoso que va a ser estudiado, como sucede en la búsqueda de abscesos postoperatorios. La esteatosis hepática causa una disminución de la atenuación del parénquima hepático en la TC, lo que se aprecia mejor utilizando una técnica sin contraste, en la que la atenuación del hígado debe ser unas 8 UH (unidades Hounsfield) mayor que la del bazo (*figura 2.14*) [23].

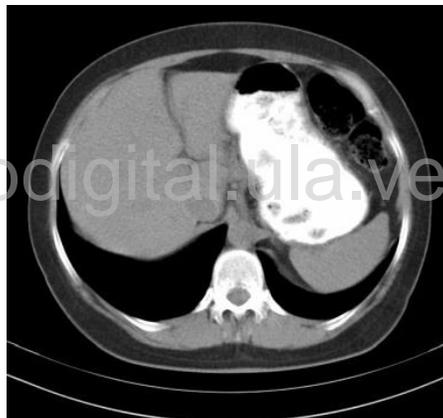


Figura 2.14: *Fase Simple de la TC, tomado de [24].*

#### ▪ TC dinámica realizada con contraste “Fase Arterial ó Vascular”

Con el fin de potenciar al máximo el contraste entre el parénquima hepático y la lesión, la inyección intravenosa de contraste yodado se ha convertido en la técnica habitual para valorar la posible afectación neoplásica del hígado. En numerosos estudios se ha demostrado que la sensibilidad de las técnicas realizadas con contraste supera en alrededor del 15 % a la de las técnicas sin contraste en la valoración de la totalidad de los pacientes enviados para estudio. Si se administran rápidamente grandes volúmenes de contraste (en general, 150 ml de agente yodado al 60 %) y se

hace inmediatamente la exploración, puede obtenerse un realce del hígado. Ello se acentúa por el hecho de que la vena porta transporta el 75 al 80 % del flujo sanguíneo hepático, lo que proporciona el realce dominante del hígado. La esteatosis hepática puede representarse en imágenes realizadas con contraste en las que el parénquima hepático aparece con una atenuación menor que la del bazo, este tipo de imágenes contrastadas son menos fiables y menos específicas. Normalmente, el bazo puede realzarse en mayor grado que el hígado, sobretodo mediante las inyecciones de contraste en bolo, lo que puede hacer pensar, erróneamente, en una esteatosis. Tras la administración del contraste, puede sospecharse una esteatosis significativa cuando la atenuación del parénquima hepático es inferior a la del músculo (*figura 2.15*) [23].

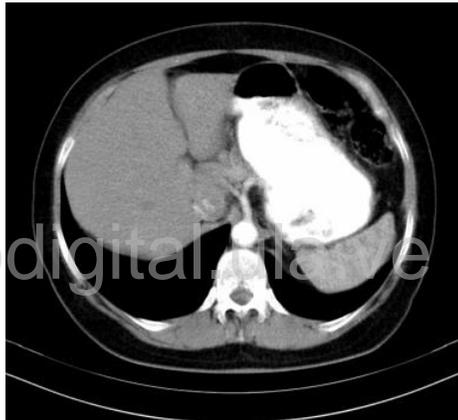


Figura 2.15: *Fase Arterial ó Vascular de la TC, tomado de [24].*

▪ **TC tardía con contraste a altas dosis “Fase Venosa ó Tardía”**

Esta técnica se basa en la capacidad del hígado para excretar un pequeño porcentaje de contraste yodado hacia la vía biliar de forma tardía. Cuando se administra una dosis alta de contraste (generalmente, 60mg de yodo) y se espera durante 4 a 6 horas, el nivel de contraste en la sangre circulante es muy bajo; sin embargo la retención del yodo en el interior de los hepatocitos aumenta la atenuación del parénquima hepático en alrededor de 20 UH (*figura 2.16*) [23].

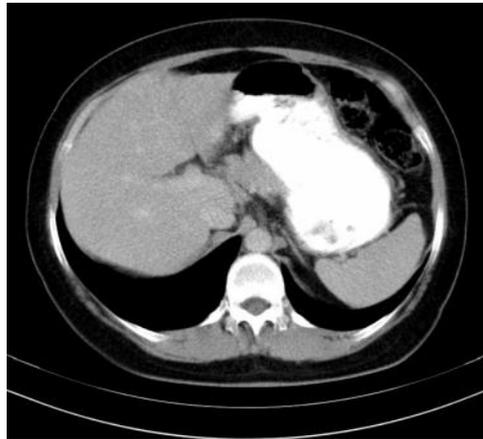


Figura 2.16: *Fase Venosa ó Tardía de la TC, tomado de [24].*

### 2.7.3. Ecografía Hepática

La infiltración grasa o esteatosis hepática es una patología relativamente frecuente, en la que el hígado muestra un mayor incremento de sus depósitos en lípidos. Suele asociarse con múltiples patologías como son la diabetes, etilismo, obesidad, alimentación parenteral, tratamiento esteroideo, malnutrición, hiperlipemia, síndrome de Reye, entre otros. Por su frecuencia hay que destacar en nuestro ambiente el alcoholismo crónico. En la infancia, la presencia de un hígado graso detectado por ecografía, si el niño no está malnutrido o recibiendo nutrición parenteral o tratamiento radioterápico, debe hacer sospechar la existencia de una enfermedad metabólica [25].

Ecográficamente se detecta un hígado heperecogénico en comparación con la corteza del riñon derecho, existiendo atenuación del sonido en las regiones más posteriores. Es lo que se ha denominado clásicamente como “hígado brillante” (*figuras 2.17 y 2.18*). Una descripción más apropiada que la de hígado brillante, a juicio de algunos expertos podría ser “highly reflective liver”, ya que en estos casos, además de unos ecos de mayor amplitud, existe un número incrementado de ellos, apareciendo en la ecografía un patrón de múltiples ecos finos estrechamente empaquetados.



Figura 2.17: *Hígado muy hiperecogénico en comparación con la corteza renal (R) en un paciente con EH, tomado de [25].*

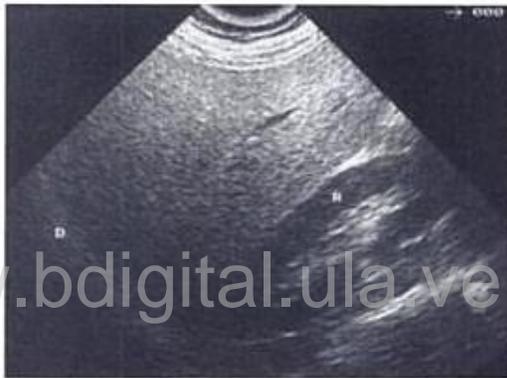


Figura 2.18: *Hígado hiperecogénico. Cuando es muy intensa la infiltración grasa se pierde la definición de los vasos portales y del diafragma (D), tomado de [25].*

Ecográficamente la infiltración grasa puede dividirse en 3 grados: grado 1 o leve, ligero incremento difuso de ecos en el parénquima hepático, con visualización normal del diafragma y de los vasos intrahepáticos; grado 2 o moderado, incremento moderado de los ecos del parénquima, con discreto empeoramiento de la visualización del diafragma y los vasos intrahepáticos; y grado 3 o severo, marcado incremento en la ecogenicidad con pobre visualización o desaparición de los vasos intrahepáticos, diafragma y porción posterior del lóbulo derecho. Aunque un hígado hiperecogénico no es específico, pudiendo aparecer en otras patologías como hepatitis

agudas y crónicas, fibrosis, entre otras, sí es muy sugerente de esteatosis en el contexto clínico adecuado.

En algunos estudios se ha propuesto como dato ecográfico de la existencia de cirrosis hepática, la presencia de un hígado brillante. Esto es debido posiblemente a que, como es sabido, la infiltración grasa o esteatosis es frecuente que coexista junto a fibrosis en un mismo paciente, principalmente en los casos en que el agente etiológico puede provocar ambas lesiones, como ocurre con el alcohol, y por tanto en los distintos trabajos realizados ha podido existir un solapamiento de dichas patologías. En los últimos años, los diversos estudios realizados han demostrado que la atenuación del sonido sólo se produce en los casos en que existe infiltración grasa, y que la fibrosis por sí sola no conduce a tal atenuación, y que por tanto, los pacientes con cirrosis hepática, sin infiltración grasa, tienen una atenuación y ecogenicidad similares a los controles sanos.

En ocasiones, la infiltración grasa hepática no es uniforme adoptando una distribución parcelar. En estos casos, se pueden originar problemas diagnósticos, dado el aspecto de lesión focal en el estudio ecográfico [25].

#### **2.7.4. Resonancia Magnética**

La resonancia magnética (RM) permite evaluar el hígado en forma meticulosa con el fin de examinar su parénquima, el sistema biliar y su vascularización. Si bien la tomografía computarizada y la ecografía del hígado suelen ser los primeros estudios empleados para evaluar este órgano, el empleo de la RM aumenta en forma progresiva como principal modalidad de diagnóstico por la imagen, uso que se agrega a la aplicación para resolver problemas específicos. La RM ofrece una descripción del tejido blando que no se puede obtener con otros métodos. La falta de radiación ionizante y la relativa ausencia de dependencia del operador son otras ventajas de la RM sobre la tomografía computarizada y la ecografía, respectivamente. Las secuencias de pulsos rápidas con respiración contenida sustituyeron en gran medida a

las secuencias de pulsos más lentas del pasado, lo que permitió reducir la duración de los estudios [26].

#### **2.7.4.1 Resonancia Magnética del Hígado**

En la resonancia magnética la esteatosis hepática puede ser difusa, focal o multifocal. Se describieron áreas cuneiformes, geográficas y nodulares en la esteatosis focal. Las imágenes de la esteatosis nodular se pueden asemejar a las de la enfermedad metastásica. La esteatosis focal o la ausencia focal de grasa tienden a desarrollarse en sitios específicos del hígado, como el borde anteromedial del segmento medial, alrededor de la vesícula biliar y dentro de las porciones subcapsulares del hígado. El drenaje anómalo de la vena gástrica hacia el segmento medial posterior que se encuentra delante del hilio hepático, la anomalía de la arteria torácica interna que irriga el área que rodea el ligamento falciforme y el drenaje de las venas vesiculares hacia el hígado circundante determinan una disminución del flujo sanguíneo en la vena porta. La ausencia focal de grasa se puede deber a la reducción del transporte de lípidos secundaria a la disminución del flujo en la vena porta. En ocasiones, estas mismas áreas revelan esteatosis focal. La ausencia focal de grasa también se puede asociar con la presencia de cortocircuitos arterioportales, oclusión de la vena porta y compresión del parénquima hepático por lesiones metastásicas. En función de su localización, las metástasis u otras masas pueden determinar la ausencia de grasa periférica, segmentaria o lobular en un hígado con esteatosis (*figura 2.19*) [26].

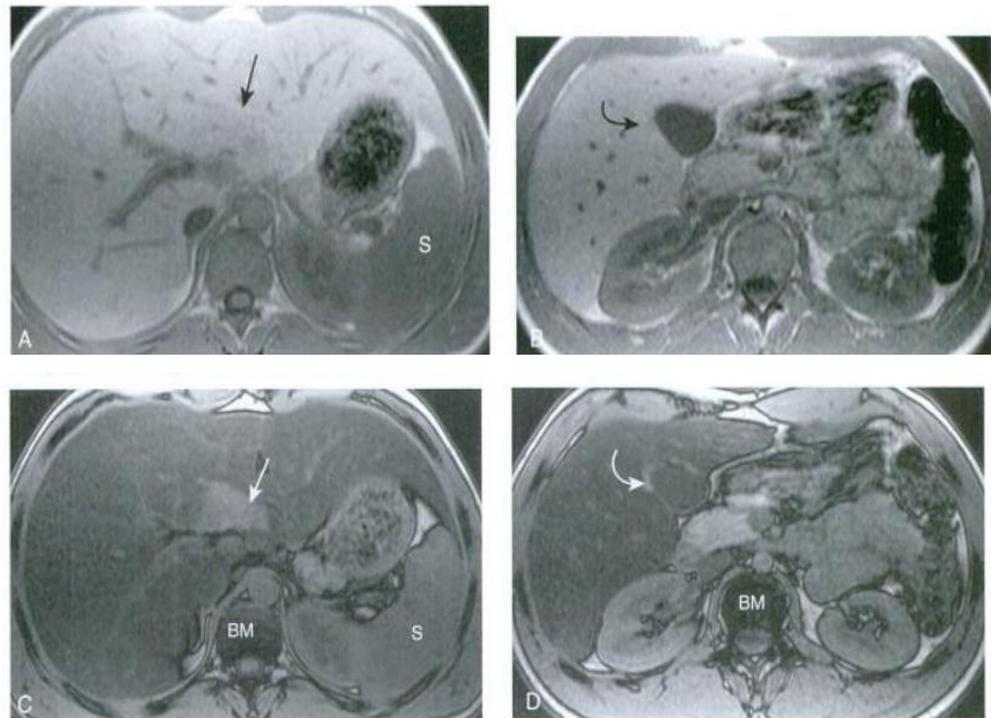


Figura 2.19: *Esteatosis hepática con conservación focal de la grasa en el segmento medial del lóbulo izquierdo y alrededor de la vesícula biliar, tomado de [26].*

## 2.8. Conclusiones

El hígado es el órgano más grande del organismo y está encargado de la mayor cantidad de procesos metabólicos. Está ubicado en el hipocondrio derecho debajo del hemidiafragma derecho, en la región supramesocólica; está constituido por dos (lóbulos derecho e izquierdo) separados entre sí por el ligamento falciforme. Su unidad funcional es la tríada portal, constituida por un hepatocito rodeado del conductillo biliar, una arteriola hepática y una vénula portal. Los desechos de su metabolismo son segregados hacia la segunda porción duodenal a través de los conductos biliares.

La Esteatosis hepática se debe a la acumulación citoplasmática de moléculas de ácidos grasos y colesterol en los hepatocitos. Su diagnóstico presuntivo es principalmente realizado a través de estudios de imágenes (ultrasonido hepático,

tomografía computarizada abdominal, resonancia magnética), realizándose sólo su confirmación a través del estudio histopatológico posterior a la toma de muestra de biopsia hepática (principalmente percutánea).

De acuerdo a los estudios de imágenes empleados para su diagnóstico el que ofrece más ventajas para establecerlo con mayor precisión es la resonancia magnética, sin embargo, la relación costo-beneficio no sustenta la posibilidad de su práctica rutinaria y la presencia de algunas contraindicaciones para su realización sólo la permiten realizar en casos muy seleccionados por el clínico tratante, permitiendo mayor indicación a otros métodos con ventajas similares en cuanto a su diagnóstico y sobre la factibilidad de su realización (por ejemplo: ultrasonido). Adicionalmente, cabe destacar la necesidad de la administración de contrastes en la TC, los cuales pueden no ser tolerados por algunos pacientes iodo-sensibles, imposibilitando su prescripción y además se debe tomar en cuenta el factor operador-dependiente al realizar el ultrasonido.

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

## Capítulo III

### Matriz de Co-ocurrencias y Análisis Multivariante

En el desarrollo de este capítulo se introducen las bases teóricas de la matriz de co-ocurrencias y del análisis multivariante que incluye las pruebas de ANOVA, MANOVA y  $T^2$  de Hotelling. A su vez se aplica la técnica de procesamiento de imágenes para extraer las características de textura con las cuales se va a realizar el análisis estadístico para determinar en cual fase de la tomografía computarizada es donde se tienen las diferencias estadísticas significativas. Para medir matemáticamente la textura se aplica la matriz de co-ocurrencia de niveles de grises o con sus siglas en ingles GLCM (Grey Level Co-occurrence Matrix), basadas en estadísticas de 2<sup>do</sup> orden. Los métodos de 2<sup>do</sup> orden consideran la relación de co-ocurrencia entre grupos de dos píxeles de la imagen original, según una distancia definida. Posteriormente se realiza el análisis multivariante con los parámetros de textura extraídos de las imágenes para contrastar la significación estadística de las diferencias entre grupos aplicando las técnicas ANOVA, MANOVA y  $T^2$  de Hotelling.

#### 3.1. Matriz de Co-ocurrencias

La matriz de co-ocurrencias fue introducida por Haralick (Haralick y otros, 1973). Los resultados de muchos trabajos demuestran la eficacia de este método para caracterizar las texturas, y por ello es la referencia para cualquier estudio de textura. La matriz de co-ocurrencia es la forma natural de estimar la estadística de segundo orden del nivel de intensidad. La base de este método es el modelo Markoviano que limita la dependencia espacial interpixel a un vecindario local [27]. La co-ocurrencia de los niveles de gris de una imagen digital se pueden especificar mediante una matriz de frecuencias relativas con elementos  $P_{ij}$  en la que dos pixeles vecinos separados por una distancia  $d$  y una dirección  $\theta$  ocurren en la imagen, uno con nivel de gris  $i$  y el otro con nivel de gris  $j$ . La matriz de co-ocurrencia considera la relación

espacial entre dos píxeles, llamados pixel de referencia o de interés y su pixel vecino. Cada elemento de la matriz de co-ocurrencia  $C_{ij}$  es el número de veces que en toda la matriz de entrada pasamos del nivel de gris  $i$  al nivel de gris  $j$  para una distancia y ángulo determinados. La matriz de co-ocurrencia normalizada se define como [28]:

$$P_{i,j} = \frac{C_{i,j}}{\sum_{i,j=0}^{N-1} C_{i,j}} \quad (3.1)$$

### 3.1.1. Parámetros de Textura

Haralick propone 14 parámetros derivados de la matriz de co-ocurrencia para caracterizar la textura. Los más interesantes por razones de interpretación y simplicidad son los siguientes [28]:

- **Contraste:** Proporciona información acerca de las variaciones bruscas del nivel de gris de la imagen.

$$CON = \sum_{i,j} |i - j|^2 P_{(i,j)} \quad (3.2)$$

- **Correlación:** Proporciona una medida de la probabilidad de la relación entre los diferentes niveles de gris de los píxeles.

$$COR = \sum_{i,j} \frac{(i - \bar{i})(j - \bar{j})P_{(i,j)}}{\sigma_i \sigma_j} \quad (3.3)$$

Dónde:

$$\bar{i} = \sum_i \sum_j i P_{(i,j)} \text{ es el valor medio de los niveles de gris.} \quad (3.3.1)$$

$$\sigma = \sum_i \sum_j (i - \bar{i})^2 P_{(i,j)} \text{ es la varianza de los niveles de gris.} \quad (3.3.2)$$

- **Energía:** Proporciona una idea de la suavidad de la textura.

$$ENER = \sum_{i,j} P_{(i,j)}^2 \quad (3.4)$$

- **Homogeneidad:** Proporciona información sobre la regularidad local de la textura.

$$HOM = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \quad (3.5)$$

- **Entropía:** En el análisis de textura es una medida de su desorden espacial.

$$ENT = \sum_{i,j} p_{i,j} (d, \theta) \ln(p_{i,j} (d, \theta)) \quad (3.6)$$

### 3.2. Análisis de Varianza Multivariante

Los modelos de análisis de varianza Multivariante (MANOVA) son una generalización directa de los univariantes. Lo único que varía es que la respuesta que se estudia es un vector para cada observación, en lugar de una variable aleatoria escalar. Ello conlleva que las sumas de cuadrados cuyos cocientes proporcionan los contrastes de las diferentes hipótesis, sean ahora formas cuadráticas generalizadas. Los estadísticos de contraste, por su parte, serán cocientes de determinantes (con distribución  $\Lambda$  de Wilks) o diferentes funciones de valores propios de ciertas matrices. La exposición que sigue presupone familiaridad con el modelo de análisis de varianza univariante [29].

#### 3.2.1. Modelo MANOVA con un tratamiento

Estudiamos una característica Multivariante  $Y_{i,j}$  que suponemos generada así:

$$Y_{i,j} = \mu_{ij} + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (3.7)$$

$$\epsilon_{ij} \sim N(0, \Sigma)$$

En la ecuación anterior,  $Y_{i,j}$  es el vector de valores que toma la v.a. multivariante estudiada para el caso  $j$ -ésimo sujeto al tratamiento  $i$ -ésimo. De existir un efecto atribuible al nivel  $i$ -ésimo del tratamiento, éste vendría recogido por el vector  $\alpha_i$ . Supondremos el mismo número de casos estudiados con cada nivel del

único tratamiento (es decir, consideraremos sólo el caso de diseño equilibrado): hay  $k$  niveles y la muestra incluye  $n$  casos tratados con cada nivel.

La hipótesis de interés más inmediato sería:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (\Leftrightarrow \alpha_i = 0 \quad \forall_i) \quad (3.8)$$

$$\text{versus } H_a: \mu_i \neq \mu_j \quad \text{para algún } i, j \quad (3.9)$$

De un modo enteramente similar a como sucede en el caso ANOVA univariante, la suma generalizada de cuadrados en torno a la media  $Y_{..}$  se descompone así:

$$\sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - Y_{..}) (Y_{ij} - Y_{..})' \quad (3.10)$$

$$= \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - Y_{i.} + Y_{i.} - Y_{..}) (Y_{ij} - Y_{i.} + Y_{i.} - Y_{..})' \quad (3.11)$$

$$= \underbrace{\sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - Y_{i.}) (Y_{ij} - Y_{i.})'}_E + \underbrace{n \sum_{i=1}^k (Y_{i.} - Y_{..}) (Y_{i.} - Y_{..})'}_H \quad (3.12)$$

Ahora bien, la teoría anterior, muestra que las matrices aleatorias E y H en la expresión anterior tienen distribuciones respectivas,

$$E \sim W(k(n-1), \Sigma) \quad (3.13)$$

$$H \stackrel{H_0}{\sim} W(k-1, \Sigma) \quad (3.14)$$

La distribución de E se sigue de los supuestos; la de H es correcta cuando la hipótesis nula es cierta. Además, hay independencia entre ambas matrices Wishart, en virtud del teorema 1.3. En consecuencia, bajo la hipótesis nula,

$$\Lambda = \frac{|E|}{|E+H|} \sim \Lambda_{p, k-1, k(n-1)} \quad (3.15)$$

Si  $H_0$  no se verifica,  $H$  “engordará”: será una Wishart no central. Son valores pequeños del estadístico  $\Lambda$  anterior los que cabe interpretar como evidencia contra la hipótesis nula [29].

### 3.2.1.1.Relación entre diversos contrastes

Observemos que si  $\delta_1, \dots, \delta_p$  son los valores propios de  $E^{-1}H$ ,

$$\Lambda = \frac{|E|}{|E+H|} = \prod_{i=1}^p \left\{ \frac{1}{1+\delta_i} \right\} \quad (3.16)$$

El estadístico de contraste es una particular función de los autovalores de  $E^{-1}H$ . No es la única elección posible: hay otras que mencionamos brevemente.

- **Estadístico máxima raíz de Roy**

$$\theta = \frac{\delta_1}{1+\delta_1} \quad (3.17)$$

- **Estadístico de Pillai**

$$V = \sum_{i=1}^p \frac{\delta_i}{1+\delta_i} \quad (3.18)$$

- **Estadístico de Lawley-Hotelling**

$$U = \sum_{i=1}^p \delta_i \quad (3.19)$$

De todos ellos hay tabulaciones que permiten contrastar  $H_0$  con comodidad. Su comportamiento es diferente dependiendo del tipo de incumplimiento de la hipótesis  $H_0$ . Por ejemplo, el estadístico de Roy está particularmente indicado cuando los vectores de medias  $\mu, \dots, \mu$  están aproximadamente alineados: esto hace crecer el primer valor propio de  $H$  y de  $E^{-1}H$ . En cambio, cuando los vectores de medias son diferentes y no están alineados, los otros estadísticos proporcionarán en general más potencia [30].

### 3.2.2. Modelos MANOVA con dos o más tratamientos

De modo análogo a como sucede en el caso univariante, un modelo MANOVA con dos tratamientos supone que la respuesta (Multivariante)  $Y_{ijk}$  (correspondiente al  $k$ -ésimo caso, tratado con los niveles  $i$  y  $j$  de los tratamientos A y B respectivamente) se genera alternativamente de una de las

Tabla 3.1: *Análisis de Varianza para un modelo con dos tratamientos e interacción.*

Fuente	Suma cuadrados	G.L
A	$H_A = KJ \sum_{i=1}^I (Y_{i..} - Y_{...}) (Y_{i..} - Y_{...})'$	I - 1
B	$H_B = KI \sum_{j=1}^J (Y_{.j.} - Y_{...}) (Y_{.j.} - Y_{...})'$	J - 1
AB	$H_{AB} = K \sum_{i=1}^I \sum_{j=1}^J (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...})$ $\times (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...})'$	(I - 1)(J - 1)
Error	$E = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (Y_{ijk} - Y_{ij.}) (Y_{ijk} - Y_{ij.})'$	IJ (k - 1)
Total	$T = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (Y_{ijk} - Y_{...}) (Y_{ijk} - Y_{...})'$	IJK - 1

siguientes formas (sin y con interacción, respectivamente):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk} \quad (3.20)$$

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (3.21)$$

El análisis es entonces reminiscente del que se realiza en el caso univariante. Las sumas de cuadrados del análisis univariante son ahora sumas de cuadrados generalizados: matrices que, bajo los supuestos de normalidad Multivariante y de vigencia de las respectivas hipótesis de contraste, se distribuyen como Wishart. A título puramente ilustrativo transcribimos en la anterior tabla la partición de la suma generalizada de cuadrados para un modelo con dos tratamientos e interacción.

Podemos ahora construir contrastes para las hipótesis de nulidad de cada uno de los efectos, empleando el estadístico  $\Lambda$  de Wilks. Si empleamos el primero tendríamos, por ejemplo, que bajo la hipótesis  $H_A: \alpha_i = 0$  para  $i = 1, \dots, I$ ,

$$\Lambda_A = \frac{|E|}{|E+H|} \sim \Lambda_{p, I-1, J(K-1)} \quad (3.22)$$

Y valores suficientemente pequeños de  $\Lambda_A$  conducirán al rechazo de la hipótesis. Similares cocientes de sumas de cuadrados generalizadas permitirían contrastar cada una de las restantes hipótesis de interés.

Salvo el contraste basado en el estadístico de Roy, los demás son bastante robustos a la no normalidad y a la heterogeneidad en las matrices de covarianzas de los vectores de observaciones. Son bastante sensibles, en cambio, a la no independencia de las observaciones. La robustez al incumplimiento de las hipótesis es en general menor cuando aumenta la dimensión [30].

### 3.3. Distribución $T^2$ de Hotelling

Sea  $W \sim W_d(n, \Sigma)$  y  $X \sim N_d(\mu, \Sigma)$ , ambas independientes. Entonces:

$$n(X - \mu)'W^{-1}(X - \mu) \quad (3.23)$$

Sigue la distribución conocida como  $T^2$  de Hotelling, de dimensión  $d$  y con  $n$  grados de libertad.

La denotaremos por  $T_{d,n}^2$ . Esta distribución puede verse como una generalización de la  $\mathcal{F}_{1,n}$  (y, por tanto,  $T$  como una generalización de la  $t$  de Student). En efecto, cuando  $d = 1$ ,

$$W \sim W_1(n, \sigma^2) = \sigma^2 x_n^2 \quad (3.24)$$

$$X \sim N(\mu, \sigma^2) \quad (3.25)$$

y:

$$n(X - \mu)'W^{-1}(X - \mu) = \frac{(X - \mu)^2}{W/n} = \frac{\left(\frac{X - \mu}{\sigma}\right)^2}{W/n\sigma^2} \sim \mathcal{F}_{1,n} \quad (3.26)$$

No es preciso contar con tablas de la distribución de Hotelling, pues una relación muy simple la liga con la distribución  $\mathcal{F}$  de Snedecor [30].

### 3.4. Metodología Aplicada

La base de datos que se utilizó está conformada por imágenes provenientes de un equipo de tomografía computarizada (TC). Los estudios tomográficos utilizados fueron provenientes de un tomógrafo General Electric light speed II de 64 cortes localizado en el Centro de Imagenología de la Policlínica Barquisimeto en el Estado Lara. Las imágenes fueron adquiridas en formato DICOM con una resolución de 512x512. Se aplicaron diferentes técnicas de pre-procesamiento, se calcula la matriz de co-ocurrencias y posteriormente se realiza el análisis estadístico de los resultados.

- ***Pre-procesamiento***

Se desarrolló un algoritmo de segmentación manual para trazar un polígono o región de interés que representa el hígado en las imágenes tomográficas de pacientes sanos y enfermos (*figura 3.1*). Dicha región se subdivide en una cuadrícula con ventanas de tamaño 4x4. Posteriormente se calcula la matriz de co-ocurrencias.

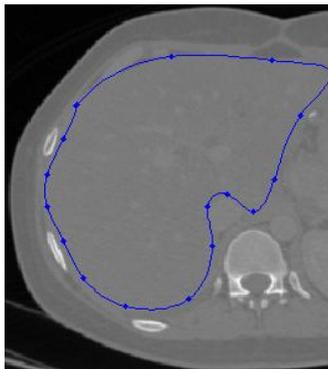


Figura 3.1: *Imagen Tomográfica, hígado, tomado de [24].*

Actualmente el diagnóstico de la esteatosis hepática a través de la tomografía computarizada la realiza un especialista observando la imagen y realizando comparaciones entre el contraste del hígado y el bazo del paciente. Hoy en día en el

área de salud se está dando uso de la tecnología para crear sistemas autónomos que permitan diagnosticar enfermedades.

En este capítulo se realiza la extracción de ventanas 4x4 del hígado las cuales contienen las características correspondientes a la matriz de co-ocurrencias como son el contraste, correlación, energía y homogeneidad. Dado a que luego se van a aplicar máquinas de soporte vectorial para realizar la clasificación de pacientes sanos y enfermos (con esteatosis hepática) por esto se necesita información correspondiente a las tomografías computarizadas para ser usadas como *input* del *software*, tanto características de las tomografías computarizadas de pacientes sanos y con esteatosis hepática.

Para garantizar un buen funcionamiento de la máquina de soporte vectorial es indispensable que las cuatro características discriminen, o caractericen, las TC de pacientes sanos y pacientes con esteatosis hepática, por tanto se hace necesario contrastar las siguientes hipótesis:

No existe diferencia significativa del vector de medias (CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD) para las poblaciones:

- Paciente Sano (Fase Simple) Vs. Paciente con Esteatosis Hepática (Fase Simple).
- Paciente Sano (Fase Arterial) Vs. Paciente con Esteatosis Hepática (Fase Arterial).
- Paciente Sano (Fase Venosa) Vs. Paciente con Esteatosis Hepática (Fase Venosa).

Lo que se quiere lograr es identificar diferencias significativas para el vector de medias formado por las características CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD de la matriz de co-ocurrencias asociada a ventanas de 4 x 4 pixeles en imágenes de tomografías computarizadas de hígado de pacientes

diagnosticados como sanos y con esteatosis hepática tomadas en tres diferentes fases simple, arterial y venosa.

Se realizó un análisis de estadísticas descriptivas (medias, desviación estándar, diagramas de caja, entre otros) para identificar indicios sobre posibles diferencias entre las poblaciones, seguido a esto se realizó una prueba de comparación de vectores de medias ( $T^2$  de Hotelling), un análisis de varianza univariante y multivariante para contrastar las hipótesis antes expuestas.

En el análisis se utilizaron 6 conjuntos de datos, cuyas descripciones son las siguientes:

Tabla 3.2: Descripción de los Datos.

ÓRGANO	DIAGNÓSTICO	FASE	NUMERO DE MATRICES DE CO-OCURRENCIAS
Hígado	Sano	Simple	914
Hígado	Enfermo con Esteatosis Hepática	Simple	920
Hígado	Sano	Arterial	894
Hígado	Enfermo con Esteatosis Hepática	Arterial	901
Hígado	Sano	Venosa	952
Hígado	Enfermo con Esteatosis Hepática	Venosa	916

### 3.5. Resultados Obtenidos

Se desarrolló un programa en Statistical Analysis System (SAS), versión remota 9.1.2, el cual es un software de análisis estadístico de datos donde se realizaron las comparaciones de los atributos obtenidos de la matriz de co-ocurrencia compuesta por contraste, correlación, energía y homogeneidad aplicando las pruebas estadísticas: análisis de varianza univariante (ANOVA), análisis de varianza multivariante (MANOVA) y  $T^2$  de Hotelling, usando un  $\alpha = 0.05$  como nivel de significación máximo admitido para el contraste de las hipótesis estadísticas.

### 3.5.1. Fase Simple

**3.5.1.1** Análisis de las características CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD de la matriz de co-ocurrencias asociados a ventanas de 4x4 pixeles en imágenes tomográficas de hígados de pacientes diagnosticados con esteatosis hepática y pacientes sanos.

Estadísticas básicas para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en las poblaciones de Pacientes con Esteatosis Hepática (PEH) y Pacientes Sanos (PS).

Tabla 3.3: *Estadísticas Básicas de Pacientes Sanos (PS) en Fase Simple de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	914	1.1725018	0.5048107	0.3333333	3.8333333
CORRELACION	914	0.2634356	0.2905475	-0.7838495	0.8778626
ENERGIA	914	0.1713773	0.04811	0.0833333	0.4722222
HOMOGENIDAD	914	0.6521472	0.0760217	0.4097222	0.8611111

Tabla 3.4: *Estadísticas Básicas de Pacientes con Esteatosis Hepática (PEH) en Fase Simple de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	920	1.2380435	0.5204605	0.25	3.5
CORRELACION	920	0.237866	0.2990756	-0.762729	0.9035943
ENERGIA	920	0.1680103	0.0465028	0.0972222	0.4722222
HOMOGENIDAD	920	0.644769	0.0794757	0.4236111	0.875

Al observar las *tablas 3.3 y 3.4* se puede analizar que los valores medios para cada una de las variables son mayores para la población PS, por otro lado las desviaciones estándar para las variables, excepto ENERGIA, son mayores en la población PEH. De acuerdo a las *tablas 3.3 y 3.4* no se observan grandes diferencias para el conjunto de datos PEH y PS en la distribución univariada de las cuatro variables.

**3.5.1.2** Análisis de Varianza Unidimensional para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS.

A continuación se muestran los resultados del análisis de varianza unidimensional para contrastar el efecto del factor DIAGNÓSTICO con dos niveles PEH y PS en cada una de las variables dependientes CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD.

Tabla 3.5: ANOVA para Contraste, Correlación, Energía, Homogeneidad en Fase Simple de la TC.

ANALISIS DE VARIANZA UNIVARIANTE			
VARIABLE	MODELO	P-Valor	CONCLUSIÓN
CONTRASTE	$DIAGNOSTICO$ $= \beta_0 + \beta_1$ $* CONTRASTE + \epsilon$	0.0063	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>CONTRASTE</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.
CORRELACIÓN	$DIAGNOSTICO$ $= \beta_0 + \beta_1$ $* CONTRASTE + \epsilon$	0.0635	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>CORRELACION</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.
ENERGIA	$DIAGNOSTICO$ $= \beta_0 + \beta_1 * ENERGIA$ $+ \epsilon$	0.1277	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>ENERGIA</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.
HOMOGENIDAD.	$DIAGNOSTICO$ $= \beta_0 + \beta_1$ $* HOMOGENIDAD + \epsilon$	0.0424	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>HOMOGENIDAD</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.

**3.5.1.3** Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS ( $T^2$  de Hotelling).

Tabla 3.6: Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS ( $T^2$  de Hotelling).

IGUALDAD DE VARIANZAS	N1	N2	p	V	$T^2$ CALCULADO	$T^2_{(4;v;0.05)}$	CONCLUSIÓN
SI	914	920	4	1832	7.7070876	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente <b>iguales</b> en las poblaciones PEH y PS
No	914	920	4	1831	7.7084481	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente <b>iguales</b> en las poblaciones PEH y PS

**3.5.1.4** Análisis de Varianza Múltiple.

Tabla 3.7: MANOVA modelo CONTRASTE CORRELACION ENERGIA HOMOGENIDAD = DIAGNÓSTICO en Fase Simple de la TC.

Criterio de test MANOVA y estadísticos F exactos para la hipótesis de efectos CATEGORIA no generales H = Tipo III Matriz SSCP para CATEGORIA Matriz SSCP de error E = S=1 M=1 N=913.5					
Estadístico	Valor	F-Valor	Num DF	Den DF	Pr > F
Wilks' Lambda	0.9958107	1.92	4	1829	0.1039
Pillai's Trace	0.0041893	1.92	4	1829	0.1039
Hotelling-Lawley Trace	0.00420693	1.92	4	1829	0.1039
Roy's Greatest Root	0.00420693	1.92	4	1829	0.1039

El análisis de varianza multivariado, *tabla 3.7*, no mostró resultados significativos para el factor DIAGNÓSTICO a un nivel de 0.05, debido a que el P-valor para los estadísticos Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace y Roy's Greatest Root son mayores a este nivel; por tanto el vector de medias formado por las variables CONTRASTE CORRELACION ENERGIA HOMOGENIDAD es estadísticamente igual en las poblaciones PEH y PS.

### 3.5.2. Fase Arterial

**3.5.2.1** Análisis de las características CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD de la matriz de co-ocurrencias asociados a ventanas 4x4 pixeles en imágenes tomográficas de hígados de pacientes diagnosticados con esteatosis hepática y pacientes sanos.

Estadísticas básicas para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en las poblaciones de Pacientes con Esteatosis Hepática (PEH) y Pacientes Sanos (PS).

Tabla 3.8: *Estadísticas Básicas de Pacientes Sanos en Fase Arterial de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	894	1.2376025	0.5214332	0.3333333	3.3333333
CORRELACION	894	0.23689	0.292875	-0.7938842	1
ENERGIA	894	0.1684377	0.0465087	0.0972222	0.4722222
HOMOGENIDAD	894	0.6437438	0.0785014	0.4097222	0.8611111

Tabla 3.9: *Estadísticas Básicas de Pacientes con Esteatosis Hepática en Fase Arterial de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	901	1.2791343	0.5449039	0.1666667	3.5
CORRELACION	901	0.2268804	0.2942679	-0.5683986	0.8783101
ENERGIA	901	0.1675916	0.0467478	0.0972222	0.3888889
HOMOGENIDAD	901	0.6364456	0.0797619	0.4236111	0.9166667

Las *tablas 3.8 y 3.9* dejan ver que no existen grandes diferencias entre los valores medios de las variables CONTRASTE, CORRELACIÓN, ENERGIA y

HOMOGENIDAD, al igual que en las desviaciones estándar, para las poblaciones PS y PEH.

**3.5.2.2** Análisis de Varianza Unidimensional para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS.

Tabla 3.10: ANOVA para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en Fase Arterial de la TC.

ANÁLISIS DE VARIANZA UNIVARIANTE			
VARIABLE	MODELO	P-Valor	CONCLUSIÓN
CONTRASTE	$DIAGNOSTICO = \beta_0 + \beta_1 * CONTRASTE + \epsilon$	0.0992	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>CONTRASTE</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.
CORRELACIÓN	$DIAGNOSTICO = \beta_0 + \beta_1 * CONTRASTE + \epsilon$	0.4702	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>CORRELACION</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.
ENERGIA	$DIAGNOSTICO = \beta_0 + \beta_1 * ENERGIA + \epsilon$	0.7007	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>ENERGIA</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.
HOMOGENIDAD.	$DIAGNOSTICO = \beta_0 + \beta_1 * HOMOGENIDAD + \epsilon$	0.0509	El análisis de varianza <b>no mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>HOMOGENIDAD</b> es estadísticamente <b>igual</b> en las poblaciones PEH y PS.

**3.5.2.3** Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS ( $T^2$  de Hotelling).

Tabla 3.11: Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en Pacientes con EH y Pacientes Sanos ( $T^2$  de Hotelling).

IGUALDAD DE VARIANZAS	N1	N2	p	V	$T^2$ CALCULADO	$T^2_{(4;v;0.05)}$	CONCLUSIÓN
SI	894	901	4	1793	6.0315866	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente iguales en las poblaciones PEH y PS
No	894	901	4	1792	6.0332705	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente iguales en las poblaciones PEH y PS

**3.5.2.4** Análisis de Varianza Múltiple.

Tabla 3.12: MANOVA modelo CONTRASTE CORRELACION ENERGIA HOMOGENIDAD= DIAGNÓSTICO en Fase Arterial de la TC.

Criterio de test MANOVA y estadísticos F exactos para la hipótesis de efectos DIAGNÓSTICO no generales					
H = Tipo III Matriz SSCP para DIAGNÓSTICO					
Matriz SSCP de error E =					
S=1 M=1 N=894					
Estadístico	Valor	F-Valor	Num DF	Den DF	Pr > F
Wilks' Lambda	0.99664731	1.51	4	1790	0.198
Pillai's Trace	0.00335269	1.51	4	1790	0.198
Hotelling-Lawley Trace	0.00336396	1.51	4	1790	0.198
Roy's Greatest Root	0.00336396	1.51	4	1790	0.198

El análisis de varianza multivariado, *tabla 3.12*, no mostró resultados significativos para el factor DIAGNÓSTICO a un nivel de 0.05, debido a que el P-valor para los estadísticos Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace y Roy's Greatest Root son mayores a este nivel; por tanto el vector de medias formado por las variables CONTRASTE CORRELACION ENERGIA HOMOGENIDAD es estadísticamente igual en las poblaciones PEH y PS.

### 3.5.3. Fase Venosa

**3.5.3.1.** Análisis de las características CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD de la matriz de co-ocurrencias asociados a ventanas de 4x4 pixeles en imágenes tomográficas de hígados de pacientes diagnosticados con esteatosis hepática y pacientes sanos.

Estadísticas básicas para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en las poblaciones de Pacientes con Esteatosis Hepática (PEH) y Pacientes Sanos (PS).

Tabla 3.13: *Estadísticas Básicas de Pacientes Sanos en Fase Venosa de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	952	1.1769958	0.5042999	0.25	3.4166667
CORRELACION	952	0.2789639	0.2865423	-0.5695898	0.9076366
ENERGIA	952	0.1712768	0.0447013	0.0972222	0.4166667
HOMOGENIDAD	952	0.6490648	0.0786761	0.3819444	0.875

Tabla 3.14: *Estadísticas Básicas de Pacientes con Esteatosis Hepática en Fase Venosa de la TC.*

Variable	Número de observaciones	Media	Desviación estándar	Mínimo	Máximo
CONTRASTE	916	1.250182	0.5312487	0.1666667	4.0833333
CORRELACION	916	0.2452152	0.3128834	-0.6350366	0.9431191
ENERGIA	916	0.1672277	0.0417387	0.0833333	0.3888889
HOMOGENIDAD	916	0.6401246	0.0825109	0.4236111	0.9166667

En las *tablas 3.13 y 3.14* se puede observar que la población PS presenta valores medios más altos, en comparación con la población PEH, para las variables CONTRASTE, CORRELACIÓN, ENERGIA y HOMOGENIDAD. Por otro lado se observa que las variables CONTRASTE, CORRELACIÓN Y HOMOGENIDAD

presentan mayor variabilidad en la población PEH, a diferencia de la variable ENERGIA cuya variabilidad es mayor en la población PS.

**3.5.3.2.** Análisis de Varianza Unidimensional para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS.

Tabla 3.15: ANOVA para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en Fase Venosa de la TC.

ANÁLISIS DE VARIANZA UNIVARIANTE			
VARIABLE	MODELO	P-Valor	CONCLUSIÓN
CONTRASTE	$DIAGNOSTICO = \beta_0 + \beta_1 * CONTRASTE + \epsilon$	0.0023	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05 por tanto el valor medio de la variable <b>CONTRASTE</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.
CORRELACIÓN	$DIAGNOSTICO = \beta_0 + \beta_1 * CONTRASTE + \epsilon$	0.0151	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>CORRELACION</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.
ENERGIA	$DIAGNOSTICO = \beta_0 + \beta_1 * ENERGIA + \epsilon$	0.0433	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>ENERGIA</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.
HOMOGENIDAD.	$DIAGNOSTICO = \beta_0 + \beta_1 * HOMOGENIDAD + \epsilon$	0.0166	El análisis de varianza <b>mostró</b> resultados significativos para el factor <b>DIAGNÓSTICO</b> a un nivel de 0.05, por tanto el valor medio de la variable <b>HOMOGENIDAD</b> es estadísticamente <b>diferente</b> en las poblaciones PEH y PS.

**3.5.3.3. Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en PEH y PS ( $T^2$  de Hotelling).**

Tabla 3.16: *Análisis Multivariante para CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD en Pacientes con EH y Pacientes Sanos ( $T^2$  de de Hotelling).*

IGUALDAD DE VARIANZAS	N1	N2	p	V	$T^2$ CALCULADO	$T^2_{(4;v;0.05)}$	CONCLUSIÓN
SI	952	916	4	1866	10.006765	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente <b>diferentes</b> en las poblaciones PEH y PS
No	952	916	4	1847	9.9892619	9.488	Los vectores de medias CONTRASTE, CORRELACIÓN, ENERGIA, HOMOGENIDAD son estadísticamente <b>diferentes</b> en las poblaciones PEH y PS

**3.5.3.4. Análisis de Varianza Múltiple.**

Tabla 3.17: *MANOVA modelo CONTRASTE CORRELACION ENERGIA HOMOGENIDAD= DIAGNÓSTICO en Fase Venosa de la TC.*

Criterio de test MANOVA y estadísticos F exactos para la hipótesis de efectos DIAGNÓSTICO no generales					
H = Tipo III Matriz SSCP para DIAGNÓSTICO					
Matriz SSCP de error E =					
S=1 M=1 N=930.5					
Estadístico	Valor	F-Valor	Num DF	Den DF	Pr > F
Wilks' Lambda	0.994666	2.5	4	1863	0.0409
Pillai's Trace	0.005334	2.5	4	1863	0.0409
Hotelling-Lawley Trace	0.005363	2.5	4	1863	0.0409
Roy's Greatest Root	0.005363	2.5	4	1863	0.0409

El análisis de varianza multivariado, *tabla 3.17*, mostró resultados significativos para el factor DIAGNÓSTICO a un nivel de 0.05, debido a que el P-valor para los estadísticos Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace y Roy's Greatest Root son menores a este nivel; por tanto el vector de medias formado por las variables CONTRASTE CORRELACION ENERGIA HOMOGENIDAD es estadísticamente diferente en las poblaciones PEH y PS.

### **3.6. Conclusiones**

En la fase simple de la tomografía computarizada solo se observó que existe diferencia estadísticamente significativa en el contraste y homogeneidad del análisis de varianza univariante (ANOVA), tanto en la  $T^2$  de Hotelling y en el análisis de varianza multivariante (MANOVA) no hubo resultados estadísticamente significativos. Al interpretar estos resultados se puede concluir que en esta fase de la TC no hay cambios en el hígado comparando pacientes sanos y pacientes con esteatosis hepática.

Los resultados obtenidos en la fase arterial de la tomografía computarizada mostró que no hay diferencias estadísticamente significativas en ninguna de las técnicas estadísticas aplicadas como son análisis de varianza univariante (ANOVA),  $T^2$  de Hotelling y en el análisis de varianza multivariante (MANOVA).

La fase venosa de la tomografía computarizada arrojó resultados estadísticamente significativos en todas las técnicas estadísticas aplicadas lo cual coincide con el diagnóstico que hacen los especialistas en esta fase. Se observó que existen diferencias estadísticamente significativas al comparar el hígado de un paciente sano y el de un paciente con esteatosis hepática lo cual confirma que la fase venosa de la TC es la que debe ser usada para detectar esta patología hepática.

## Capítulo IV

# Máquinas de Soporte Vectorial y Regresión Logística Binaria

En este capítulo se explica la teoría de las máquinas de soporte vectorial y de la regresión logística binaria. Se aplican las máquinas de soporte vectorial y la regresión logística binaria para realizar la clasificación de los pacientes con esteatosis hepática y pacientes sanos. Para ello se uso como datos de entrada las características extraídas de la matriz de co-ocurrencias. A continuación se realizo la comparación de la clasificación de las dos técnicas aplicadas tanto para las imágenes de tomografía computarizada como para imágenes de ultrasonido hepático.

### 4.1. Conceptos Fundamentales Máquinas de Soporte Vectorial

Las máquinas de soporte vectorial (SVM) son una técnica de clasificación, basadas en la Teoría de Aprendizaje Estadístico (SLT). Constituyen una de las principales herramientas de aprendizaje que implementan el principio inductivo de minimización de riesgo estructural, para obtener buenas cualidades de generalización sobre un número limitado de patrones de aprendizaje [31].

Esta teoría toma como base matemática la dimensión de Vapnik-Chervonenkis y la existencia de un hiperplano medio que sea capaz de establecer una frontera que maximice un margen de separación entre dos clases. Matemáticamente esto se representa haciendo una formulación lagrangiana del problema cuya solución final solo depende de los puntos muestrales que determinan los límites del margen, a los que la técnica denomina vectores de soporte [31].

Por otra parte, SVM resuelve el problema de la no linealidad de los datos incluyendo el uso de las funciones kernel. La idea de estos es lograr las transformaciones necesarias en el espacio de entrada sin tener que operar

directamente en un espacio de características de mayor dimensionalidad (si los puntos de entrada están en  $\mathbb{R}^2$  entonces con mapeados por la SVM a  $\mathbb{R}^3$ ) como se observa en la *figura 4.1*. Toda esta teoría está fundamentada sobre espacios de Hilbert con kernels reproductores entre los que se destacan: el kernel lineal, polinomial, de base radial y Sigmoid [31].

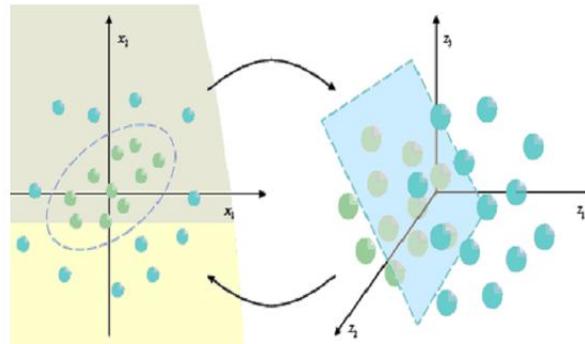


Figura 4.1: Transformación del espacio de entrada al espacio característico, tomado de [31].

www.bdigital.ula.ve

#### 4.1.1 Clasificación con Vectores de Soporte

Considérese un grupo de entrenamiento de tamaño  $N$  compuesto por pares atributo-etiqueta  $(\vec{x}_i, y_i)$ , donde  $1 \leq i \leq N$ ,  $\vec{x}_i \in \mathbb{R}^n$  y cada escalar  $y_i \in \{-1, 1\}$ , se desea estimar una ecuación del hiperplano que separe dicho grupo de entrenamiento, de manera que aquellos puntos con igual etiqueta se encuentren a un mismo lado del hiperplano. De esta manera, se puede presentar dos casos de forma general; cuando las clases son linealmente separables y cuando no lo son [31].

- **Conjuntos linealmente separables**

Dos conjuntos son linealmente separables cuando existe un hiperplano en  $\mathbb{R}^n$  determinado por un vector  $\vec{w}$  (vector normal que define al hiperplano) y un escalar  $b$  como se observa en la ecuación xx:

$$\vec{w}^T \vec{x}_i + b = 0, \quad \vec{w}, \vec{x}_i \in \mathbb{R}^n \quad y \quad b \in \mathbb{R} \quad (4.1)$$

para cualquier punto  $\vec{x}$  en el hiperplano.

La existencia de un hiperplano determinado por un vector  $\vec{w}$ , no es única, de hecho existen una infinidad de éstos vectores, por lo que es necesario establecer un criterio que permita determinar cuál de ellos se debe seleccionar para la clasificación. Éste criterio se basa en que si ambos conjuntos son separables entonces existe una distancia máxima o margen máximo entre ambos conjuntos [31] como se observa en la *figura 4.2*.

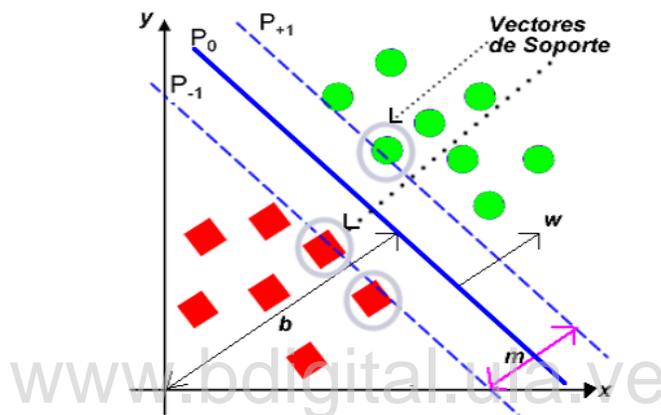


Figura 4.2: Esquema de dos conjuntos linealmente mediante un hiperplano ( $P_0$ ) de separación óptimo  $w^T \cdot x + b$  que maximiza el margen  $m$ , tomado de [31].

Así entre todos los posibles hiperplanos, aquel cuya distancia al patrón de entrenamiento más cercano es máxima se denomina el hiperplano óptimo de separación ( $P_0$ ). Cómo se puede observar en la *figura 4.2* paralelo al hiperplano ( $P_0$ ) se encuentran dos hiperplanos ( $P_{-1}$ ,  $P_{+1}$ ) definidos por las siguientes ecuaciones 4.2 y 4.3 [31].

$$P_{+1} = (\vec{w}^T \vec{x}_{+1} + b) = +1 \quad (4.2)$$

$$P_{-1} = (\vec{w}^T \vec{x}_{-1} + b) = -1 \quad (4.3)$$

Siendo el margen  $m$  la distancia  $\perp$  entre 4.2 y 4.3:

$$[(\vec{w}^T \cdot \vec{x}_{+1}) + b] - [(\vec{w}^T \cdot \vec{x}_{-1}) + b] = (+1) - (-1) \quad (4.4)$$

$$\vec{w}^T (\vec{x}_{+1} - \vec{x}_{-1}) = 2 \quad (4.5)$$

$$\frac{\vec{w}^T}{\|\vec{w}\|} \cdot (\vec{x}_{+1} - \vec{x}_{-1}) = \frac{2}{\|\vec{w}\|} \quad (4.6)$$

Todos los datos de entrenamiento deben cumplir:

$$(\vec{w}^T \cdot \vec{x}_{+1}) + b \geq +1 \quad \text{para } y_i = +1 \quad (4.7)$$

$$(\vec{w}^T \cdot \vec{x}_{-1}) + b \leq -1 \quad \text{para } y_i = -1 \quad (4.8)$$

Combinando las ecuaciones 4.7 y 4.8 se obtiene la ecuación 4.9

$$y_i (\vec{w}^T \cdot \vec{x}_i + b) \geq 1 \quad (4.9)$$

Si existe un hiperplano que satisfaga la ecuación 4.9 los datos son linealmente separables. Para encontrar el hiperplano óptimo se debe maximizar el margen (ecuación 4.6), teniendo en cuenta la restricción (ecuación 4.9) equivalente a resolver el problema:

$$\min_{\{w\}} \frac{1}{2} \vec{w}^T \vec{w} \quad (4.10)$$

sujeto a

$$y_i (\vec{w}^T \cdot \vec{x}_i + b) \geq 1, \quad \forall_i \quad (4.11)$$

- **Conjunto linealmente no separables**

El tratamiento para el caso de dos conjuntos no separables linealmente consiste en realizar una correspondencia del espacio de entrada a un espacio de representación de dimensión alta a través de una función no lineal elegida a priori, donde por medio de una función  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , con  $n, m \geq n$  ó  $m = \alpha$ , se trazan los datos de entrada  $\vec{x}_i \in \mathbb{R}^n$  a algún espacio de mayor dimensión llamado espacio característico (F) como se observa en la *figura 4.3*, con la variante que se introducen

variables de relajación ( $\xi_1, \dots, \xi_N$ ) en la ecuación 4.11 para ampliar la región factible, formulando así una nueva restricción [31].

$$y_i (\bar{w}^T \cdot \bar{x}_i + b) \geq 1 - \xi_i, \quad \forall_i \quad (4.12)$$

El propósito de las variables es permitir puntos erróneamente clasificados, los cuales correspondan a  $\xi_i > 1$ , por lo tanto, la suma de todos los  $\xi_i > 1$  es una cota superior del número de errores de entrenamiento.

Para encontrar el hiperplano clasificador óptimo se debe minimizar la función:

$$\min_{\{w\}} \left[ \frac{1}{2} \bar{w}^T \cdot \bar{w} + C \sum_{i=1} \xi_i \right] \quad (4.13)$$

sujeto a

$$y_i (\bar{w}^T \cdot \bar{x}_i + b) \geq 1 - \xi_i, \quad \forall_i \quad (4.14)$$

donde el parametro C es elegido a priori por el usuario de tal manera que un valor grande es una alta penalización a los errores.

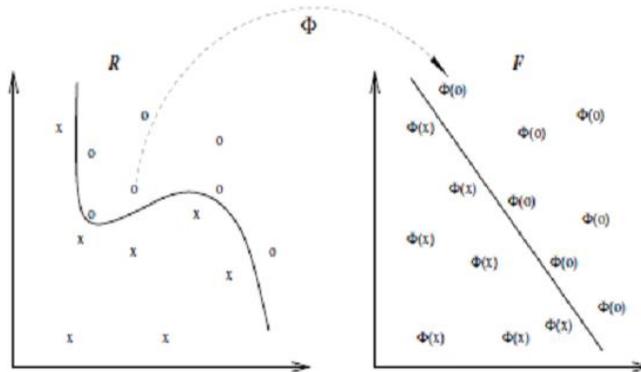


Figura 4.3: Transformación del espacio de entrada al espacio característico, tomado de [31].

#### 4.2. Máquinas de Soporte Vectorial de Mínimos Cuadrados

Uno de los inconvenientes de las SVM es su carga computacional en la resolución de restricciones para resolver el problema de optimización. Este

inconveniente ha sido superado por máquinas de soporte vectorial de mínimos cuadrados (LS-SVM), las cuales son una variante de las SVM que resuelven ecuaciones lineales en lugar de un problema de programación cuadrática [27].

#### 4.2.1. Software Máquinas de Soporte Vectorial de Mínimos Cuadrados (LSSVM)

El software de LSSVM se encuentra en la web en dos modalidades: un toolkit para Matlab bajo el Sistema Operativo Windows (utilizado en el presente trabajo) y el código fuente en lenguaje C para ser compilado en Linux.

Para ejecutar la primera modalidad se debe ejecutar el script denominado *installwindows.m* en el entorno de trabajo Matlab a través del cual se generan los archivos.dll necesarios para utilizar luego las funciones de LSSVM con Matlab.

Es importante acotar que en la carpeta de las LSSVM se puede apreciar un número significativo de archivos, sin embargos las funciones principales a utilizar por el usuario son: Para Matlab en Windows; *trainlssvm.m* para realizar el entrenamiento y *simlssvm.m* para la validación y clasificación, para Octave en Linux, *lssvmFILE.x* y *simFILE.x* para entrenamiento y validación respectivamente [31].

La función *trainlssvm.m* ó *lssvmFILE.x* se define de la siguiente manera  $[w,b] = \text{trainlssvm}(X,Y, \text{classification}, \text{gamma}, \text{sigma}, \text{RBFkernel}, \text{preprocess})$ . Ésta función realiza varias operaciones como por ejemplo:

1. Crea la LSSVM, usando Data (X), Etiquetas (Y) correspondiente a los valores de -1 (Marcador) y 1 (No Marcador), el tipo de proceso (clasificador ó estimación) a realizar por la máquina en los patrones, el valor de los parámetros del kernel (gamma y sigma) y tipo de kernel (para procesar la data en otro espacio).
2. Preprocesa estadísticamente la Data (Media cero y Varianza uno), el cual se representa por el último parámetro de la función (*preprocess*).
3. Con la data preprocesada entrena la LSSVM como clasificador.

4. Resuelve el sistema lineal de ecuaciones. Para ello, implementa dos métodos que pasan a ser activos de acuerdo a la complejidad del problema. Así, se obtiene los valores de los vectores de soporte ( $w$ ) y del umbral ( $b$ ).

Para la validación se utiliza la función  $Yh = \text{simlsvm}(X, Y, \text{classification}, \text{gam}, \text{sig2}, \text{kernel}, \text{preprocess}, w, b, Xt)$  mediante la cual, se clasifican las observaciones presentes en el conjunto de validación  $Xt$ . Ésta función recibe como parámetros el arreglo de Marcadores y No Marcadores  $X$ , el vector de etiquetas  $Y$  correspondiente a los valores de -1 para Marcador y 1 para No Marcador,  $\text{classification}$  es el procedimiento a aplicar a los patrones, valor de los parámetros  $\text{gamma}$  y  $\text{sigma}$  del kernel, tipo de kernel para procesar la data en otro espacio, tipo de data;  $\text{preprocessada}$  que es la data normalizada (Media cero y Varianza uno), los valores de los vectores de soporte  $w$  y del umbral  $b$  determinados en la etapa de entrenamiento, conjunto de validación  $Xt$ ,  $Yh$  representa el conjunto de etiquetas asignadas automáticamente por la LSSVM a cada una de las observaciones presentes en  $Xt$  [31].

### 4.3. Regresión Logística

La regresión logística es un instrumento estadístico de análisis bivariado o multivariado, de uso tanto explicativo como predictivo. Resulta útil su empleo cuando se tiene una variable dependiente dicotómica (un atributo cuya ausencia o presencia se ha puntuado con los valores cero y uno, respectivamente) y un conjunto de  $m$  variables predictoras o independientes, que pueden ser cuantitativas (que se denominan covariables o covariadas) o categóricas.

El propósito del análisis es:

- Predecir la probabilidad de que a alguien le ocurra cierto evento: por ejemplo, “estar desempleado” = 1 o “no estarlo” = 0; “ser pobre” = 1 o “no ser pobre” = 0; “graduarse” = 1 o “no graduarse” = 0.
- Determinar que variables pesan más para aumentar o disminuir la probabilidad de que a alguien le suceda el evento en cuestión.

Esta asignación de probabilidad de ocurrencia del evento a un cierto sujeto, así como la determinación del peso de cada una de las variables dependientes en esta probabilidad, se basan en las características que presentan los sujetos a los que, efectivamente, les ocurren o no estos sucesos. Por ejemplo, la regresión logística tomará en cuenta los valores que asumen en un serie de variables (edad, sexo, nivel educativo, posición del hogar, origen migratorio, etc) los sujetos que están efectivamente desocupados (= 1) y los que no lo están (= 0). En base a ello, predecirá a cada uno de los sujetos – independientemente de su estado real y actual – una determinada probabilidad de ser desocupado (es decir, de tener valor 1 en la variable dependiente). Es decir, si alguien es un joven no amo de casa, con baja educación y de sexo masculino y origen emigrante (aunque esté ocupado) el modelo le predecirá una alta probabilidad de estar desocupado (puesto que la tasa de desempleo de el grupo así definido es alta), generando una variable con esas probabilidades estimadas. Y procederá a clasificarlo como desocupado en un nueva variable, que será el resultado de la predicción. Además, analizará cuál es el peso de cada uno de estas variables independientes en el aumento o la disminución de esa probabilidad. Por ejemplo, cuando aumenta la educación disminuirá en algo la probabilidad de ser desocupado. En cambio, cuando el sexo pase de 0 = “mujer” a 1 = “varón”, aumentará en algo la probabilidad de desempleo porque la tasa de desempleo de los jóvenes de sexo masculino es mayor que la de las mujeres jóvenes. El modelo, obviamente, estima coeficientes de tales cambios.

Cuanto más coincidan los estados pronosticados con los estados reales del sujeto, mejor ajustará el modelo. Uno de los primeros indicadores de importancia para apreciar el ajuste del modelo logístico es el doble logaritmo del estadístico de verosimilitud (likelihood). Se trata de un estadístico que sigue una distribución similar a  $\chi^2$  y compara los valores de predicción con los valores observados en dos momentos: (a) en el modelo sin variables independientes, sólo con la constante y (b) una vez introducidas las variables predictoras. Por lo tanto, el valor de verosimilitud

debiera disminuir sensiblemente entre ambas instancias e, idealmente, tender a cero cuando el modelo predice bien [32].

#### 4.3.1. Regresión Logística Binaria

Los modelos de regresión logística binaria resultan los de mayor interés ya que la mayor parte de las circunstancias analizadas en medicina responden a este modelo (presencia o no de enfermedad, éxito o fracaso, etc). Como se ha visto, la variable dependiente será una variable dicotómica que se codificará como 0 ó 1 (respectivamente, “ausencia” y “presencia”). Este aspecto de la codificación de las variables no es banal (influye en la forma en que se realizan los cálculos matemáticos), y habrá que tenerlo muy en cuenta si se emplean paquetes estadísticos que no recodifican automáticamente las variables cuando éstas se encuentran codificadas de forma diferente (por ejemplo, el uso frecuente de 1 para la presencia y -1 ó 2 para la ausencia) [32].

La ecuación de partida en los modelos de regresión logística es:

$$\Pr(y = 1|x) = \frac{\exp(b_0 + \sum_{i=1}^n b_i x_i)}{1 + \exp(b_0 + \sum_{i=1}^n b_i x_i)} \quad (4.15)$$

Donde:  $\Pr(y = 1 | X)$  es la probabilidad de que  $y$  tome el valor 1 (presencia de la característica estudiada), en presencia de las covariables  $X$ :

- $X$  es un conjunto de  $n$  covariables  $\{x_1, x_1, \dots, x_n\}$  que forman parte del modelo.
- $b_0$  es la constante del modelo o término independiente.
- $b_1$  los coeficientes de las covariables.

Si se divide la expresión 4.15 por su complementario, es decir, si se construye su odds (en el ejemplo de presencia o no de enfermedad, la probabilidad de estar enfermo entre la probabilidad de estar sano), se obtiene una expresión de manejo matemático más fácil:

$$\frac{\Pr(y=1 | x)}{1-\Pr(y=1 | x)} = \exp(b_0 + \sum_{i=1}^n b_i x_i) \quad (4.16)$$

Si ahora se realiza su transformación logaritmo natural, se obtiene una ecuación lineal que lógicamente es de manejo matemático aún más fácil y de mayor comprensión:

$$\log\left(\frac{\Pr(y=1 | x)}{1-\Pr(y=1 | x)}\right) = b_0 + \sum_{i=1}^n b_i x_i \quad (4.17)$$

O simplificando:

$$\log\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_i x_i \quad (4.18)$$

En la expresión (4.18) se ve a la izquierda de la igualdad el llamado logit, es decir, el logaritmo natural de la odds de la variable dependiente (esto es, el logaritmo de la razón de proporciones de enfermar, de fallecer, de éxito, etc). El término a la derecha de la igualdad es la expresión de una recta, idéntica a la del modelo general de regresión lineal:

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n \quad (4.19)$$

Pero la regresión lineal presenta una diferencia fundamental respecto al modelo de regresión logística. En el modelo de regresión lineal se asume que los errores estándar de cada coeficiente siguen una distribución normal de media 0 y varianza constante (homoscedasticidad). En el caso del modelo de regresión logística no pueden realizarse estas asunciones pues la variable dependiente no es continua (sólo puede tomar dos valores, 0 ó 1, pero ningún valor intermedio). Llamando  $\varepsilon$  al posible error de predicción para cada covariable  $x_i$  se tendrá que el error cometido dependerá del valor que llegue a tomar la variable dependiente, tal como se ven en (4.20) [32].

$$y = \Pr(x) + \varepsilon \begin{cases} y = 1 \Rightarrow \varepsilon = 1 - \Pr(x) \\ y = 0 \Rightarrow \varepsilon = -\Pr(x) \end{cases} \quad (4.20)$$

Esto implica que  $\varepsilon$  sigue una distribución binomial, con media y varianza proporcionales al tamaño muestral y a  $\Pr(y = 1|x_i)$  (la probabilidad de que  $y = 1$  dada la presencia de  $x_i$ ).

#### 4.3.1.1. Estimación de los coeficientes del modelo y de sus errores estándar

Para la estimación de los coeficientes del modelo y de sus errores estándar se recurre al cálculo de estimaciones de máxima verosimilitud, es decir, estimaciones que hagan máxima la probabilidad de obtener los valores de la variable dependiente y proporcionado por los datos de nuestra muestra. Estas estimaciones no son de cálculo directo, como ocurre en el caso de las estimaciones de los coeficientes de regresión de la regresión lineal múltiple por el método de los mínimos cuadrados. Para el cálculo de estimaciones máximo – verosímiles se recurre a métodos iterativos, como el método de Newton-Raphson. Dado que el cálculo es complejo, normalmente hay que recurrir al uso de rutinas de programación o a paquetes estadísticos. De estos métodos surgen no sólo las estimaciones de los coeficientes de regresión, sino también de sus errores estándar y de las covarianzas entre las covariables del modelo.

El siguiente paso será comprobar la significación estadística de cada uno de los coeficientes de regresión del modelo. Para ello se pueden emplear básicamente tres métodos: el estadístico de Wald, el estadístico  $G$  de razón de verosimilitud y la prueba Score [32].

#### 4.3.1.2. Estadístico de Wald

Contrasta la hipótesis de que un coeficiente aislado es distinto de 0, y sigue una distribución normal de media 0 y varianza 1. Su valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo. En modelos con errores estándar grandes, el estadístico de Wald puede proporcionar falsas ausencias de significación (es decir, se incrementa el error tipo II) [33].

#### **4.3.1.3. El estadístico $G$ de razón de verosimilitud**

Se trata de ir contrastando cada modelo que surge de eliminar de forma aislada cada una de las covariables frente al modelo completo. En este caso cada estadístico  $G$  sigue una  $\chi^2$  con un grado de libertad (no se asume normalidad). La ausencia de significación implica que el modelo sin la covariable no empeora respecto al modelo completo (es decir, da igual su presencia o su ausencia), por lo que según la estrategia de obtención del modelo más reducido (principio de parsimonia), dicha covariable debe ser eliminada del modelo ya que no aporta nada al mismo. Esta prueba no asume ninguna distribución concreta, por lo que es la más recomendada para estudiar la significación de los coeficientes [33].

### **4.4. Metodología Aplicada**

#### **4.4.1. Máquinas de Soporte Vectorial de Mínimos Cuadrados**

Para la detección automática de pacientes con esteatosis hepática se usa una técnica de clasificación, Máquinas de Soporte Vectorial (SVM). Existe una variante de las SVM, denominada Máquina de Soporte Vectorial de Mínimos Cuadrados (LSSVM), resuelven un sistema equivalente de ecuaciones en un tiempo considerablemente menor comparado con las SVM. Con el propósito de realizar la detección automática de pacientes enfermos, se construye un clasificador LSSVM con un Kernel Gaussiano de base radial, utilizando el toolbox de Matlab (LS-SVM Versión 7.12, 32bits). El clasificador toma los atributos obtenidos de la matriz de co-ocurrencia y estadísticas descriptivas de interés. La clasificación fue aplicada a un conjunto de imágenes médicas obtenidas a través de tomografía computarizada y ultrasonido conformadas por dos grupos: pacientes con esteatosis hepática y pacientes sanos. El objetivo del entrenamiento de la LSSVM es construir una superficie de decisión que permita la clasificación de los patrones de entrada estos son los Marcadores y los No Marcadores, utilizando un kernel gaussiano y los hiperparámetros sigma y gamma. Donde sigma representa la varianza del kernel y

gamma es un parámetro de regularización. Estos hiperparámetros se consiguen de manera heurística hasta lograr un entrenamiento que clasifique los patrones [35].

#### **4.4.2. Análisis de Regresión Logística Binaria**

Son modelos estadísticos que buscan la relación entre 1) una variable dependiente cualitativa dicotómica (dos categorías) y 2) la o las variables explicativas independientes que pueden ser continuas o categóricas. Para esta investigación, la variable dependiente es el diagnóstico de Esteatosis Hepática, el cual toma dos posibles alternativas: enfermo o sano codificadas con 1 y 0 respectivamente.

Según Carrasco y Hernán [28], este tipo de análisis permite:

- 1) Cuantificar la importancia entre cada covariable y la variable dependiente, lo que lleva implícito clarificar relaciones confusas entre covariables y la variable dependiente a través de la detección de interacciones.
- 2) Clasificar individuos dentro de las categorías de la variable dependiente (enfermo/sano), según la probabilidad que tenga de pertenecer a una de ellas en función de las covariables que tengan efecto significativo.

La curva Característica de Operaciones del Receptor (COR), permite evaluar gráficamente la capacidad del modelo para discriminar. El área bajo la curva ROC muestra la probabilidad que tiene un individuo enfermo elegido al azar de presentar mayor probabilidad estimada de padecer la enfermedad que un individuo no enfermo elegido al azar. En la práctica, esta área es el porcentaje de pares de individuos enfermos y no enfermos en los que el enfermo tiene mayor probabilidad estimada de padecer la enfermedad que él no enfermo [30].

#### **4.4.3. Curvas ROC**

La metodología de curvas ROC aparece por primera vez durante la segunda guerra mundial como una técnica estadística para la detección y análisis de señales de radar. Tiene un uso extensivo en medicina donde recibe también el nombre de “Performance Characteristic Curves” o curva de rendimiento diagnóstico [32]. Las

curvas ROC tiene tres propósitos específicos: establecer un punto de equilibrio para el cual la sensibilidad y la especificidad alcanzan simultáneamente sus máximos valores (punto de corte óptimo), evaluar la capacidad discriminante de la prueba alternativa o diagnóstico y comparar dos o más test alternativos, a través de su capacidad discriminante [36].

Para la construcción de la curva ROC se toma la sensibilidad y la especificidad, construyendo un gráfico que deja bajo la curva un área que puede variar entre  $[0,1]$ , en este caso se considera el método no paramétrico para la estimación del área, la cual se debe alejar del valor 0.5. Una prueba tiene buena capacidad discriminante si es posible rechazar la hipótesis nula  $H_0: \text{área}=0.5$ . El punto de corte en la escala continua se calcula como el punto en el gráfico con el mayor índice de Youden, el cual se calcula como sensibilidad + especificidad – 1 [37].

#### **4.4.4. Prueba de T de Wilcoxon.**

La prueba de Wilcoxon es una comparación de tipo no paramétrico. Se basa en las diferencias de los pares de observaciones sobre un mismo individuo y la hipótesis de contraste está formulada sobre las distribuciones de las que provienen los datos. Es una técnica estadística alternativa a la aplicación de la prueba t de Student para muestras pareadas, cuando las diferencias  $d_i$  no se distribuyen normales.

##### **Hipótesis:**

$H_0$ : La distribución de las poblaciones I y II son idénticas Vs.

$H_1$ : La distribución de las poblaciones I y II tienen localizaciones diferentes.

#### **4.4.5. Estadístico de contraste:**

Para calcular el estadístico de contraste, se ordenan las distancias  $|d_i|$  de menor a mayor y se obtienen sus rangos, y se suman las diferencias  $d_i$  cuyo signo (positivo o negativo) tiene mayor frecuencia (se excluyen las  $d_i = 0$ ). A esta suma la llamaremos T.

De igual forma es necesario el cálculo de la cantidad  $T'$ , que corresponde a la suma de los rangos de las observaciones con signo de  $d_i$  de mayor frecuencia:

$$T' = m(n + 1) - T \quad (4.21)$$

Dónde  $m$  es el número de rangos con signo de  $d_i$  con menor frecuencia.

**Regla de decisión:**

Si  $T$  o  $T'$  es menor o igual que la cantidad tabulada de la Distribución Wilcoxon [33], se rechaza la hipótesis nula de contraste, afirmando que las distribuciones difieren de forma significativa.

**4.4.6. Prueba U de Mann-Whitney**

El estadístico  $U$  de Mann-Whitney es una prueba no paramétrica que plantea el contraste de hipótesis similar al planteado en la prueba  $T$  de Wilcoxon. La  $U$  se obtiene al ordenar las  $(n_1 + n_2)$  observaciones según su magnitud y al contar el número de observaciones en la muestra I que preceda a cada observación de la muestra II. El estadístico  $U$  es la suma de estas cantidades.

**4.4.7 Estadístico de contraste:**

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W, \quad (4.22)$$

Dónde:

$n_1$ = número de observaciones en la muestra I.

$n_2$ = número de observaciones en la muestra II,

$W$ = suma de rango para la muestra I.

El contraste de la prueba se hace comparando los resultados de  $U$  con los resultados de la distribución asociada al estadístico [33].

#### 4.5. Análisis de la Base de Datos

La base de datos que se utilizó está conformada por dos tipos de imágenes médicas provenientes de un equipo de tomografía computarizada y un equipo de ultrasonido. Los estudios tomográficos utilizados fueron provenientes de un tomógrafo General Electric light speed II de 64 cortes. Las imágenes fueron adquiridas en formato DICOM con una resolución de 512x512. Las imágenes de ecografía fueron adquiridas en un equipo de ultrasonido Esaote Maylab 50, en formato bmp y con una resolución de 800x600. La base de datos incluye imágenes tomográficas de 6 sujetos sanos y 8 pacientes con EH. Los pacientes son heterogéneos y van desde esteatosis leve a severa. Los pacientes con cualquier otra enfermedad hepática fueron excluidos de este estudio. Las imágenes de tomografía con contraste fueron adquiridas considerando tres fases: simple, arterial y venosa. En las imágenes de TC se consideró sólo la fase venosa, por presentar los mejores resultados de clasificación [38]. Adicionalmente 18 sujetos sanos y 29 pacientes con EH fueron adquiridas utilizando imágenes 2D de ultrasonido. Las imágenes de ultrasonido fueron adquiridas sin contraste y sin ningún protocolo de preparación especial. Este grupo también es heterogéneo y pacientes con otras enfermedades hepáticas fueron excluidos del estudio. Todas las imágenes fueron estudiadas por al menos tres médicos para establecer un diagnóstico clínico que se considera como la realidad sobre el terreno para el análisis estadístico. Los médicos trazaron las regiones de interés en cada imagen. Cada ROI se subdivide en un conjunto de ventanas de tamaño 4x4. Se estimó una matriz de co-ocurrencias para cada una de las ventanas. En esta investigación se consideró una cuantificación de 4 niveles de gris ( $N_g = 4$ ),  $\theta = 0^\circ$  y  $d = 1$  para el cálculo de los parámetros de la matriz de co-ocurrencias. Cada una de las imágenes de tomografía seleccionadas generó entre 166 y 336 matrices para un total de 10393 matrices y para las imágenes de ultrasonido se generó entre 116 y 694 matrices para un total de 16551 matrices. A partir de estas matrices, se estimaron 9 parámetros estadísticos descriptivos sobre la base de la matriz de co-ocurrencias de niveles de gris.

#### **4.5.1. Entrenamiento del Modelo**

Para el entrenamiento del modelo la población de las matrices se subdividen aleatoriamente en dos grupos: a) 5167 matrices de imágenes de tomografía y b) 8322 matrices de imágenes de ultrasonido. Un modelo de regresión logística binaria [39] fue ajustado considerando como variable dependiente la presencia o ausencia de la enfermedad y los parámetros estadísticos descriptivos como variables predictoras. Los parámetros descriptivos considerados fueron el contraste de textura (CON), la correlación de textura (COR), la energía de textura (ENER), la homogeneidad (HOM), la entropía (ENT), la media del nivel de gris en vecindades 4x4 (M), la varianza de niveles de gris (VAR), la desviación estándar de nivel de gris (DEV), el mínimo (MIN) y máximo nivel de gris (MAX). Con el objetivo de analizar si la EH se puede determinar eficazmente teniendo en cuenta los parámetros descriptivos extraídos de la textura y la información del nivel de gris también se realizó un procedimiento por etapas para la selección de variables significativas. La probabilidad de entrada fue de 0,1 y la probabilidad de salida fue de 0,01 para alcanzar el 99% de confianza en los resultados [40].

#### **4.5.2. Validación del Modelo**

Durante la validación todas las matrices fueron consideradas, así como su relación con respecto a un paciente particular. De esta manera el umbral óptimo para la clasificación de las matrices también proporciona un porcentaje de matrices clasificadas dentro de un sujeto determinado. La calidad de la clasificación se evaluó mediante la característica operativa del receptor (curva ROC).

#### **4.5.3. Clasificación de los Pacientes usando el Modelo de RLB**

Se realizó una prueba adicional para evaluar la capacidad del modelo RLB para la clasificación de pacientes. Siete muestras aleatorias, incluyendo 11 pacientes fueron extraídas del grupo de imágenes de ultrasonido. Cada paciente en la muestra

fue clasificado mediante el modelo RLB considerando el umbral previamente seleccionado. Un procedimiento similar se realizó considerando las imágenes de TC; sin embargo el número de pacientes en cada grupo fue de 6.

#### 4.5.4. Modelo Estadístico para Imágenes de Tomografía

El primer conjunto de imágenes analizadas fue tomado en las fases simple, arterial y venosa. La influencia de la fase se estudió mediante la realización de un análisis estadístico multivariante sobre muestras de matrices de co-ocurrencias extraídas al azar. Un análisis estadístico multivariante usando ANOVA, MANOVA y  $T^2$  de Hotelling [29] se llevó a cabo en cada una de las fases. Se obtuvieron los mejores resultados para la fase venosa coincidiendo con los médicos, los cuales usan esta fase para el diagnóstico de la esteatosis hepática mediante las imágenes de TC. Se estimó un modelo RLB como se explicó anteriormente. Las variables COR, ENT y MIN fueron excluidos del modelo, ya que no fueron significativas. Después de excluir las variables no significativas el ajuste del modelo es:

$$\begin{aligned} \ln \left[ \frac{P_r (EH)}{(1 - P_r (EH))} \right] \\ = 7.28 CON + 1.8 ENER + 4.82 HOM - 0.158 M - 0.02 VAR \\ + 0.499 DE + 0.0398 MAX \end{aligned}$$

dónde  $P_r (EH)$  representa la probabilidad de esteatosis hepática.

#### 4.5.5. Modelo Estadístico para Imágenes de Ultrasonido

En cuanto al modelo estadístico para las imágenes de ultrasonido, las variables significativas para la predicción de EH usando el modelo RLB fueron las mismas variables utilizadas en las imágenes de TC. Sin embargo, después de aplicar el método de selección por etapas, las variables correlación (COR), la energía (ENER) y la desviación estándar (DE) fueron excluidas. El modelo estimado es el siguiente:

$$\begin{aligned} \text{Ln} \left[ \frac{P_r (EH)}{(1 - P_r (EH))} \right] \\ = -12.21 - 1.62 \text{ CON} - 1.52 \text{ HOM} + 2.75 \text{ ENT} + 0.071 \text{ M} \\ - 0.0031 \text{ VAR} + 0.028 \text{ MIN} + 0.029 \text{ MAX} \end{aligned}$$

#### 4.5.6. Validación del Modelo

Los resultados de la clasificación de matrices 4x4 se muestran en la *tabla 4.1*. El área ROC es 0.833, con un rango de 0.679 y 0.987 con un intervalo de confianza del 95%. La tasa de clasificación correcta (TCC) fue de 82.14%. La sensibilidad  $S_e$  fue 93.75% y la especificidad  $S_p$  fue de 66.67%. La curva ROC para las imágenes de TC se muestra en la *figura 4.4*. En el caso de imágenes de ultrasonido el área ROC fue de 0.986, con un rango entre 0.887 y 0.999 para un intervalo de confianza asintótico de 95%. La tasa de clasificación correcta (TCC) fue 95.74%. La sensibilidad de  $S_e$  fue 95,45% y la especificidad  $S_p$  fue 96.00%. La curva ROC para las imágenes de ultrasonido se muestra en la *figura 4.5*. Los resultados muestran que ambos modelos son apropiados para la clasificación de las matrices de co-ocurrencias procedente de un sujeto enfermo o sano. Estos modelos de regresión clasifican matrices 4x4, pero ningún paciente. Para la clasificación de los pacientes, fue necesario establecer un umbral que representa el porcentaje necesario de matrices clasificadas en cada imagen para alcanzar la clasificación del paciente. Este umbral se establece en base a la curva ROC. Los umbrales óptimos seleccionados fueron 52% para las imágenes de ultrasonido y el 51,30% para las imágenes de TC. Los resultados obtenidos para la clasificación de los pacientes se muestran en la *tabla 4.2*. En cuanto a las imágenes de ultrasonido, la RLB clasifica correctamente EH en el 80% de las muestras seleccionadas y detecta los sujetos sanos en el 100% de las muestras. Los resultados de clasificación para la TC en fase venosa alcanzaron valores de sensibilidad y especificidad de 100%.

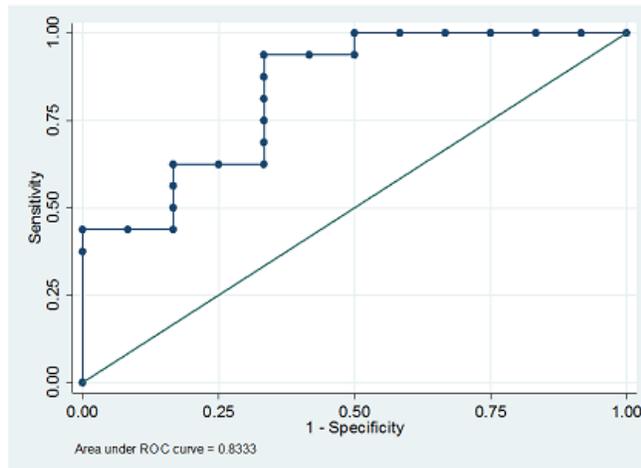


Figura 4.4: Curva ROC para clasificación de matrices 4x4 en imágenes de TC.

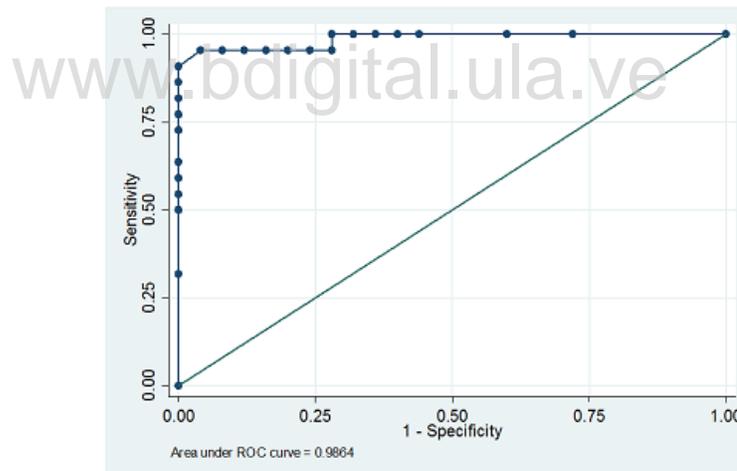


Figura 4.5: Curva ROC para clasificación de matrices 4x4 en imágenes de ultrasonido.

Tabla 4.1: *Puntos de corte óptimos para clasificar a los pacientes con esteatosis hepática, en base al modelo de RLB.*

	<b>Ultrasonido (n=16651)</b>	<b>Tomografía (n=10393)</b>
<b>Punto de corte</b>	52.00%	51.3%
<b>S<sub>e</sub></b>	95.45%	93.75%
<b>S<sub>p</sub></b>	96.00%	66.67%
<b>Tasa CC</b>	95.74%	82.14%
<b>Área</b>	0.986	0.833
<b>P</b>	0.000	0.003
<b>IC asintótico del LI</b>	0.887	0.679
<b>95% para el área LS</b>	0.999	0.987

Tabla 4.2: *Resultados de clasificación de pacientes con esteatosis hepática. La sensibilidad (S<sub>e</sub>) y la especificidad (S<sub>p</sub>) son calculados por cada modalidad de imagen.*

	<b>Ultrasonido</b>		<b>Tomografía</b>	
	<b>S<sub>e</sub></b>	<b>S<sub>e</sub></b>	<b>S<sub>p</sub></b>	<b>S<sub>p</sub></b>
<b>Muestra 1</b>	80	80	100	100
<b>Muestra 2</b>	60	100	100	100
<b>Muestra 3</b>	100	100	100	100
<b>Muestra 4</b>	80	100	100	50
<b>Muestra 5</b>	100	100	100	100
<b>Muestra 6</b>	100	100	100	100
<b>Muestra 7</b>	60	100	100	100

#### 4.5.7. Análisis de Resultados

Los resultados de las imágenes de TC con contraste muestran que el modelo de RLB es útil para la detección de EH con una tasa de clasificación correcta superior al 80% cuando se clasifican las matrices de tamaño 4x4. Aun cuando los resultados de clasificación de los pacientes usando imágenes de tomografía, (tabla 4.2) son

excelentes pero no son concluyentes por cuanto el número de pacientes es pequeño. Estos resultados sin embargo están de acuerdo con los reportados por [41] y [42] donde se muestra que la fase venosa es óptima para la detección de EH. Los resultados de las imágenes de ultrasonido para clasificación de las matrices son mejores que los resultados en imágenes de TC. En las imágenes de ultrasonido la TCC es de 95.74%, con altos valores de sensibilidad y especificidad. Los resultados de clasificación de los pacientes también son excelentes con una sensibilidad superior a 60% y especificidad del 100% en todas las muestras. Estos resultados concuerdan con los resultados reportados en [43] utilizando MSV. Sin embargo, en este caso no estamos clasificando las diferentes etapas de la enfermedad. Este método difiere del método propuesto en [44] porque se está usando sólo un pequeño conjunto de parámetros y adicionalmente en este caso la región de interés se subdivide en pequeñas ventanas. De esta manera se aprovecha de la información local para la realización de la clasificación con un área bajo la curva ROC mayor de 0.986. En esta investigación las matrices de co-ocurrencias se calcularon utilizando  $d = 1$ ,  $\theta = 0^\circ$  y  $N_g = 4$ , sin embargo se requiere más investigación para la comprensión de la influencia de estos parámetros en la precisión global del método. La RLB tiene bajo costo computacional, el software SAS (Statistical Analysis Software) funciona rápido, el procesamiento de los datos fueron segundos y su aplicación no es muy compleja.

#### **4.6. Clasificación de los Pacientes usando Máquinas de Soporte Vectorial**

Para la detección automática de pacientes con esteatosis hepática se usa una técnica de clasificación, Máquinas de Soporte Vectorial (SVM). Existe una variante de las SVM, denominada Máquina de Soporte Vectorial de Mínimos Cuadrados (LSSVM), resuelven un sistema equivalente de ecuaciones en un tiempo considerablemente menor comparado con las SVM. Con el propósito de realizar la detección automática de pacientes enfermos, se construye un clasificador LSSVM con un Kernel Gaussiano de base radial, utilizando el toolbox de Matlab (LS-SVM

Versión 7.12, 32bits). El clasificador toma los atributos obtenidos de la matriz de co-ocurrencias y estadísticas descriptivas de interés. La clasificación fue aplicada a un conjunto de imágenes médicas obtenidas a través de tomografía computarizada y ultrasonido conformadas por dos grupos: pacientes con esteatosis hepática y pacientes sanos. El objetivo del entrenamiento de la LSSVM es construir una superficie de decisión que permita la clasificación de los patrones de entrada estos son los Marcadores y los No Marcadores, utilizando un kernel gaussiano y los hiperparámetros sigma y gamma. Donde sigma representa la varianza del kernel y gamma es un parámetro de regularización. Estos hiperparámetros se consiguen de manera heurística hasta lograr un entrenamiento que clasifique los patrones [35].

#### **4.6.1. Etapa de Entrenamiento para Imágenes de Tomografía**

Para esta etapa el conjunto de datos de entrenamiento fue proveniente de los atributos extraídos de la matriz de co-ocurrencias (contraste, correlación, energía, homogeneidad, entropía) y estadísticas descriptivas de interés como son: media, varianza, desviación estándar, mínimo y máximo valor de pixel, de 4 pacientes con esteatosis hepática y 4 pacientes sanos. Se variaron los hiperparámetros sigma y gamma hasta que se logro encontrar los valores óptimos de entrenamiento.

#### **4.6.2. Etapa de Validación para Imágenes de Tomografía**

En esta etapa el conjunto de datos de validación fue proveniente de los atributos extraídos de la matriz de co-ocurrencias (contraste, correlación, energía, homogeneidad, entropía) y estadísticas descriptivas de interés como son: media, varianza, desviación estándar, mínimo y máximo valor de pixel, de 4 pacientes con esteatosis hepática y 2 pacientes sanos, los cuales no fueron usados en el conjunto de entrenamiento.

#### **4.6.3. Etapa de Entrenamiento para Imágenes de Ultrasonido**

Para esta etapa el conjunto de datos de entrenamiento fue proveniente de los atributos extraídos de la matriz de co-ocurrencias (contraste, correlación, energía, homogeneidad, entropía) y estadísticas descriptivas de interés como son: media, varianza, desviación estándar, mínimo y máximo valor de pixel, de 24 pacientes con esteatosis hepática y 12 pacientes sanos. Se variaron los hiperparámetros sigma y gamma hasta que se logro encontrar los valores óptimos de entrenamiento.

#### **4.6.4. Etapa de Validación para Imágenes de Ultrasonido**

En esta etapa el conjunto de datos de validación fue proveniente de los atributos extraídos de la matriz de co-ocurrencias (contraste, correlación, energía, homogeneidad, entropía) y estadísticas descriptivas de interés como son: media, varianza, desviación estándar, mínimo y máximo valor de pixel, de 5 pacientes con esteatosis hepática y 6 pacientes sanos, los cuales no fueron usados en el conjunto de entrenamiento.

#### **4.6.5. Análisis de Resultados**

Para realizar la clasificación de pacientes con esteatosis hepática a través de máquinas de soporte vectorial se usaron 7 grupos de combinaciones. Esto se realizó con el fin de verificar como clasificaba las LSSVM y poder estimar la sensibilidad y especificidad. Se entreno con un grupo conformado por 4 pacientes enfermos y 4 pacientes sanos y se validó con otro grupo diferente de 4 pacientes enfermos y 2 pacientes sanos para la TC. Para ultrasonido se trabajó con un grupo de entrenamiento de 24 pacientes con esteatosis hepática y 12 pacientes sanos y se validó con 5 pacientes con esteatosis hepática y 6 pacientes sanos. Los resultados obtenidos para TC se muestran en la *tabla 4.3* y para ultrasonido se muestran en la *tabla 4.4*. En las máquinas de soporte vectorial se requiere un mayor costo

computacional aunque el tiempo de respuesta fue bastante rápido, luego de obtener un buen entrenamiento la parte de la validación se hace bastante sencilla.

Tabla 4.3: *Resultados de la clasificación de pacientes con esteatosis hepática de imágenes de TC usando máquinas de soporte vectorial.*

	$S_e$	$S_p$
<b>Muestra 1</b>	100	100
<b>Muestra 2</b>	100	100
<b>Muestra 3</b>	100	100
<b>Muestra 4</b>	100	100
<b>Muestra 5</b>	100	100
<b>Muestra 6</b>	100	100
<b>Muestra 7</b>	100	100
<b>Mediana</b>	<b>100</b>	<b>100</b>

Tabla 4.4: *Resultados de la clasificación de pacientes con esteatosis hepática de imágenes de ultrasonido usando máquinas de soporte vectorial.*

	$S_e$	$S_p$
<b>Muestra 1</b>	100	83
<b>Muestra 2</b>	100	83
<b>Muestra 3</b>	100	83
<b>Muestra 4</b>	100	83
<b>Muestra 5</b>	100	83
<b>Muestra 6</b>	80	83
<b>Muestra 7</b>	80	100
<b>Mediana</b>	<b>100</b>	<b>83</b>

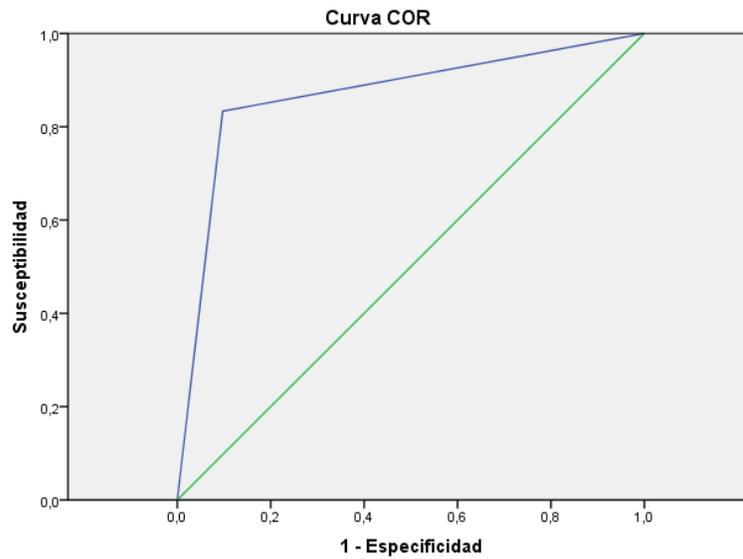


Figura 4.6: Curva ROC de TC.

Tabla 4.5: Resultado del Área bajo la Curva para TC.

**Área bajo la curva**

Variables resultado de contraste:MSV

Área	Error típ. <sup>a</sup>	Sig. asintótica <sup>b</sup>	Intervalo de confianza asintótico al 95%	
			Límite inferior	Límite superior
,868	,029	,000	,811	,925

Tabla 4.6: Tabla de Contingencia para TC.

		MSV				Total
		Enfermos		Sanos		
Dx	Positivos	(VP)	60	(FP)	12	72
	Negativos	(FN)	14	(VN)	130	144
Total			74		142	216

En la *figura 4.6* se muestra la curva ROC de TC. Se obtuvo un área bajo la curva de 0.868 (*tabla 4.5*) y con la tabla de contingencia (*tabla 4.6*) podemos

obtener la sensibilidad ( $S_e$ ) la cual indica la capacidad de clasificar correctamente a un individuo enfermo, su resultado es:

$$S_e = \frac{VP}{VP+FN} = \frac{60}{60+14} = 0.8108 = 81.08 \%$$

A su vez podemos obtener el valor de la especificidad ( $S_p$ ) la cual indica la probabilidad de clasificar correctamente a los sanos y su resultado es:

$$S_p = \frac{VN}{VN+FP} = \frac{130}{130+12} = 0.9154 = 91.54 \%$$

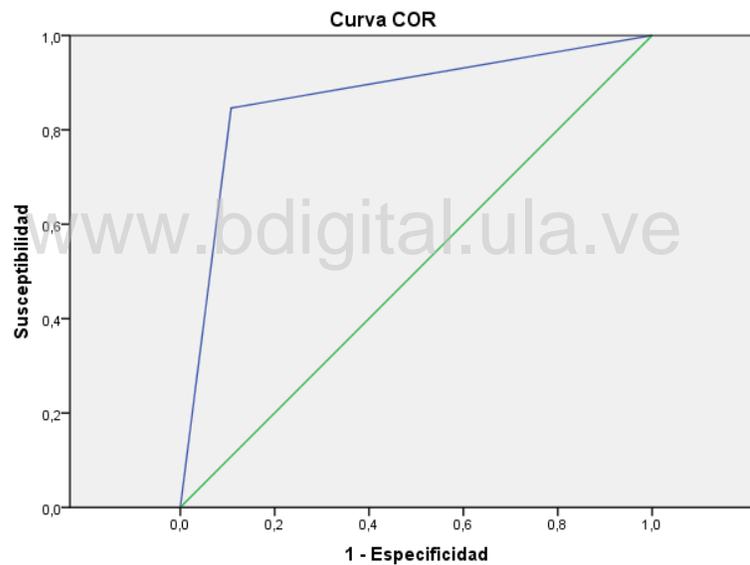


Figura 4.7: Curva ROC de US.

Tabla 4.7: Resultado del Área bajo la Curva para US.

**Área bajo la curva**

Variables resultado de contraste:MSV

Área	Error típ. <sup>a</sup>	Sig. asintótica <sup>b</sup>	Intervalo de confianza asintótico al 95%	
			Límite inferior	Límite superior
,869	,033	,000	,805	,933

Tabla 4.8: *Tabla de Contingencia para US.*

		MSV				Total
		Enfermos		Sanos		
Dx	Positivos	(VP)	66	(FP)	12	78
	Negativos	(FN)	7	(VN)	58	65
Total			73		70	143

La curva ROC de ultrasonido se muestra en la *figura 4.7*, en la cual se obtuvo un área bajo la curva de 0.869 (*tabla 4.7*) y con la tabla de contingencia (*tabla 4.8*) podemos obtener la sensibilidad ( $S_e$ ) la cual indica la capacidad de clasificar correctamente a un individuo enfermo, su resultado es:

$$S_e = \frac{VP}{VP+FN} = \frac{66}{66+7} = 0.9041 = 90.41 \%$$

A su vez podemos obtener el valor de la especificidad ( $S_p$ ) la cual indica la probabilidad de clasificar correctamente a los sanos y su resultado es:

$$S_p = \frac{VN}{VN+FP} = \frac{58}{58+12} = 0.8285 = 82.85 \%$$

En regresión logística binaria se trabaja con los individuos sobre los cuales se miden las características de estudio. Esos individuos en este caso para efectos estadísticos están representados por las matrices de co-ocurrencias, por esto la curva ROC para RLB se usan para determinar el punto de corte sobre el cual una cantidad de matrices de co-ocurrencias van a clasificar a un paciente como sano ó enfermo. En el caso de las máquinas de soporte vectorial se trabajo con todas las matrices de co-ocurrencias que contienen la región de interés de cada paciente, por lo tanto en MSV se clasifican pacientes debido a que la esteatosis hepática se encuentra de forma difusa en todo el hígado por esto se

hizo difícil realizar la identificación de cada matriz de co-ocurrencias, lo cual se deja como parte para otro estudio de investigación.

#### **4.7. Conclusiones**

Los modelos RLB son útiles para detectar EH basados en imágenes de TC en la fase venosa y en imágenes de ultrasonido. La tasa de clasificación correcta es de 82.14% en las imágenes de TC y de 95.74% en las imágenes de ultrasonido. La sensibilidad para TC fue de 93.75% y para US fue de 95.45% y la especificidad para TC fue de 66.67% y para US fue de 96%. Sin embargo, los resultados indican que la detección de EH en imágenes de ultrasonido es más precisa que en las imágenes de TC. Con las máquinas de soporte vectorial también se puede realizar la detección de esteatosis hepática tanto en imágenes de TC como de ultrasonido. Arrojando como resultado que se obtiene una mayor precisión en la detección de pacientes con EH por medio de imágenes de ultrasonido con una sensibilidad de 90.41% que en las imágenes de TC, la cual fue de 81.08%. En cuanto a la especificidad se puede señalar que en las imágenes de TC fue superior con un valor de 91.54% comparado con ultrasonido que fue de 82.85%. Al comparar ambas técnicas como es la regresión logística binaria y las máquinas de soporte vectorial se puede concluir que la RLB detecta mejor la esteatosis hepática y ambas técnicas coinciden con que la clasificación para las imágenes de ultrasonido es superior que para la tomografía computarizada.

## Capítulo V

### Conclusiones Generales y Recomendaciones

#### 5.1. Conclusiones Generales

- El resultado de la presente investigación consistió en el desarrollo de una herramienta de software para la detección de la esteatosis hepática en imágenes de tomografía computarizada y en imágenes de ultrasonido. Se puede utilizar como herramienta para la clasificación de pacientes con esta patología hepática asociada.
- El ultrasonido, la tomografía computarizada y la resonancia magnética son métodos de imágenes que permiten visualizar con facilidad la existencia de esteatosis hepática. Sin embargo, estas técnicas de imágenes no permiten diferenciar los diferentes grados de la enfermedad. La evolución de esta patología va a depender del estadio histológico de la enfermedad hepática. El examen histológico de la biopsia hepática es el que permite detectar con seguridad esta patología. Finalmente el trasplante hepático es una opción terapéutica para aquellos pacientes que se encuentren en un estadio muy avanzado de la Esteatosis Hepática como lo es la cirrosis avanzada.
- La confiabilidad de esta herramienta se validó con datos reales provenientes de imágenes de tomografía y de ultrasonido tanto de pacientes sanos como de pacientes con esteatosis hepática, las cuales fueron revisadas por médicos especialistas en el área.
- Los parámetros de textura como son el contraste, correlación, energía y homogeneidad presentan diferencias estadísticas significativas durante la fase venosa. El análisis estadístico que se aplicó en el capítulo III demuestra que existe una

diferencia estadística significativa entre los pacientes sanos y pacientes con esteatosis hepática durante la fase venosa coincidiendo con el diagnóstico que realizan los especialistas durante esta fase de la tomografía para diagnosticar la patología descrita.

- Se logró clasificar los pacientes con esteatosis hepática pero no la clasificación de los estadios de la enfermedad.
- La clasificación de los pacientes con esteatosis hepática fue satisfactoria con las máquinas de soporte vectorial y con el modelo de regresión logística binaria, ya que ambas presentaron valores de sensibilidad superiores al 80% y de especificidad superiores al 66.67%. Al comparar las dos técnicas como es la RLB y las MSV se puede concluir que la RLB detecta mejor la esteatosis hepática tanto en imágenes de tomografía como de ultrasonido y que ambas técnicas coinciden con que la detección de EH para las imágenes de ultrasonido es superior que para las imágenes de tomografía computarizada.

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)

## 5.2. Recomendaciones

- Como trabajo futuro se propone aumentar la muestra de imágenes tomográficas y realizar la clasificación de los pacientes mediante modelos multinivel, con el fin de comparar esos resultados con los encontrados hasta ahora.
- Para una mejor clasificación con Regresión Logística Binaria se propone agregar más datos a la muestra entre estos el sexo y la edad del paciente.
- Adicionalmente se recomienda aumentar el número de imágenes de cada modalidad para obtener una detección más precisa, con esto se puede ampliar el conjunto de entrenamiento y de validación de las máquinas de soporte vectorial.
- Se propone realizar más investigación para clasificar la esteatosis hepática en sus diferentes grados como son: leve, moderada y severa. Para esto se necesita obtener imágenes tanto de tomografía como de ultrasonido de diferentes pacientes que presenten todos los estadios de esta enfermedad hepática.
- Se plantea realizar una validación más completa que incluya un mayor número de datos (sexo, edad, estadio de la enfermedad) con lo cual se puede realizar un mejor diagnóstico.

## Bibliografía

- [1]. N. Chalasani., Younossi., Joel E., A. Mae Diehl., E. Brunt., K. Cursi., M. Charlton., A. Sanyal. “The Diagnosis and Management of Nonalcoholic Fatty Liver Disease. Practice Guideline by The American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology”. AGA. *Institute American College of Gastroenterology and American Society for the study of liver disease*. Gastroenterology 2012; 142:1592-1609.
- [2]. C. Williams., J. Stenge., M. Asike., D. Torres., J. Shaw., M. Contreras., C. Landt., S. Harrison. “Prevalence of Noalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis Among a Largely Middle-Aged Population Utilizing Ultrasonnd and Liver Biopsy: A Prospective Study”. *Clinical Advances in Liver, Pancreas, and Biliary Tract*. Gastroenterology 2011; 140:124-131.
- [3]. A. Awai., K. Newton., C. Sirlin., C. Behling., J. Schwimmer., “Evidence and Recommendations for Imaging Liver Fat in Children Based of Systematic Review”. *Clinical Gastroenterology and Hepatology* 2014; 12:765-773.
- [4]. Pedrosa, C., Casanova R. “Diagnóstico por Imagen Tratado de Radiología Clínica”, volumen 2. Mc Graw Hill, 2000.
- [5]. C. García Monzón. “Enfermedad Hepática Grasa no Alcohólica”. *Gastroenterología y Hepatología. Problemas comunes de la práctica clínica*. 2<sup>a</sup> Edición. España. 2012.
- [6]. L. Douglas. “Enfermedad del hígado graso no alcohólico y esteatohepatitis no alcohólica”. *Guías de la Organización Mundial de Gastroenterología*, Junio 2012.
- [7]. Qayyma A., Chena D., Breimana R., Westphalena A., Yeha B., Jonesb K., Lua Y., Coakleya F., Callena P. “Evaluation of diffuse liver steatosis by ultrasound,

computed tomography, and magnetic resonance imaging: which modality is best”. *Clinical Imaging*, (33): 110-115. 2009.

[8]. Davidson L., Kuk J., Church T., Ross R. “Protocol for measurement of liver fat by computed tomography”. *School of Physical and Health Education and Division of Endocrinology and Metabolism, Department of Medicine, Queen’s University, Kingston, Ontario, Canada. The Cooper Institute, Dallas, Texas (100): 864-868. 2006.*

[9]. Lupsor M., Badea R. “Imaging diagnosis and quantification of hepatic steatosis: Is it an accepted alternative to needle biopsy”. *Clinical Imaging. Department of Ultrasonography, University of Medicine and Pharmacy, Cluj-Napoca, Romania*, (4): 419-425. 2005.

[10]. Chun Yeh W., Jeng Y., Han Li C., Huang Lee P., Li P. “Liver steatosis classification using high frequency ultrasound”. *Ultrasound in Med and Biol*, (31): 599-605. 2005.

[11]. Gómez W., W.C.A Pereira., A.F.C Infantosi. “Analysis of Co-occurrence Texture Statistics as a Function of Gray-Level Quantization for Classifying Breast Ultrasound”. *IEEE TMI-2012-0297.R1.*

[12]. A. Bouchet, J. Cuilleret. “Anatomía Descriptiva, Topográfica y Funcional”. Editorial Médica Panamericana. Argentina, 1996.

[13]. Biblioteca Nacional de Medicina de EEUU. Medline Plus [Online]. Disponible: <https://www.nlm.nih.gov/medlineplus/spanish/ency/>

[14]. [Online]. Disponible: <http://es.slideshare.net/CiinthiaLunaa/anatomia-del-higado-39673459>

[15]. Henrie Rouvière, Andrés Delmas. “Anatomía Humana descriptiva, topográfica y funcional”. Editorial Masson. Tomo 2. España, 2005.

- [16]. B. Young, J. W Heath. Wheater's. "Histología Funcional". Editorial: Churchill Livingstone. Cuarta Edición. España, 2007. Pág 274.
- [17]. Joaquín Berenguer, Pascual Parrilla. "Transplante hepático". Editorial Elsevier. Segunda edición. España, 2008.
- [18]. Revista de Gastroenterología del Perú. [Online]. Disponible: [http://www.scielo.org.pe/scielo.php?pid=S102251292008000400002&script=sci\\_arttext](http://www.scielo.org.pe/scielo.php?pid=S102251292008000400002&script=sci_arttext)
- [19]. Kris Kowdley. "Esteatosis hepática no alcohólica y hepatopatías nutricionales". Capítulo 231. Martin H. Floch. Netter Gastroenterología. Editorial Masson. España, 2006.
- [20]. Juan Jaramillo Antillón. "Historia y Filosofía de la Medicina". Primera edición. Costa Rica, 2005.
- [21]. F. J Cabrero Fraile. "Imagen Radiológica principios físicos e instrumentación". Editorial Masson. España, 2006.
- [22]. Revista Española de Medicina Nuclear e Imagen Molecular. [Online]. Disponible: <http://www.elsevier.es/es-revista-revista-espanola-medicina-nuclear-e-125-articulo-introduccion-tomografia-computarizada-13088421>
- [23]. John Haaga, Charles F. Lanzieri, David J. Sartoris and Elias A. Zerhouni. "Tomografía Computarizada y Resonancia Magnética Diagnóstico por Imagen Corporal Total", volumen 2. Mosby, Madrid, 1994.
- [24]. Imágenes de Tomografía Computarizada suministradas por el Departamento de Imagenología de la Policlínica Barquisimeto. Estado Lara - Venezuela.
- [25]. José María Segura. "Ecografía Abdominal". Ediciones Norma. 2da edición. Madrid.

- [26]. Evan S. Siegelman. “Resonancia Magnética. Tórax, abdomen y pelvis. Aplicaciones Clínicas”. Editorial Médica Panamericana. España, 2005.
- [27]. G. Passariello., F. Mora. Imágenes Médicas. “Adquisición Análisis Procesamiento Interpretación”. Ediciones Equinoccio de la Universidad Simón Bolívar, Venezuela. 1.995.
- [28]. A. Pino., J.P. Mallofré., J.A. Roca. “Determinación de propiedades superficiales del papel utilizando el análisis de la textura del patrón de speckle”. Universidad Politécnica de Cataluña. Barcelona, España. 2009.
- [29]. Rencher, A. “Methods of Multivariate Analysis”. *Briagham Young University. WileyInterscience*. Canada. 2002.
- [30]. L. Díaz. “Estadística Multivariada: Inferencia y Métodos”. Universidad Nacional de Colombia. Bogotá, Colombia. 2007.
- [31]. J. Cuadros, R. Medina, “Procesamiento Avanzado de la Señal Electrocardiográfica de Pacientes Chagásicos”, Universidad de Los Andes, pp.34-38, Febrero 2012.
- [32]. Carrasco J. *et al* (1993). “Estadística multivariante en las ciencias de la vida”. Editorial Ciencia 3. Madrid, España.
- [33]. Hosmer D. *et al* (2000). “Applied Logistic regression”. Inc. 2ª Ed. John Willey & Sons.
- [34]. Dennis W, Mendenhall W, Scheafffer RL. “Estadística Matemática con aplicaciones”. Cengage Learning Editores S.A. 7ª Ed. 2010.

- [35]. Durán F, Medina R. “Visualización y Cuantificación de Espinas Dendríticas de Celulas de Purkinje en Imágenes por Microscopía de Luz”. *Tesis de Maestría. Dpto. de Ingeniería*. Universidad de Los Andes. Mérida-Venezuela, 2011.
- [36]. Burgueño M, García J, González J. “Las curvas ROC en la evaluación de las pruebas diagnósticas”. [Online] *MedClin*. 104:661-670. Disponible: [http://www.isodocsoft.com/downloads/documents/productossanitarios/curva\\_roc\\_pruebas\\_diagnosticas.pdf](http://www.isodocsoft.com/downloads/documents/productossanitarios/curva_roc_pruebas_diagnosticas.pdf). 1995.
- [37]. Cerda, Jaime. Y Cifuentes, Lorena. “Uso de Curvas ROC en Investigación clínica. Aspectos teórico-prácticos”. *Laboratorio de Infectología*. 29(2): 138-141. [Online]. Disponible: [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0716-10182012000200003](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-10182012000200003). 2012.
- [38]. Rivas, E., Ortiz, J., Medina, R., Moreno, F. “La Matriz de Co-ocurrencia como Herramienta para la Detección de la Esteatosis Hepática en Imágenes Tomográficas”. *Congreso Internacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas – Cimenics 2012*. Margarita, Venezuela.
- [39]. D. Hosmer, S. Lemeshow, and R. Sturdivant. “Applied Logistic Regression”. Wiley Series in Probability and Statistics. Wiley, 2013.
- [40]. C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll. “An introduction to logistic regression analysis and reporting”. *The Journal of Educational Research*, 96(1):3–14, 2002.
- [41]. R. d. F. Monjardim, D. M. C. Costa, R. F. T. Romano, P. S. Salvadori, J. d. V. C. d. Santos, A. C. V. Atzingen, D. C. Shigueoka, and G. D’Ippolito. “Diagnosis of hepatic steatosis by contrast-enhanced abdominal computed tomography”. *Radiología Brasileira*, 46:134 – 138, 06 2013.

- [42]. D. A. Lawrence, I. B. Oliva, and G. M. Israel. “Detection of hepatic steatosis on contrast-enhanced ct images: diagnostic accuracy of identification of areas of presumed focal fatty sparing”. *American Journal of Roentgenology*, 199(1):44–47, 2012.
- [43]. D. M. Mihilescu, V. Gui, C. I. Toma, A. Popescu, and I. Sporea. “Computer aided diagnosis method for steatosis rating in ultrasound images using random forests”. *Medical Ultrasound*, 15(3):184–190, 2013.
- [44]. A. Andrade, J. S. Silva, J. Santos, and P. Belo-Soares. “Classifier approaches for liver steatosis using ultrasound images”. *Procedia Technology*, 5:763–770, 2012.

[www.bdigital.ula.ve](http://www.bdigital.ula.ve)