

**Desarrollo de algoritmos de procesamiento de imagen
avanzado para interpretación de imágenes médicas.**

**Aplicación a segmentación de hígado sobre imágenes
de Resonancia Magnética multisequencia**



Arantza Bereciartua Pérez

Tesis Doctoral

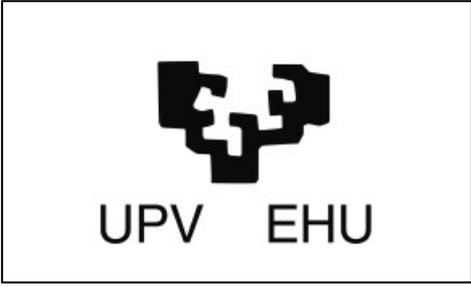
Departamento de Ingeniería de Sistemas y Automática

Universidad de País Vasco (UPV/EHU)

Tecnalia Research and Innovation

Directores: Dr. Pedro M^a Iriondo (UPV/EHU) y Dr. Artzai Picón
(Tecnalia)

Bilbao, 2016



Agradecimientos

Probablemente, este sea uno de los apartados más difíciles de la memoria. No sabe uno por dónde empezar.

En primer lugar, quiero agradecer a mis directores de tesis, Dr. Pedro Iriondo y Dr. Artzai Picón, toda la dedicación y apoyo en estos años de trabajo. Los principios fueron duros hasta enfocar el desarrollo, y luego vinieron las pruebas, el texto, las revisiones, y las exigentes publicaciones. Horas y horas de trabajo. En este proceso también ha participado activamente nuestro matemático favorito, Adrián Galdrán, que trajo las técnicas variacionales a mi vida, ya para quedarse. Gracias por esas revisiones minuciosas y siempre constructivas de las publicaciones y respuestas a editores.

También quiero agradecer al Dr. Gonzalo Solís, jefe del servicio de diagnóstico por imagen de CVSS Radiología Clínica, del IMQ, su disponibilidad a participar en este proyecto. Gracias por la aportación de los estudios para la realización de los desarrollos y las pruebas de esta tesis, así como por la fracción de conocimiento médico recibido, necesario para interpretar las imágenes radiológicas.

Los desarrollos realizados en esta tesis han sido parcialmente soportados por el proyecto HEPATOOL, financiado por el Ministerio de Ciencia e Innovación, con identificador de contrato IPT-2011-1514-900000.

La tesis propone un método concreto para un problema concreto, pero cuenta con todo el bagaje no escrito de mis 16 años en el grupo Computer Vision de Tecnalia, al que quiero recordar aquí: Esti, Jone, Jose, Antonio, Jan, Albertos, Cristina, Aitor, Miguel, Sergio (we miss you), y más que han pasado por él, con los que he disfrutado del trabajo y del ocio, con los que he compartido momentos interesantes en las instalaciones del cliente, que son y serán nuestras batallitas de abuelo. También quiero acordarme de Ana Ayerbe, directora del Área de Negocio de Tecnalia al que pertenecía en el momento de comienzo de la tesis, y de Patxi Estensoro, mi director actual, por su apoyo firme y necesario para poder completar este largo y a veces, tortuoso camino.

Gracias también a Aingeru por su ayuda insustituible. Acabar la tesis a los 40 con 2 hijos pequeños es misión imposible si alguien no aligera las tareas logísticas del hogar, y te permite disponer de huecos donde ya no los hay. Gracias a mis dos incansables terremotos, Aitor y Asier, por llenar de alegría mis días, y motivarme a esforzarme, queriendo ser un ejemplo para ellos en todos los aspectos, y también en la dedicación y el buen hacer en el trabajo.

Finalmente, pero no menos importante, quiero agradecer a mis padres todo lo que han hecho por mí. Sin ellos, no estaría aquí. Apostaron por la educación y la formación desde siempre, el mejor regalo que se le puede hacer a un hijo. Gracias a ti también, Javi, hermano favorito, por las chapas editoriales, pero sobre todo, por las pizzas.

Resumen:

La imagen médica se ha convertido en los últimos años en una potente herramienta de ayuda al diagnóstico. Gracias a los avanzados escáneres y software de reconstrucción de imágenes disponibles es posible la identificación de distintos órganos y tejidos, así como la obtención de datos que ayuden a caracterizar y cuantificar las patologías. Los radiólogos son los responsables del uso e interpretación de dichas imágenes y demandan herramientas que les permitan localizar órganos y tejidos con mayor precisión y rapidez, así como la identificación y caracterización cuantitativa de las patologías presentes en ellos, con el fin de realizar un diagnóstico preciso.

Por otra parte, el cáncer de hígado es una de las principales causas de muerte por cáncer en todo el mundo. Las técnicas invasivas utilizadas para su diagnóstico, tales como biopsias quirúrgicas, a veces pueden ser reemplazadas por técnicas no invasivas con imagen médica como la tomografía axial computerizada (TAC o CT, en sus siglas en inglés) y la resonancia magnética (RM o MRI, en inglés), con claros beneficios para el paciente. Con el fin de ayudar a los radiólogos y cirujanos en una planificación fiable de la intervención, son necesarios nuevos métodos y herramientas para localizar y segmentar adecuadamente el órgano de interés y las patologías presentes.

La segmentación (delimitación) automática del hígado es un problema complejo. Se han alcanzado resultados parciales principalmente sobre imágenes obtenidas mediante CT. La técnica de MRI ofrece mayor información para fines de diagnóstico. Sin embargo, la segmentación del hígado en imágenes de MRI representa un desafío debido a la presencia de artefactos característicos de dicha tecnología de adquisición, como es el caso de los volúmenes parciales, el ruido, y en general, la baja nitidez y el escaso contraste existente entre órganos, de manera que el límite entre los diferentes tejidos suele ser confuso. Existen menos desarrollos sobre imágenes de MRI, aunque éstos han ido en aumento progresivo en los últimos años.

En este trabajo, se presenta un nuevo método para la segmentación automática de hígado sobre imagen multicanal obtenida mediante resonancia magnética. El método propuesto consiste en la minimización de una superficie activa 3D mediante la aproximación dual a la formulación variacional subyacente del problema. Esta superficie activa evoluciona sobre un mapa de probabilidad que se basa en un nuevo descriptor compacto propuesto que contiene la información espacial y multiseuencia de cada píxel en relación a un modelo estadístico multivariable de hígado generado previamente. Esta superficie activa 3D integra de manera natural la regularización volumétrica. El descriptor visual compacto junto con el enfoque propuesto constituye un método de segmentación 3D rápido y preciso. El método fue probado en 18 estudios de hígado sano y los resultados se compararon con una segmentación de referencia realizada por expertos radiólogos. Las comparaciones con otros métodos del estado del arte se realizan mediante la obtención de 9 métricas establecidas. Los resultados obtenidos son comparables, incluso mejores en algunos casos, a los de otras técnicas del estado del arte. Se ha obtenido un coeficiente de similaridad de Dice de 98.59.

Abstract:

Medical imaging has become in recent years a powerful tool to support diagnosis. Thanks to advanced scanners and image reconstruction software available, it is possible to identify different organs and tissues, as well as obtaining data that may help to characterize and quantify the pathologies. Radiologists are responsible for the use and interpretation of these images and require tools that allow them to locate organs and tissues with greater accuracy and speed, as well as the identification and quantitative characterization of the pathologies present in them, in order to make an accurate diagnosis.

Moreover, liver cancer is one of the leading causes of cancer death worldwide. Invasive techniques used for diagnosis, such as surgical biopsies can sometimes be replaced by non-invasive techniques in medical imaging such as computed axial tomography (CT) and magnetic resonance imaging (MRI), with clear benefits for the patient. In order to assist radiologists and surgeons in a reliable intervention planning, new methods and accurate and efficient tools are needed to locate and segment properly the organ of interest and the pathologies inside.

Automatic segmentation (delimitation) of the liver is a complex problem. Partial results have been achieved mainly on images obtained by CT. MRI technique provides more information for diagnostic purposes. However, liver segmentation in MRI images is a challenge due to the presence of characteristic artifacts, such as the partial volumes, the noise, and generally, the low sharpness and the low contrast between organs, so that the boundary between different tissues is often confusing. There are fewer developments on MRI, although these have been steadily increasing in recent years.

In this thesis, we present a novel method for multichannel MRI automatic liver segmentation. The proposed method consists of the minimization of a 3D active surface by means of the dual approach to the variational formulation of the underlying problem. This active surface evolves over a probability map that is based on a new compact descriptor comprising spatial and multisequence information of every pixel which is further modeled by means of a liver multivariate statistical model that has been previously generated. This proposed 3D active surface approach naturally integrates volumetric regularization in the statistical model. The advantages of the compact visual descriptor together with the proposed approach result in a fast and accurate 3D segmentation method. The method was tested on 18 healthy liver studies and results were compared to a gold standard made by expert radiologists. Comparisons with other state-of-the-art approaches are provided by means of nine well established quality metrics. The obtained results are in line with the state-of-the-art methodologies, and are even better than them in some cases. A Dice Similarity Coefficient of 98.59 has been achieved.

INDICE

INDICE.....	vii
<i>Capítulo 1: INTRODUCCIÓN.....</i>	1
1.1. Motivación y trayectoria	3
1.2. Objetivos de la tesis.....	3
1.3. Contenido de la memoria	5
1.4. Herramientas empleadas y desarrollo SW para la tesis.....	6
<i>Capítulo 2: ESTADO DEL ARTE: Imágenes médicas. Tecnologías de adquisición, limitaciones y retos. Caso de aplicación: segmentación hepática en MRI.....</i>	7
2.1. Imagen Médica.....	8
2.2. Limitaciones y retos en procesamiento de imagen médica.....	11
2.3. Tecnologías de adquisición de imagen médica	14
2.3.1. Tomografía Axial Computerizada (TAC o CT).....	14
2.3.2. Resonancia Magnética Nuclear (RMN o MRI)	15
2.3.2.1. Comparativa MRI y CT.....	18
2.3.2.2. Problemas y Artefactos en imágenes de MRI	20
2.3.3. Otras tecnologías: Ultrasonido (US), PET, combinadas	23
2.4. Sistemas comerciales existentes para adquisición y procesamiento de imagen médica.....	24
2.5. Caso de aplicación: Segmentación hepática en MRI.....	28
2.5.1. Origen y Descripción del problema.....	28
2.5.2. Problemas y Limitaciones	29
<i>Capítulo 3: ESTADO DEL ARTE: Descriptores</i>	33
3.1. Concepto de descriptor.....	34
3.2. Descriptor de Nivel de gris.....	34
3.3. Descriptor de Forma.....	37
3.4. Descriptor de Borde o Gradiente.....	39
3.4.1. Sobel.....	39
3.4.2. Kirsch.....	40
3.4.3. Canny	40
3.4.4. SIFT.....	40

3.4.5. DSIFT.....	41
3.5. Descriptor de Textura.....	41
3.5.1. Textura basada en estadísticos	43
3.5.1.1. Estadísticos de 1 ^{er} orden: media y varianza	43
3.5.1.2. Estadísticos de 2 ^o orden: Matrices de co-ocurrencia y LBP	44
3.5.1.2.1. Matrices de co-ocurrencia o GLCM	44
3.5.1.2.2. Local Binary Patterns (LBP)	46
3.5.1.3. Estadísticos de 3 ^{er} y 4 ^o orden y superiores.....	48
3.5.2. Textura basada en métodos geométricos o estructurales	49
3.5.3. Textura basada en modelos.....	49
3.5.3.1. Campos aleatorios de Markov	50
3.5.3.2. Fractales.....	50
3.5.4. Textura basada en tratamiento de señal	51
3.5.4.1. Filtros en el dominio de Fourier	52
3.5.4.2. Filtros de Gabor	52
3.5.4.3. Wavelets.....	54
3.5.4.4. Transformada discreta del coseno.....	56
3.6. Aplicación en imagen médica.....	57
3.7. Aplicación a segmentación hepática	59
3.8. Conclusiones	61
Capítulo 4: ESTADO DEL ARTE: Segmentación.....	65
4.1. En qué consiste la segmentación	67
4.2. Introducción al problema de segmentación de hígado en imagen médica.....	67
4.3. Métodos de segmentación.....	69
4.3.1. Umbralización o <i>thresholding</i>	70
4.3.2. Detección de bordes	72
4.3.3. Crecimiento de regiones (<i>region growing</i>)	74
4.3.4. Métodos estadísticos.....	76
4.3.5. Agrupamiento o Clustering.....	77
4.3.5.1. Clustering no supervisado.....	78
4.3.5.2. Clustering supervisado o clasificación	80
4.3.6. Conocimiento a priori.....	84
4.3.7. Basada en modelos (a partir de atlas)	85
4.3.8. Cortes de grafos (graph cuts).....	88
4.3.9. Modelos Deformables	91
4.3.9.1. Contornos activos o <i>snakes</i>	92

4.3.9.2.	Isosurfaces y Level sets.....	99
4.3.9.3.	Modelos Activos de Forma o Active Shape Models (ASM)	103
4.3.9.4.	Modelos Activos de Apariencia o Active Appearance Models (AAM)	108
4.3.9.5.	Contornos Activos Sin Bordes o <i>Active Contours Without Edges</i> (ACWE)	109
4.3.9.6.	Contornos Geodésicos Activos o <i>Geodesic Active Contours</i> (GAC)	111
4.4.	Técnicas variacionales para cálculo de la solución óptima de segmentación	114
4.5.	Segmentación hepática en 3D	115
4.6.	Segmentación hepática sobre estudios multiseuencia en MRI.....	117
4.7.	Limitaciones de los métodos existentes para la segmentación de hígado en MRI	120
4.8.	Conclusiones	121
Capítulo 5: APROXIMACIÓN TEÓRICA		123
5.1.	Concepto de monosecuencia y multiseuencia	125
5.2.	Descriptor compacto basado en modelo: generación de un mapa de distancias.....	127
5.3.	Modelo de contorno activo	132
5.4.	Minimización del contorno activo: aproximación dual al problema variacional para minimización de funcionales de energía en 2D.....	133
5.5.	Extensión a 3D del modelo de contorno activo (superficie activa) y su proceso de minimización del funcional de energía.....	137
5.6.	Algoritmo completo de segmentación	138
5.6.1.	Entrada monosecuencia o multiseuencia	140
5.6.2.	Generación de un modelo gaussiano multivariante	140
5.6.3.	Descriptor compacto: Generación de un mapa de distancias basado en modelo a partir de la entrada	140
5.6.4.	Minimización del contorno activo definido mediante los métodos descritos en 2 y 3 dimensiones.....	141
5.6.5.	Etapas de post-procesado.....	141
5.6.6.	Ajuste de parámetros.....	143
Capítulo 6: PRUEBAS Y RESULTADOS		145
6.1.	Metodología de las pruebas.....	146
6.1.1.	Conjuntos de datos. Monosecuencia y multiseuencia	147
6.1.2.	Segmentación de referencia	148
6.1.3.	Métricas	150

6.1.4. Problemas que afectan los resultados: el registro de las secuencias y la calidad de la segmentación de referencia.....	155
6.2. Generación del modelo estadístico multivariable de hígado.....	157
6.3. Segmentación en 2D por minimización de contorno activo mediante técnicas variacionales y basado en modelo	157
6.2.1. Entrada Monosecuencia (raw VIBE data)	158
6.2.2. Entrada Monosecuencia (con mapa de distancias)	160
6.2.3. Entrada Multisección.....	163
6.4. Método de segmentación 3D por minimización de contorno activo mediante técnicas variacionales y basado en modelo	167
6.4.1. Entrada Monosecuencia	167
6.4.2. Entrada Multisección.....	171
6.5. Análisis de resultados	174
6.5.1. Contribución del descriptor compacto a los resultados de la segmentación ..	177
6.5.2. Contribución de la extensión a 3D a los resultados de la segmentación	182
6.6. Comparación con el estado del arte	190
6.7. Conclusiones	193
<i>Capítulo 7: CONCLUSIONES, APORTACIONES Y LÍNEAS FUTURAS.....</i>	195
7.1. Comentarios sobre el trabajo realizado	196
7.2. Conclusiones	197
7.3. Aportaciones de la tesis	199
7.4. Líneas de trabajo futuras	201
Referencias	204

Capítulo 1: INTRODUCCIÓN

La imagen médica se ha convertido en los últimos años en una potente herramienta de ayuda al diagnóstico. Gracias a los avanzados escáneres y software de reconstrucción de imágenes disponibles es posible la identificación de distintos órganos y tejidos, así como la obtención de datos que ayuden a caracterizar y cuantificar las patologías.

Los radiólogos son los responsables del uso e interpretación de dichas imágenes, cada vez más extendidas para el diagnóstico de enfermedades y su seguimiento posterior. Dichos facultativos demandan herramientas que les permitan localizar órganos y tejidos con mayor precisión y rapidez, así como la identificación y caracterización cuantitativa de las patologías presentes en ellos, con el fin de realizar un mejor diagnóstico.

El término *Medical Imaging* en inglés comprende todas las disciplinas técnicas necesarias para dotar la práctica clínica de dichas herramientas. *Medical Imaging* abarca el conjunto de técnicas específicas para la adquisición y reconstrucción de imagen por las múltiples tecnologías disponibles; la visualización de dichas imágenes; la compresión y almacenamiento de las mismas; la mejora de la calidad y el contraste; las tareas de registro junto con el procesamiento, segmentación y extracción de información de dichas imágenes, como puede ser la medida y cuantificación de parámetros anatómicos y fisiológicos. La integración de todas ellas permite abordar múltiples aplicaciones médicas.

El procesamiento y análisis de imagen médica es un problema complejo que requiere de un conocimiento especializado. Como se explicará más adelante, son imágenes en mapa de grises en las que existen muchos problemas derivados del proceso de captación (artefactos), así como ruidos y efectos de iluminación no homogénea. Esto hace necesario el uso de técnicas avanzadas para la segmentación correcta y precisa de los tejidos presentes en las imágenes.

Por otra parte, el cáncer de hígado es una de las principales causas de muerte por cáncer en todo el mundo. Las técnicas invasivas utilizadas para su diagnóstico, tales como biopsias quirúrgicas, a veces pueden ser reemplazadas por técnicas no invasivas con imagen médica como la tomografía axial computerizada (TAC o CT, en sus siglas en inglés) y la resonancia magnética (RM o MRI, en inglés), con beneficios evidentes para el paciente. Con el fin de ayudar a los radiólogos y cirujanos en una planificación fiable de la intervención, son necesarios nuevos métodos y herramientas precisas y eficientes para localizar y segmentar adecuadamente el órgano de interés y las patologías presentes.

La segmentación (delimitación) automática del hígado es un problema complejo. Se han alcanzado resultados parciales principalmente sobre imágenes obtenidas mediante CT. La técnica de MRI ofrece mayor información para fines de diagnóstico. Sin embargo, la segmentación del hígado en imágenes de MRI representa un desafío debido a la presencia de artefactos característicos, como los volúmenes parciales, ruido, y en general, baja nitidez y escaso contraste entre órganos, de manera que el límite entre los diferentes tejidos suele ser confuso. Existen menos desarrollos sobre MRI, aunque éstos han ido en aumento progresivo en los últimos años.

Además, en cada captura realizada mediante CT sólo es posible la adquisición de un conjunto de imágenes del cuerpo humano. Este conjunto de imágenes se denomina secuencia. Por el contrario, las técnicas de adquisición de resonancia magnética (MRI) están basadas en la respuesta de los tejidos a diferentes campos electromagnéticos. Cada configuración de un conjunto de parámetros de adquisición determinados provoca respuestas funcionales de los tejidos distintas, dando lugar a secuencias diferentes en el mismo estudio clínico. Este hecho permite a los radiólogos disponer de múltiples imágenes sobre los mismos tejidos, en las que los órganos y tejidos aparecen representados de manera diferente, lo que proporciona información complementaria valiosa para el diagnóstico.

1.1. Motivación y trayectoria

La doctoranda comenzó en el mundo de la investigación en el ámbito del procesamiento de imagen en el año 2000, en el centro tecnológico Robotiker, que en el año 2011 pasaría a formar parte de Tecnia. Por tanto, son 16 años trabajando en aplicaciones de visión artificial, en todas las fases, desde la conceptualización de la solución, pasando por el diseño y el desarrollo hasta la implantación del sistema en el cliente. Asimismo, durante este tiempo la doctoranda ha realizado numerosas publicaciones científicas, ha participado en conferencias nacionales e internacionales y ha contribuido a 2 patentes.

Si bien en los comienzos el procesamiento de imagen en las aplicaciones industriales fue el foco principal de los proyectos en los que participó la doctoranda, en el año 2010 abordó un nuevo campo de aplicación, el procesamiento de imagen médica. La industria biomédica estaba apostando fuerte por la automatización de parte de sus procesos, y la digitalización creciente de las imágenes diagnósticas posibilitaba el desarrollo de algoritmos que permitieran ayudar a los facultativos en su labor diaria. En este contexto, gran parte de los algoritmos que han sido aplicados con éxito en el mundo industrial pueden ser adaptados al procesamiento de las imágenes médicas.

En el proyecto HEPATOOL, financiado por el Ministerio de Ciencia e Innovación, con identificador de contrato IPT-2011-1514-900000, Tecnia participa con CVSS Radiología Clínica, que forma parte de IMQ, en concreto con el servicio de diagnóstico por imagen, liderado por el Dr. Gonzalo Solís, que planteó un caso de uso, esto es, la segmentación del hígado en MRI multisequencia. La dificultad en la solución del problema permitiría el desarrollo de algoritmos lo suficientemente novedosos para constituir una tesis. El reto era interesante, y ha sido abordado contando con la experiencia acumulada en procesamiento de imagen. En esta memoria se recoge el trabajo realizado.

1.2. Objetivos de la tesis

Los objetivos de esta tesis son:

- Proponer, desarrollar y validar nuevos métodos para segmentar el hígado de manera precisa en imágenes abdominales obtenidas mediante la técnica de resonancia magnética

- Proponer, desarrollar y validar una metodología para el manejo adecuado de la entrada multiseuencia de manera que contribuya a una segmentación más precisa.

En esta tesis, se presenta un nuevo método automático de segmentación hepática en 2D y 3D sobre imágenes de resonancia magnética multicanal (multiseuencia) por medio de la minimización de un modelo de contorno activo mediante técnicas variacionales. La minimización de un modelo de contorno activo volumétrico 3D (lo que se conoce como superficie activa) integra la regularización volumétrica en los resultados de la segmentación.

A pesar de que la combinación de múltiples secuencias simultáneas (conjuntos de imágenes distintos obtenidos sobre el mismo sujeto) es habitual en la práctica clínica mediante tecnología MRI, como se ha indicado previamente, los métodos automáticos que las manejan para extraer información son reducidos, y habitualmente consisten en añadir un número de dimensiones directamente proporcional al número de secuencias empleado en la adquisición. En esta tesis, se propone un nuevo descriptor compacto capaz de combinar la información multicanal mediante un modelo estadístico basado en distribución gaussiana multivariante y multiseuencia. Este enfoque aborda la distribución espacial de los valores de intensidad en las múltiples secuencias y captura al mismo tiempo la variabilidad de los datos en cada secuencia. El descriptor compacto es capaz de transformar la entrada multicanal en un único canal de entrada con mayor contraste y calidad que el que ofrece cada uno por separado. De esta manera, la dimensión del problema de segmentación no aumenta.

El método de segmentación propuesto permite una segmentación rápida y precisa. El método ha sido propuesto y desarrollado en 2D y 3D. El método ha sido validado mediante estudios reales y los resultados son comparados con respecto a una segmentación de referencia con ayuda de radiólogos expertos. Se han establecido nueve métricas para la valoración de la calidad de la segmentación.

Los resultados obtenidos son comparables, e incluso mejores en varios casos, a otros métodos existentes en el estado del arte y ponen de manifiesto la validez del uso de múltiples secuencias simultáneas de entrada para conseguir una segmentación más precisa, siendo además, muy competitivos en cuanto a tiempo de ejecución para su aplicación en la práctica clínica.

Esto ratifica que los métodos propuestos en el marco de esta tesis son válidos y demuestran su aplicabilidad y eficiencia en el proceso de segmentación de imagen médica, y en concreto, de segmentación hepática sobre imagen de resonancia magnética. Sin duda, la metodología diseñada, implementada y validada en esta tesis puede ser extendida a la segmentación de otros órganos en imágenes de MRI.

1.3. Contenido de la memoria

Este documento de tesis está distribuido en siete capítulos cuyo contenido se describe brevemente a continuación.

Capítulo 1. INTRODUCCIÓN. En este capítulo se hace una introducción al procesamiento de imagen médica y problemática existente. Se resume la trayectoria y motivación de la doctoranda. Se indican los objetivos de la tesis. Se describe el contenido de los capítulos de la memoria. Por último se enumeran las herramientas software empleadas para la implementación de los métodos propuestos en esta tesis.

Capítulo 2. ESTADO DEL ARTE: Imágenes médicas. Tecnologías de adquisición, limitaciones y retos. Caso de aplicación: segmentación hepática en MRI. En este capítulo se recoge el estado del arte en tratamiento de imágenes médicas, así como las limitaciones y retos existentes en el procesamiento de dichas imágenes. Se describen brevemente las tecnologías de adquisición más extendidas para centrarse principalmente en las tecnologías de CT y MRI, comparando las ventajas e inconvenientes de ambas, ya que son las predominantes para el diagnóstico de enfermedades hepáticas. La problemática concreta de la técnica de adquisición de resonancia magnética (MRI) es descrita en mayor detalle. Finalmente, se introduce el problema de segmentación hepática en imágenes de MRI y las dificultades existentes, lo que constituye el origen del problema que se aborda en esta tesis.

Capítulo 3. ESTADO DEL ARTE: Descriptores. En este capítulo se introduce el concepto de descriptor y se presentan los más relevantes relacionados con el problema de segmentación. Se muestran agrupados en base a si emplean el valor del nivel de gris, la forma, el valor de gradiente o la textura. Se referencian ejemplos en los que dichos descriptores han sido aplicados a segmentación de imagen médica en general, y a segmentación hepática en particular. Se añaden las limitaciones de los mismos y las conclusiones extraídas del análisis.

Capítulo 4. ESTADO DEL ARTE: Segmentación. En este capítulo se introduce el concepto de segmentación, así como la problemática inherente a la segmentación del hígado en imagen médica. Se describen brevemente los principales métodos de segmentación de la literatura que han sido aplicados a segmentación hepática. Los métodos de segmentación en 3D así como el tratamiento de las múltiples secuencias de MRI son tratados en detalle. Finalmente se indican las limitaciones de los métodos indicados para la segmentación hepática multisequencia.

Capítulo 5. APROXIMACIÓN TEÓRICA. En este capítulo se recoge la descripción teórica del método propuesto, tanto el modelo de contorno activo en 2D y 3D como el nuevo descriptor compacto para el manejo de la información multisequencia y que conduce a la generación de un mapa de distancia, o nueva secuencia mejorada. El método de minimización del contorno activo en 2D o superficie activa en 3D es descrito. El algoritmo completo de segmentación propuesto e implementado en el marco de esta tesis es descrito.

Capítulo 6. PRUEBAS Y RESULTADOS. Este capítulo recoge la metodología de pruebas establecida para la evaluación de la bondad del nuevo método de segmentación propuesto. Se describe el conjunto de datos disponible para la realización de las pruebas, la segmentación de referencia así como las métricas establecidas para la validación de la calidad de la segmentación. Se muestran los resultados de la segmentación obtenidos en base a dichas métricas. Se analizan los resultados y se discuten los factores que han podido influenciarlos.

Capítulo 7. APORTACIONES Y LÍNEAS FUTURAS. En este último capítulo se recogen las conclusiones extraídas a partir del desarrollo, pruebas y validación de los métodos propuestos. Se indican las aportaciones realizadas en esta tesis. Asimismo, se identifican posibles futuras líneas de trabajo que complementarían dichos desarrollos.

1.4. Herramientas empleadas y desarrollo SW para la tesis

El desarrollo de diversas pruebas preliminares y las comparaciones con métodos del estado del arte han sido desarrollados en Visual Studio C++ 2010 con el apoyo de ITK (versión 3.16.0) y VTK (versión 5.4.2). Los algoritmos numéricos que componen el núcleo de esta tesis (que se detallan en el capítulo 5) han sido desarrollados íntegramente en Matlab R2012b (versión 8.0).

Para la validación de los resultados, se ha desarrollado un interfaz de usuario avanzado en Visual Studio C++ 2010 que permite la visualización 3D de las diferentes secuencias, la reconstrucción 3D y las vistas de imágenes 2D en los tres planos axial, sagital y frontal. Adicionalmente, se ha desarrollado en C en el entorno de desarrollo LabWindows CVI 2012 una interfaz de usuario ad-hoc para la realización de la segmentación manual por parte del radiólogo, lo que constituirá la segmentación de referencia (*ground-truth* o *gold estándar*) con respecto a la cual se obtienen los resultados de la segmentación que se muestran en el capítulo 6.

*Capítulo 2: ESTADO DEL ARTE:
Imágenes médicas. Tecnologías de
adquisición, limitaciones y retos.
Caso de aplicación: segmentación
hepática en MRI*

El objetivo de este capítulo 2 es describir las líneas de trabajo que comprende la Imagen Médica, junto con las limitaciones y retos actuales. A continuación, se describen brevemente las tecnologías de adquisición más usadas en la práctica clínica, profundizando en los problemas presentes en la técnica de resonancia magnética (MRI), mediante la cual se han obtenido las imágenes con las que se trabaja en esta tesis. El procesamiento de la imagen médica se identifica como un área especial dentro del ámbito de la imagen médica. Finalmente, se presenta el caso concreto de aplicación de segmentación hepática sobre MRI, en el que se enmarcan los desarrollos realizados en esta tesis.

2.1. Imagen Médica

El término Imagen Médica, *Medical Imaging* en inglés, abarca distintas líneas de trabajo complementarias entre sí. Imagen Médica comprende el conjunto de técnicas específicas para la adquisición y reconstrucción de imagen por las múltiples tecnologías disponibles; la visualización de dichas imágenes; la compresión y almacenamiento de las mismas; la mejora de su calidad y contraste; las tareas de registro de imágenes junto con su procesamiento para la segmentación y extracción de información de dichas imágenes, como puede ser la cuantificación de parámetros anatómicos y fisiológicos. La integración de todas estas disciplinas permite abordar múltiples aplicaciones, como la ayuda en la planificación quirúrgica e intervenciones guiadas por imagen, la descripción de regiones anatómicas basadas en atlas, la visualización y análisis de órganos y tejidos, entre otras. Un repaso muy interesante de la evolución de las técnicas de análisis y procesamiento de imagen médica desde los años 80 puede encontrarse en (Duncan and Ayache, 2000).

La imagen médica no puede ser tratada de igual manera que otras imágenes de escenas naturales. Presenta una problemática específica, que varía dependiendo de la técnica mediante la que se ha generado la imagen. En general, se puede hablar de imágenes en niveles de gris, con ruido, poco contraste, bordes mal definidos y borrosos, y por tanto, aparecen problemas en la diferenciación de los órganos representados en la imagen. Si, además, se añade el hecho de que el principal objetivo de la imagen médica es el diagnóstico y el conocimiento de lo que sucede dentro del cuerpo humano, la precisión y exactitud de lo que se identifica y mide resulta de vital importancia. Las máquinas de adquisición de imágenes médicas y sus módulos software asociados usados por personal médico están sometidos a rigurosos controles, y para ser comercializados necesitan un certificado especial, que es otorgado por el organismo FDA (Food and Drug Administration) en Estados Unidos o la European Medicines Agency en Europa.

Si bien las técnicas de procesamiento de imagen se remontan a finales de los 70, su aplicación a imagen médica comienza en los años 80, suponiendo la inclusión de la tecnología de Resonancia Magnética o Magnetic Resonance Imaging (MRI) en el panorama un gran impulso que se produce aproximadamente a partir de 1985.

Un hito importante tiene lugar en 1991, en EEUU, con el proyecto Visible Human Project Male and Female (VHP) de la National Library of Medicine (NLM), cuyo objetivo era recopilar de manera minuciosa el mayor número de datos posible sobre el cuerpo humano.

La cantidad de información almacenada comenzaba a ser enorme y era necesario disponer de herramientas automatizadas para manejar y reconstruir la información, así como para visualizarla y tratarla con el objeto de dar soporte a las necesidades de la comunidad médica. El consorcio Insight Software se constituye en 1999, con el objetivo de crear un software gratuito capaz de visualizar, segmentar y registrar imágenes en 3D para analizar toda la información almacenada en la base de datos del VHP. A partir de esta iniciativa surge la comunidad ITK (Insight Toolkit) que hoy en día aporta la mayor librería de software libre disponible con una amplia funcionalidad para el manejo y procesamiento de imágenes médicas, cualquiera que sea su fuente, que usan miles de investigadores en todo el mundo. Esta librería es fácilmente integrable en varios entornos de programación y están implementadas en muchos módulos software comerciales. El número de organismos implicados en el desarrollo de ITK continúa siendo enorme en la actualidad, entre los que destacan por su reconocido prestigio la Clínica Mayo, la universidad de Harvard, Columbia, Utah, o el Imperial College de Londres.

Estos primeros pasos dieron lugar a abordar la problemática de la imagen médica de manera global, a realizar un diseño específico y establecer unos estándares para el manejo y el procesamiento de la imagen médica, que a su vez determina las características de configuración y adquisición de los sistemas hardware existentes. Como punto importante de partida, se definió un protocolo concreto de imagen, el estándar **DICOM** (Digital Imaging and Communication in Medicine), cuya primera versión relevante data de 1.993. Este es el estándar reconocido mundialmente para el intercambio de imágenes médicas. Ha sido diseñado para su manejo, almacenamiento, impresión y transmisión. Incluye la definición de un formato de fichero y de un protocolo de comunicación de red. El protocolo de comunicación es un protocolo de aplicación que usa TCP/IP para la comunicación entre sistemas. El fichero DICOM tiene dos grandes bloques: la cabecera y los datos. En la cabecera se almacena toda la información relativa al paciente y las condiciones de captura. En el bloque de datos se almacena la imagen. Los ficheros DICOM pueden intercambiarse entre dos entidades que tengan capacidad de recibir imágenes y datos de pacientes en este formato DICOM. Este estándar ha sido adoptado ampliamente por hospitales, donde las imágenes se almacenan y manejan en los denominados PACS, acrónimo de Picture Archiving and Communication System.

Una vez que las imágenes están disponibles es necesario procesarlas. El procesamiento de imagen aglutina toda o parte de la secuencia siguiente: enfatizado y mejora de la calidad de la imagen, extracción de características, segmentación, cuantificación y clasificación. En el caso de imagen médica pueden requerirse operaciones de registro si se trabaja con varias imágenes de la misma región física. En líneas generales, cada una de estas fases consiste en:

- Pre-procesamiento de la imagen. La calidad de una imagen médica es variable dependiendo de la técnica y máquina de adquisición. Pueden encontrarse determinados artefactos (característicos de cada tecnología), ruido no homogéneo, iluminación no uniforme, bordes difusos y bajo contraste entre los elementos presentes. Por ello, la extracción de información de la imagen se convierte en una tarea ardua y difícil. La

mejora de la calidad de una imagen consiste en abordar uno o varios problemas, como la eliminación del ruido, el aumento del contraste, el enfatizado de los bordes y de estructuras relevantes, la detección de puntos de interés, la uniformización de la iluminación, la eliminación de artefactos de adquisición, etc.

En el caso de las imágenes de resonancia magnética, los filtros genéricos no resultan suficientes para realizar un correcto enfatizado de las imágenes, ya que, aunque permiten una mejor visualización de la misma, eliminan muchos de los detalles en ellas contenidas. Es necesario abordar técnicas más avanzadas, filtros no lineales o filtros adaptativos, capaces de detectar las zonas de ruido y uniformizarlas, y al mismo tiempo, identificar los bordes y resaltarlos. De esta manera no se pierde nitidez en los límites de los órganos. Es lo que se denomina filtrado no homogéneo en todas las direcciones, como son los filtros de difusión anisotrópicos o los filtros de Wiener.

En las imágenes de resonancia magnética se dispone de un conjunto de imágenes 2D que constituyen un volumen 3D, lo que hace posible la aplicación de métodos de filtrado avanzados que tengan en cuenta los 3 ejes para mejorar la imagen. De la misma forma, en el caso de las imágenes de resonancia magnética, donde puede disponerse de varias capturas en el mismo estudio, así como en cualquier aplicación donde se combinan varias tecnologías de adquisición (MRI, CT, ultrasonidos...) el hecho de que un mismo punto esté asociado a varios valores, permitirá generar métodos de realce que tengan en cuenta la naturaleza multi-variable del mismo. El pre-procesamiento o enfatizado de la imagen constituye en sí misma una línea profunda de trabajo e investigación, que no es objeto de estudio en esta tesis, si bien puede incluirse en las líneas de trabajo futuras.

Algunas referencias de interés donde se puede encontrar un compendio de técnicas de mejora sobre imágenes de MRI son (Sonka et al., 1993), (Clarke et al., 1995), (Vovk et al., 2007), y para eliminación de ruido de la imagen se proponen las siguientes referencias (Bao and Zhang, 2003) (Buades et al., 2005), (Manjón et al., 2012).

- Extracción de características: Comprende el conjunto de técnicas para la identificación y extracción de características y descriptores relevantes de la imagen, que permitan representar y diferenciar unas zonas de otras. Estas características de la imagen dependen de los niveles de gris, de su distribución estadística (histogramas), de sus gradientes (bordes), y de la textura en la vecindad de cada píxel. El capítulo 3 está exclusivamente dedicado a este tema.
- Segmentación: Son técnicas que permiten separar diferentes elementos en una imagen. La segmentación de una imagen consiste en la división o separación de la imagen en regiones de atributos similares. Es decir, la segmentación subdivide una imagen en sus partes constituyentes u objetos. El grado de subdivisiones depende del problema a resolver, por lo que la segmentación deberá detenerse cuando los objetos de interés hayan sido aislados. La segmentación consiste en que a cada píxel de una imagen se le

asigna una etiqueta correspondiente a cada región. Todos los píxeles con la misma etiqueta (región) son similares en una serie de características: intensidad, color y/o textura. Los algoritmos para segmentar imágenes monocromáticas generalmente se basan en una de las dos propiedades básicas de los valores del nivel de gris: discontinuidad y similitud. En la primera categoría, se divide la imagen basándose en cambios abruptos en el nivel de gris. La segunda categoría se basa en segmentar por niveles de umbral, en regiones de crecimiento y en corte y pegado de regiones. La segmentación es uno de los grandes problemas de la visión artificial, ya que en general, no existe una solución global, y cada caso concreto debe ser abordado por separado.

Las técnicas de segmentación abarcan desde las más sencillas, basadas en diferencias en niveles de gris, hasta las más complejas basadas en textura, imágenes vectoriales y modelos deformables. Los métodos de segmentación son el elemento esencial de esta tesis, con lo que merecen un capítulo dedicado. El capítulo 4 contiene información detallada sobre las técnicas de segmentación más frecuentes y su aplicación específica al caso de segmentación hepática.

- **Clasificación:** Las técnicas de clasificación permiten, a partir de un conjunto de datos obtenidos de la región segmentada, la asignación de dicha región a un grupo u otro, lo que se denomina clase. Esta asignación a una clase, en el caso de imagen médica puede ser la identificación de un tipo de tejido o lesión. Además se pueden añadir operaciones de medida y cuantificación sobre la imagen para obtener datos más exactos, que sean de valor en el diagnóstico, la planificación de tareas de cirugía o en el tratamiento posterior.

En la presente tesis se van a abordar en detalle las tareas de extracción de características y los métodos de segmentación, en los capítulos 3 y 4, respectivamente. No se profundiza en la aplicación de técnicas de mejora de la calidad de la imagen o de técnicas de clasificación, que escapan del alcance de esta tesis.

2.2.Limitaciones y retos en procesamiento de imagen médica

Los algoritmos desarrollados en imagen médica no son aplicables y transferibles directamente a cualquier problema, sino que, por regla general, son desarrollados ad-hoc para la caracterización y segmentación de un órgano o patología específica, y también son dependientes de la tecnología de adquisición, dado que la misma condiciona las características y calidad de la imagen y por tanto, influye en el algoritmo. Los desarrollos existentes en el estado del arte son dedicados. La interpretación automática y universal está lejos de ser una realidad a día de hoy.

En el procesamiento de imagen médica, técnicamente hablando, el reto es triple:

- 1) realizar una segmentación precisa,
- 2) que dicha segmentación sea obtenida mediante un proceso automático,
- 3) y que dicho proceso automático se ejecute rápido.

En cualquier problema de segmentación el mayor reto es conseguir una segmentación precisa, lo que se traduce en que si se compara el resultado de una segmentación propuesta con una segmentación de referencia generada manualmente por un experto, las áreas de las imágenes 2D y por tanto, el volumen 3D final segmentado y reconstruido coincidan.

La segmentación perfecta implica que no existan fugas o sobre-segmentaciones y tampoco que la zona segmentada resultante sea inferior a la superficie real del órgano o elemento de análisis, lo cual, dadas las características de las imágenes, la similitud en los niveles de gris entre los órganos adyacentes, los bordes difusos, y el bajo contraste existente entre algunas estructuras, hacen que sea un problema complejo. De hecho, en el proceso de generación de la segmentación de referencia (lo que se conoce como *gold standard* o *ground truth*), los propios expertos no se ponen de acuerdo, y a veces, incluso el mismo experto en distintos momentos de tiempo puede segmentar el mismo órgano de manera diferente (Göçeri et al., 2015).

Existen tres tipos de segmentación en base al grado de intervención del usuario: manual, semiautomático y automático. En una **segmentación manual**, el usuario, normalmente un radiólogo experto, realiza la totalidad de la segmentación, es decir, existen unas herramientas de anotación y dibujo que posibilitan que esta persona pueda ir seleccionando las áreas y puntos de interés. Por ejemplo, en este caso, se asigna hígado o no hígado. Estas regiones contenidas en los contornos marcados al respecto son etiquetadas como pertenecientes a determinada clase.

En los **métodos semiautomáticos** existe una parte que es realizada por el ordenador y otra en la que el usuario interactúa con el sistema. Es característico de algunas técnicas de segmentación, como son los modelos estadísticos de forma, indicar los puntos de interés o aquellos más significativos. El usuario debe seleccionarlos para que el sistema pueda computar los valores asociados. También son frecuentes los métodos semiautomáticos para la selección de semillas en métodos de agrupamiento o clustering. Otras veces en estos métodos, la interacción del usuario es requerida al final, cuando ya se han propuesto los resultados de una segmentación automática y se pide al usuario que la valide y refine.

En los **métodos automáticos** de segmentación, el usuario no hace absolutamente nada, ni aporta ninguna información al sistema. Una vez cargadas las imágenes, el proceso de segmentación se desencadena automáticamente y presenta unos resultados. Evidentemente, desarrollar un método que realice una segmentación de manera automática y satisfactoria es más complejo.

Otro de los retos en la segmentación de imagen médica en general, y hepática en particular, es el diseño e implementación de algoritmos relativamente rápidos. El tiempo real está muy lejos de poder ser aplicado en segmentaciones de grandes volúmenes de datos como los de las imágenes médicas a día de hoy, pero debe perseguirse algo razonable que no implique horas y horas de procesamiento, para que a corto o medio plazo, puedan ser incluidos en la práctica clínica.

Una limitación importante en la segmentación de la imagen médica es la dificultad para comparar la bondad de los resultados. Para evaluar cómo de bueno es un método con respecto a otro, es necesario, como se ha indicado en el párrafo anterior, un conjunto de datos de prueba y validación adecuados. El *gold standard* o *ground truth* se entiende como la referencia absoluta con la que comparar la segmentación realizada y se calculan unas métricas o indicadores de la magnitud del error con respecto a ella. Ninguno de ambos elementos está estandarizado a día de hoy en la mayor parte de los casos. No existen bases de datos públicas con imágenes de los diferentes órganos y elementos captadas por las diferentes tecnologías de adquisición, y además, anotadas, de manera que se pueda disponer de una segmentación de referencia. Cada estudio y trabajo debe agenciarse esta referencia por los medios de que se disponga, normalmente, mediante trabajo conjunto con radiólogos o especialistas que amablemente anoten las imágenes para generar esta referencia. Además, las métricas existentes para valorar la calidad de la segmentación son múltiples y diversas, atendiendo cada una a una medida de error, y no existe una metodología determinada extendida que indique cuáles usar. Cada investigador decide cuáles le parecen más adecuadas para sus propósitos y muestra los resultados en relación a ellas. Como consecuencia, la comparación entre diversos estudios se torna difícil ya que ni la segmentación de referencia ni las métricas son iguales, salvo en casos excepcionales, lo que dificulta mucho la validación de los nuevos métodos propuestos. Sería necesario habilitar bases de datos de pruebas, y una metodología de validación específica para cada caso.

En el caso de segmentación hepática sobre MRI no existen conjuntos de datos con su segmentación de referencia y unas métricas establecidas, sino que cada uno aplica su propio criterio. Hay algunas metodologías implantadas, como la que se estableció en MICCAI Challenge 2007, con un conjunto adoptado de 5 métricas y un conjunto de test de validación que son un compendio de imágenes obtenidas mediante CT, no sobre MRI.

En esta tesis se ha preparado un conjunto de datos propio de prueba y se han implementado y calculado un conjunto de 9 métricas muy extendidas en la literatura.

Otro de los retos a futuro es el trabajo conjunto de los desarrolladores de hardware, las máquinas de adquisición de imagen propiamente dichas; y los desarrolladores software, de manera que se puedan optimizar las funciones, y secuencias (en el caso de MRI) para caracterizar mejor y discriminar los diferentes tipos de tejidos, sanos y tumorales, objetos de interés. En la medida en que se trabaje conjuntamente, la extracción de características, segmentación y medida permitirá obtener mejores resultados.

Finalmente, otro de los retos es que el fin último del compendio de técnicas de procesamiento, segmentación y análisis de imagen médica debiera ser la inclusión futura de las mismas en software para el uso de los radiólogos, de manera que les ayudaran en la práctica clínica diaria. Por ello, deben perseguir ser útiles y obtener información valiosa para el diagnóstico y cuantificación posterior, así como la asistencia en la intervención quirúrgica y su correspondiente planificación. Desde las fases tempranas de desarrollo es necesario el trabajo conjunto con personal médico para dar respuesta real a las necesidades de estos usuarios.

2.3. Tecnologías de adquisición de imagen médica

Existen diversas tecnologías de adquisición de imagen médica, desde la obtención de imágenes de rayos X hasta llegar a los modernos dispositivos multimodales.

Los rayos X (también denominados rayos Röntgen en Europa central y del este) fueron descubiertos en 1895 por Wilhelm Conrad Röntgen, que recibió el premio Nobel de Física en 1901. Estas primeras investigaciones condujeron al desarrollo de varias tecnologías de adquisición de imagen médica, lo que permite elegir la técnica más adecuada en cada ocasión en base al propósito de estudio.

Las máquinas que obtienen hoy en día las imágenes médicas proveen datos anatómicos en dos y tres dimensiones. Estos avances en el hardware de adquisición de imágenes ofrecen un gran potencial para mejorar la precisión del diagnóstico, prevenir el riesgo de padecer algunas enfermedades y seleccionar las terapias más adecuadas. Junto con dichas máquinas, las empresas comercializadoras suelen proporcionar un SW con funcionalidad básica que permite visualizar las imágenes en 2D y a veces, en 3D, y realizar un procesamiento de las mismas más o menos avanzado. En definitiva, las nuevas tecnologías médicas permiten obtener mejores resultados en el diagnóstico de los pacientes, lo cual influye en la eficacia del tratamiento recomendado y en la eficiencia del sistema sanitario, con la consiguiente reducción de los costes asociados.

Existen varias tecnologías de adquisición de las imágenes. Para el estudio del hígado, las dos técnicas más extendidas son CT y MRI. Su esquema básico de funcionamiento se expone en los siguientes apartados. Se mencionan algunas otras tecnologías a modo ilustrativo, pero que no aplican en el marco de la presente tesis.

2.3.1. Tomografía Axial Computerizada (TAC o CT)

La Tomografía Axial Computerizada (TAC) o más conocida por sus siglas en inglés (CT) que se corresponden con *Computerized Tomography*, o simplemente escáner, es una tecnología sanitaria de exploración por rayos X que produce imágenes detalladas de cortes axiales del cuerpo.

Los primeros estudios cotejados clínicamente de pacientes datan de principios de los años 70 en el hospital general de Massachusetts (KISTLER et al., 1975) sobre imágenes de cerebro, y debido a la bondad de los resultados de validación, a partir de ese momento, su extensión y aplicación a todo tipo de órganos y patologías fue inevitable.

Tomografía viene del griego *tomos* que significa corte o sección y de *grafía* que significa representación gráfica. Por tanto, tomografía es la obtención de imágenes de cortes o secciones del cuerpo.

En lugar de obtener una imagen como la radiografía convencional, la tomografía obtiene múltiples imágenes alrededor del cuerpo. La respuesta de los órganos y tejidos humanos es recogida por los sensores de la máquina, y mediante la intervención de un ordenador

("computerized") cada corte se reconstruye como un mapa de grises en una imagen 2D. El conjunto de todos los cortes resultantes de un barrido en el eje axial (plano axial es aquel que es perpendicular al eje longitudinal de un cuerpo), constituye el volumen 3D adquirido como resultado de la inspección, y es el conjunto de datos que se deben analizar para que el radiólogo pueda extraer conclusiones. Los píxeles representan la información en las imágenes 2D. La información en 3D se recoge en los denominados vóxeles.

Esta técnica de adquisición tiene un inconveniente y es que siempre existe la leve posibilidad de provocar un cáncer como consecuencia de la exposición excesiva a la radiación. Esta es una desventaja con respecto a la resonancia magnética.

2.3.2. Resonancia Magnética Nuclear (RMN o MRI)

La Resonancia Magnética Nuclear (RMN), o sus siglas en inglés, NMR, en los comienzos, que corresponden a Nuclear Magnetic Resonance, fue sustituida por Magnetic Resonance Imaging (MRI), que es la terminología más extendida hoy en día. Se trata de una técnica radiológica no invasiva, e indolora, que permite obtener información sobre la estructura interna del cuerpo. Para la obtención de dicha información, se utiliza el fenómeno de la resonancia magnética, de tal forma que se evita el uso de radiaciones ionizantes, como es el caso de CT. El cuerpo inspeccionado no es sometido a ningún tipo de radiación potencialmente perjudicial para la salud.

El físico Damadian demostró también que dos tejidos, uno sano y otro maligno se comportaban en vivo de manera diferente bajo la tecnología MRI (entonces conocida como Resonancia Magnética Nuclear), de manera que esta tecnología podía ser empleada con fines diagnósticos (Damadian, 1971). La primera máquina de MRI para escanear el cuerpo humano la construyó él y fue patentada. Sin embargo, sus métodos de interpretación no eran adecuados tal y como se conocen ahora. En 1976, Mansfield y Lauterbur propusieron la técnica de MRI en la manera en que se emplea actualmente. Lauterbur publicó los fundamentos de reconstrucción de una imagen 2D y 3D basadas en gradientes mediante MRI (LAUTERBUR, 1973); y Mansfield diseñó un método matemático para realizar un escaneado en segundos con imágenes de más calidad que las existentes, en detrimento de un tiempo de proceso de varias horas como necesitaban las imágenes generadas por Damadian. En 2003, Mansfield y Lauterbur recibieron el Nobel de Medicina por su aportación de la técnica de Magnetic Resonance Imaging. Hubo bastante controversia dado que Damadian también fue considerado por muchos el precursor de la tecnología mientras que los laureados refinaron la técnica.

Sea como fuere, MRI se ha convertido a lo largo de los años en una tecnología con un potencial tremendo para el diagnóstico por imagen, y es usada en el análisis de cualquier parte del cuerpo.

El uso de la región de radiofrecuencia en el espectro electromagnético para generar una imagen es bastante sorprendente e inimaginable hace 50 años. En una imagen visible, la radiación es reflejada por el cuerpo. En una imagen generada mediante rayos X, la radiación

es transmitida a través del cuerpo. Sin embargo, en una imagen de MRI, la persona es estimulada de manera que una radiación electromagnética es emitida por el cuerpo. Con esta base física y mediante métodos muy avanzados, la señal emitida es receptionada, interpretada y manejada hábilmente hasta poder generar una imagen con información relevante.

Las imágenes de resonancia dependen de una serie de parámetros del campo electromagnético y de la respuesta emitida por el cuerpo que varía según el tejido. Cada una de las diferentes secuencias que se emplean tiene una selección de patrones de tiempo para los pulsos de radiofrecuencia y los gradientes del campo electromagnético para optimizar el contraste entre los diferentes elementos anatómicos presentes en las imágenes (Bushong, 2003). Para generar las imágenes, el sistema somete al paciente a un campo electromagnético de determinado valor, como son 1.5 teslas en las máquinas más extendidas en la mayoría de hospitales a día de hoy, aunque se empiezan a incorporar máquinas con resonador de 3 teslas. La diferencia en el número de teslas del campo magnético influye en la precisión de la información que la imagen puede proporcionar, de manera que se puede apreciar mayor resolución de píxel con máquinas de 3 teslas que las de 1.5 teslas no son capaces de proporcionar.



Figura 2.1. Imagen de un escáner de MRI Abanto de Siemens
[\(https://www.healthcare.siemens.com/\)](https://www.healthcare.siemens.com/)

El principio físico de funcionamiento es el siguiente. El campo electromagnético es generado por un gran imán que atrae a los átomos, principalmente de hidrógeno, contenidos en el agua que conforma los tejidos humanos. De esta forma, al ser estimulados por la radiación electromagnética (a una determinada frecuencia de resonancia), salen de su alineamiento normal y modifican su momento magnético. Cuando el estímulo se suspende, los átomos regresan a su posición original, liberando energía en forma de ondas de radiofrecuencia. Los diferentes tipos de tejidos devuelven señales diferentes. La señal emitida por cada átomo es capturada por un conjunto de bobinas. Esta información es procesada por un ordenador, que reconstruye las señales de respuesta recibidas en un mapa

de niveles de gris, de manera que es posible generar imágenes detalladas de los órganos, tejidos blandos, huesos, y prácticamente el resto de las estructuras internas del cuerpo. La resolución de la imagen, en cuanto a precisión en la información proporcionada depende del número de bits que representen la información contenida en cada píxel. De este modo, un píxel representado por 8 bits sólo podrá presentar 256 valores de gris (2^8). Con 12 bits se podrán representar 4.096 niveles de gris (2^{12}) y en el caso de emplear 16 bits, se podrán mapear hasta 65.536 niveles (2^{16}), con lo que podrán apreciarse diferentes matices entre imágenes, como se puede apreciar en la figura 2.2.

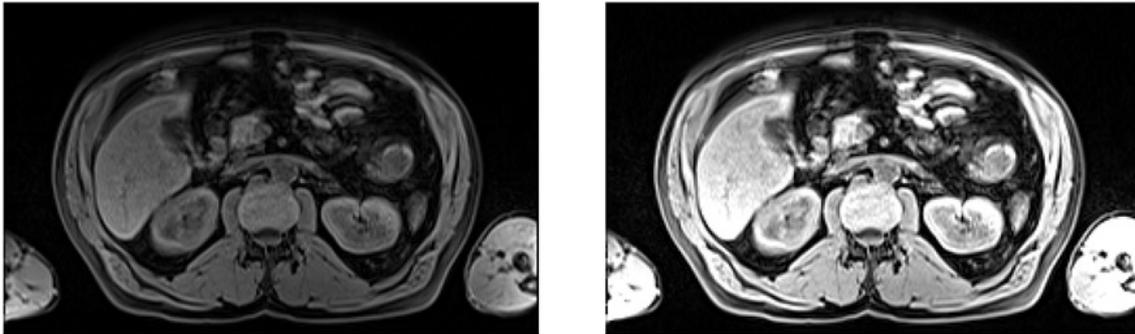


Figura 2.2. Comparativa imagen MRI almacenada con 12 bits (izda) y 8 bits (dcha)

Las imágenes obtenidas mediante esta tecnología no son de mucha resolución, pudiendo disponer de imágenes de 160 x 256 píxeles, por ejemplo, o incluso menos. La resolución va a depender de la configuración introducida y de la zona del cuerpo que se desee visualizar. La resolución espacial a la que se corresponde dicho número de píxeles en su representación del cuerpo humano es de 1 o 2 mm/píxel en el eje x e y de los planos axiales (cortes transversales del individuo). El eje longitudinal, el eje z de la representación puede presentar diversos espaciados, hasta de varios mm, con lo que se pierde información.

El fundamento de la mayor información que alcanza la MRI es la utilización de unas secuencias de pulsos de radiofrecuencia determinadas, durante la aplicación de los campos electromagnéticos. Hay parámetros que pueden ser configurados, como son el tiempo de eco, tiempo de repetición o tiempo de relajación entre otros. De la modificación y combinación de determinados valores de los parámetros configurables, que permiten modificar el campo electromagnético generado sobre el cuerpo, se puede establecer un conjunto de secuencias diferentes. Existen numerosos tipos de secuencias, cada una optimizada para obtener un contraste determinado en la imagen. Cada fabricante proporciona una serie de secuencias diseñadas para un funcionamiento óptimo, siendo algunas comunes a todas las máquinas y muy extendidas en la práctica clínica por la información valiosa que proporcionan.

MRI es una tecnología flexible que permite modificar el contraste de la imagen mediante el uso de diferentes secuencias de pulsos y cambiando los parámetros de la imagen correspondientes al tiempo de relajación longitudinal (T1), tiempo de relajación transversal (T2), y las intensidades de la señal en imágenes ponderadas T1 y T2 en relación a características específicas de los tejidos. El contraste de la imagen depende de parámetros

en la secuencia de pulsos. Las secuencias más comunes son las denominadas T1 y T2. Mediante este escaneado, las zonas que contienen fluidos, especialmente agua, son más brillantes; mientras que las zonas que contienen grasa son más oscuras.

Los radiólogos expertos tienen identificado el comportamiento de las patologías en cada una de las secuencias, con lo que la realización del diagnóstico puede ser más sencilla y fiable. Este es otro de los grandes potenciales que presenta MRI frente a CT, donde sólo se obtiene una secuencia en cada estudio. Este aspecto será abordado y explotado en el marco de la presente tesis.

2.3.2.1. Comparativa MRI y CT

Las dos grandes tecnologías más empleadas para la adquisición de imagen médica que sirven de herramienta de ayuda al diagnóstico a miles de facultativos en todo el mundo son CT y MRI. Ambas están muy extendidas. Existen muchos más estudios de procesamiento sobre imágenes de CT en la literatura, al menos en el caso analizado de segmentación de hígado. Esto es debido a que esta tecnología apareció antes (principios de los 70) y a que una máquina de CT es mucho más barata que una de MRI, con lo que más hospitales pueden disponer de ella. Sin embargo, la tecnología MRI ha ido tomando fuerza poco a poco debido a sus múltiples ventajas. Se enumeran a continuación las principales ventajas e inconvenientes de las tecnologías de adquisición MRI y CT (Duncan and Ayache, 2000), (Sharma and Aggarwal, 2010), (James and Dasarathy, 2014).

Las principales ventajas de MRI son:

- Representa muy bien los tejidos blandos, proporcionando muchos matices y una información muy rica.
- Posibilita una alta resolución de imagen, de hasta 1 mm cúbico en vóxeles.
- Presenta un ratio señal ruido elevado.
- Permite la adquisición multisequencia modificando las características de captura. El uso combinado de las secuencias facilita las tareas de segmentación y análisis.

Las principales desventajas de MRI son:

- La adquisición de una secuencia tarda más tiempo que con CT, al menos 45-60 segundos. Por tanto, la posibilidad de realizar la adquisición de una secuencia en apnea es muy reducida o prácticamente inexistente. Este hecho implica que en las imágenes se puedan apreciar desplazamientos debidos al movimiento respiratorio involuntario u otros movimientos.
- Si se usan varias secuencias para analizar la misma zona física se añaden problemas de registro que dificultan la segmentación y clasificación posterior; el registro consiste en asegurar que cada píxel en cada imagen de cada secuencia se corresponde con el mismo punto físico en todas las secuencias disponibles.
- Es difícil obtener una imagen de calidad, y de iluminación uniforme.
- Existe más presencia de ruido.

Por contra, la tecnología CT presenta las siguientes ventajas:

- Está más extendida en hospitales, ya que es más barata y por tanto, es más fácil disponer de una máquina.
- Los últimos modelos ofrecen alta resolución espacial.
- El tiempo de realización de un estudio es corto, 15-20 segundos aproximadamente, con lo que existe posibilidad de realizarlo en apnea.
- En algunos casos concretos, como en el análisis de huesos, músculos, o en la detección de calcificaciones intra-craneales, presenta una sensibilidad muy elevada, pero no en el caso de tejidos blandos como el hígado.

Las desventajas de emplear la tecnología CT son:

- Ofrece menor contraste entre tejidos blandos, y menor precisión en la información en comparación a la que proporciona MRI.
- Implica una exposición a la radiación (rayos X), con su potencial riesgo de provocar cáncer.

En la siguiente tabla se muestra una comparativa de las prestaciones de ambas tecnologías:

	CT	MRI
Tiempo adquisición	15-20 s. Muy corto	45-60 s. Largo
Radiación para el paciente	Existe radiación ionizante	Sin peligro
Calidad de información de las imágenes	Calidad buena	Calidad muy buena, mucha precisión en la información por píxel. Es posible distinguir pequeños matices entre tejidos.
	Bordes nítidos, gran contraste entre elementos.	Bajo contraste entre diferentes tejidos, bordes difusos, no uniformidad en la imagen.
Ruido	Existe	Existe ruido y hay presencia de artefactos.
Movimientos en imágenes	Pocos	Muchos
Posibilidad de obtener múltiples secuencias del mismo tejido	No es posible	Es posible disponer de información múltiple de la misma zona.

		Surge el problema de registro de imágenes de diferentes secuencias para poder ser usadas de manera combinada
--	--	--

Tabla 2.1. Comparativa de las características de las imágenes CT y MRI

Se muestran a modo de ejemplo una imagen de CT y otra de MRI en la que se pueden visualizar algunas de las características mencionadas para cada una de estas tecnologías.

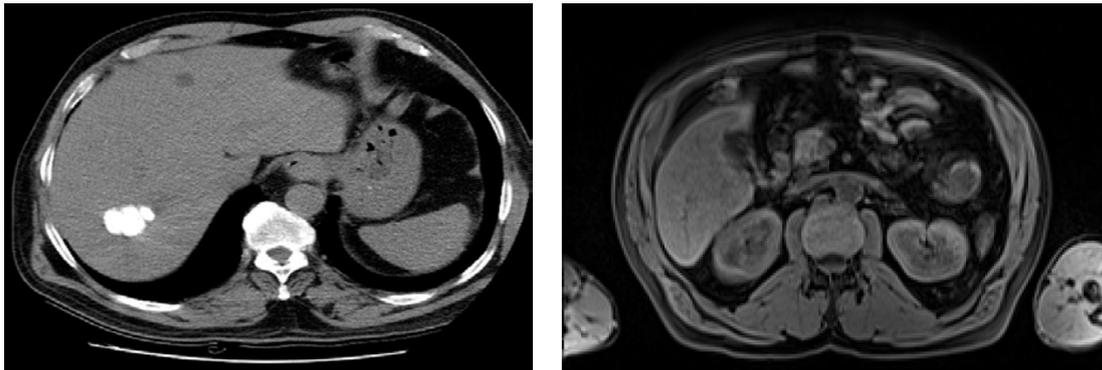


Figura 2.3. Comparativa imagen abdominal de CT (izda) y MRI (dcha)

Se puede apreciar en la imagen que los matices disponibles son mucho más ricos en las imágenes de MRI si bien presentan bordes más difusos y menor nitidez y contraste entre los elementos presentes.

El uso de varias secuencias simultáneas de MRI para la segmentación se abordará en esta tesis, ya que la interpretación combinada de varias secuencias es una práctica clínica extendida y eficaz en la caracterización de los órganos blandos, como se ha explicado anteriormente.

2.3.2.2. Problemas y Artefactos en imágenes de MRI

Las imágenes adquiridas mediante tecnología MRI (al igual que las de CT) presentan una serie de problemas y artefactos que se describen a continuación (Pham et al., 1998) (Sharma and Aggarwal, 2010).

Uno de los problemas más notables es el efecto del volumen parcial. En una imagen de MRI, cada píxel (o vóxel en 3D) representa la señal devuelta por un tipo de tejido. Si un píxel representa la información de un solo tejido, presenta un nivel de intensidad único característico. Cuando hay varios tejidos próximos, es posible que la señal devuelta en respuesta al campo electromagnético sea la combinación de las señales devueltas por varios tejidos diferentes adyacentes. El problema del volumen parcial consiste en que existe una zona difusa con bordes no nítidos en la que un vóxel (o más) representa la señal combinada

devuelta por varios tejidos, de manera que es difícil determinar dónde está el límite entre ellos. La imagen aparece como si estuviera desenfocada. Dependiendo del caso de aplicación, es relevante la técnica de segmentación que se utiliza y el tratamiento que se le da al efecto de volumen parcial, porque la asignación de unos píxeles a una clase u otra puede derivar en conclusiones e implicaciones totalmente dispares. El volumen parcial ha intentado ser cuantificado en algunos trabajos para corregir los resultados de una segmentación (González Ballester et al., 2002). Este efecto se puede apreciar en la imagen mostrada en la figura 2.4.

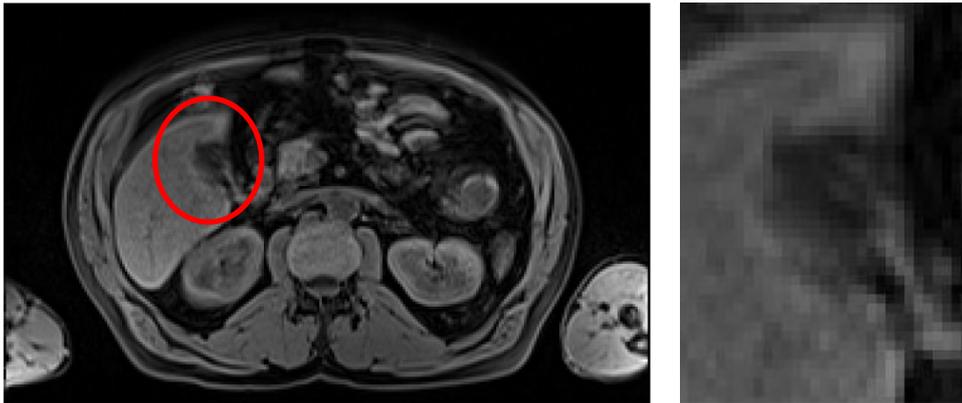


Figura 2.4. Efecto del volumen parcial en una región de la imagen de MRI (izda) y dicha región aumentada (dcha).

En las imágenes de MRI, este hecho provoca que los bordes no sean nítidos, sino bastante difusos, y sea difícil determinar dónde empieza una estructura y termina otra. En la figura 2.4 se aprecia que la delimitación de lo que es hígado respecto del fondo no es sencilla ya que hay varios píxeles que representan varias estructuras.

Otro tipo de artefactos que aparecen en las imágenes son los artefactos de movimiento, o los artefactos de anillo, mostrado en la figura 2.5. Se precisa de la aplicación de técnicas avanzadas para la eliminación de dichos problemas.

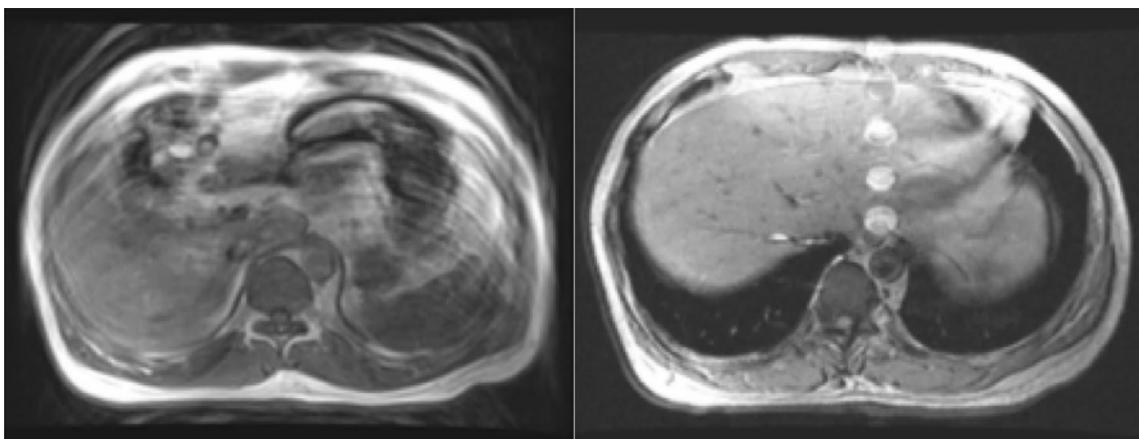


Figura 2.5. Efecto de movimiento como artefacto de MRI.(Rosenkrantz et al., 2010)

La presencia de ruido debida a los sensores y al conjunto electrónico relacionado también es un problema reseñable. En MRI, se identifica el denominado “*bias field noise*”, causado por in-homogeneidades en el campo electromagnético generado, lo que se traduce en zonas de iluminación no uniforme en la imagen. La aparición de zonas de iluminación no uniforme puede ser debido a varias causas solapadas, constituyendo este artefacto una de ellas. Este efecto dificulta la segmentación e implica la necesidad de aplicación de técnicas de pre-procesamiento para uniformizar la imagen o el desarrollo de algoritmos complejos capaces de dar respuesta a este problema. Se muestra en la figura 2.6 un ejemplo. Se aprecia que la parte superior de la imagen abdominal es más oscura que la parte inferior.



Figura 2.6. Artefacto de ruido bias-field

Se pueden revisar en (Vovk et al., 2007) un compendio de los métodos propuestos para corregir este artefacto y las no homogeneidades en las imágenes de MRI, en general. En este trabajo se describen técnicas de pre-procesamiento de imagen, lo cual constituye una línea de investigación en sí misma.

Como se ha indicado previamente, los algoritmos diseñados para el procesamiento de una imagen médica dependen fuertemente de la modalidad de adquisición elegida, del órgano del cuerpo estudiado y de la aplicación concreta, es decir, del tipo de información que se desee extraer en cada caso. Hay que tener en cuenta los artefactos y problemas que se presentan a la hora de diseñar el algoritmo de procesamiento específico (Sharma and Aggarwal, 2010). Por ejemplo, el efecto de volumen parcial es más habitual en imágenes de cerebro, mientras que en imágenes abdominales aparece el artefacto de movimiento y sobre todo, la iluminación no uniforme en la totalidad de la imagen, lo que dificulta mucho los procesos de segmentación automática. No es posible, a día de hoy, un algoritmo único para la segmentación automática de imágenes médicas, cualquiera que sea su técnica de captura u órgano de análisis.

Además de todos los problemas mencionados, cabe recordar que a la hora de procesar e interpretar las imágenes, una de las grandes dificultades es que tejidos diferentes presentan un valor de nivel de gris similar e incluso igual, con lo que la segmentación automática se convierte en un verdadero reto y precisa de técnicas avanzadas.

2.3.3. Otras tecnologías: Ultrasonido (US), PET, combinadas

Existen otras tecnologías para la adquisición de imágenes médicas. Una de ellas es la técnica de **ultrasonidos**. El ultrasonido es una onda sonora cuya frecuencia está por encima de la que registra el oído humano. Esta tecnología basa su funcionamiento en que las ondas de ultrasonidos de alta frecuencia son emitidas y se reflejan en las estructuras del cuerpo humano, y dependiendo de la distancia, son recogidas en momentos temporales diferentes, con lo que es posible reconstruir la imagen como un mapa de distancias. Las técnicas de ultrasonido son ampliamente utilizadas en diagnóstico por imagen, ya que no suponen ningún riesgo ni emiten radiaciones ionizantes como es el caso de los rayos X o CT. La frecuencia de emisión del ultrasonido varía dependiendo de la estructura del cuerpo que se quiera analizar. En general, a menor frecuencia de emisión, más penetración, y una mayor frecuencia conlleva menor penetración.

Las imágenes de ultrasonidos, obtenidas por los ecógrafos, se caracterizan, en general, por ser muy ruidosas, y es difícil la identificación de los elementos presentes salvo que se sea un experto. La figura 2.7 muestra una imagen de hígado obtenida mediante ultrasonido.



Figura 2.7. Imagen de hígado adquirida mediante ultrasonidos (<http://www.healthcare.philips.com>)

Otra tecnología extendida de adquisición de imagen médica es la tomografía por emisión de positrones o más conocida por sus siglas en inglés, PET, *Positron Emission Tomography*. Se trata de una técnica no invasiva usada para el diagnóstico, al igual que CT y MRI, cuyo objetivo es medir la actividad metabólica del cuerpo. Pertenece a la denominada medicina nuclear. En este caso, la tecnología PET detecta y muestra la distribución en el cuerpo de un radiofármaco específico que es administrado previamente a la prueba por inyección intravenosa. Existen diferentes tipos de fármacos dependiendo del objetivo de la prueba. Los sensores del tomógrafo recogen los fotones gamma emitidos por el paciente. El principio físico en que se basa es que estos fotones gamma son el producto de una aniquilación entre un positrón, emitido por el radiofármaco, y un electrón del cuerpo del paciente. Esta aniquilación da lugar a la emisión, fundamentalmente, de dos fotones. Para que estos fotones acaben por conformar la imagen deben detectarse en coincidencia, es decir, al mismo tiempo; en una ventana de tiempo adecuada (nanosegundos), deben provenir de la misma dirección y sentidos opuestos, pero además su energía debe superar un

umbral mínimo que certifique que no ha sufrido dispersiones energéticas de importancia en su trayecto hasta los detectores (fenómeno de dispersión o *scatter*). Los detectores de un tomógrafo PET están dispuestos en anillo alrededor del paciente, y conforman la imagen cuando detectan en coincidencia a los fotones generados en cada aniquilación. Para la generación de la imagen estos fotones detectados son convertidos en señales eléctricas. Esta información se somete posteriormente a procesos de filtrado y reconstrucción para obtener la imagen. Esta tecnología PET se está implantando con fuerza en oncología. En la figura 2.8 se muestra un ejemplo de cómo la metástasis presente en un hígado enfermo es resaltada mediante la tecnología PET, donde la tecnología CT es limitada.

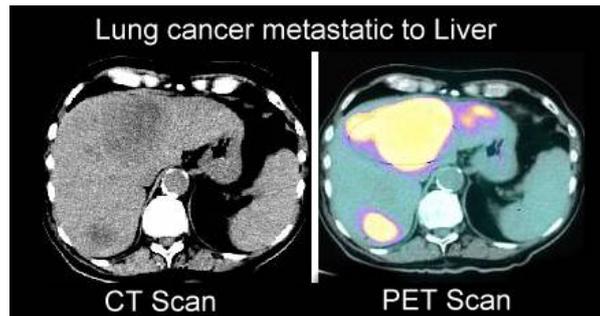


Figura 2.8. Comparativa imagen CT y PET donde se enfatiza la patología de metástasis (www.aboutcancer.com)

En los últimos años también ha tomado fuerza la posibilidad de aunar diferentes tecnologías de adquisición para poder contar con las ventajas de una y de otra simultáneamente. Este es el caso de PET-MRI que aúna la tecnología MRI para análisis morfológico de tejido blando y tecnología funcional de PET. Pocas empresas fabrican las máquinas capaces de componer este tipo de imagen, siendo Philips, Siemens y GE las que dominan el sector. También existe la tecnología híbrida PET-CT. El estado del arte de estas técnicas y algunas experiencias de aplicación se recogen en (James and Dasarathy, 2014).

2.4. Sistemas comerciales existentes para adquisición y procesamiento de imagen médica

Las máquinas de resonancia magnética son diseñadas y fabricadas por gigantes de la tecnología médica, como son Philips, Siemens y GE Healthcare. El mercado está dominado por unos pocos. Estas grandes empresas, junto con la tecnología de adquisición, proporcionan algunas herramientas SW con cierta funcionalidad para realizar una serie de acciones.

Philips es uno de los grandes proveedores de equipamiento médico para la adquisición de imágenes del cuerpo humano. Esta empresa es la que ha fabricado el primer equipo mencionado anteriormente que funciona fusionando las técnicas PET-MRI. A día de hoy, Philips provee una variedad enorme de equipamiento en todas las tecnologías existentes. En el caso de la resonancia magnética, junto con el equipo de adquisición, proporciona el software SmartExam. Este software posibilita el manejo de las imágenes de resonancia. Las

tareas que dicho software permite realizar son la reconstrucción del volumen 3D, y la visualización de la información en distintos planos, así como el establecimiento de unos puntos de referencia para seguimiento de las patologías.

Siemens es otro grande de la fabricación de equipamiento médico. De hecho, se ha contado en esta tesis con su equipo MAGNETOM Avanto de 1.5 teslas para la realización de resonancias, en las instalaciones que CVSS Radiología Clínica tiene en IMQ. Dispone de un software específico para el manejo de la información adquirida. El alcance del software es similar al de Philips. Existe la posibilidad de adquirir paquetes complementarios del software como REVEAL para la mejora de la calidad de imágenes abdominales, y ayuda en la detección; o Tissue4D, que enfatiza y contrasta las imágenes abdominales para facilitar la localización de tumores.

Los sistemas comerciales actuales ofrecen herramientas para la segmentación manual de las lesiones. Sin embargo, el volumen de información manejada, el consumo de tiempo del especialista y la tarea repetitiva justifica las técnicas de segmentación automática.

En general, las acciones que los radiólogos pueden realizar, mediante las herramientas software que son proporcionadas junto con el escáner de adquisición, están relacionadas con la modificación de ciertos parámetros de configuración de la adquisición de los escáneres (ajuste del campo electromagnético generado) y visualización. El módulo de visualización permite la navegación entre las imágenes 2D que constituyen el volumen completo de las múltiples secuencias adquiridas, y su vista en los diferentes planos, axial, frontal y sagital. También es posible obtener ciertas medidas adicionales, como el volumen existente en determinada región seleccionada, el valor de nivel de gris de una determinada zona, etc. Asimismo, es posible mejorar ciertas características de la imagen mediante la aplicación de un filtrado para la eliminación de ruido o mejora del contraste. Dichos módulos no ofrecen funcionalidades de segmentación.

Por otra parte están los CADs (Computer Aided Diagnosis) que son módulos que permiten localizar y clasificar unas patologías determinadas. Esos módulos son específicos de la patología y no suelen ser generalistas, necesitando cada órgano y patología un entrenamiento previo y desarrollo ad-hoc, debido a la complejidad del problema.

Dado el interés y la importancia del análisis y el procesamiento de la imagen médica, existen muchos grupos trabajando en diferentes órganos, como son cerebro, corazón, pulmón...y también el hígado. Estos grupos han desarrollado algunos módulos software que pretenden ser generalistas en la aplicación de transformaciones y algoritmos para realizar labores de segmentación y clasificación, así como el propio procesamiento previo de la imagen para mejorar su calidad. Si bien las técnicas de pre-procesamiento pueden ser aplicables a todas las disciplinas, la segmentación y tratamiento de cada órgano y sus patologías asociadas es una tarea altamente específica donde unos algoritmos funcionan mejor en unos casos que en otros. Esta especificidad y la dificultad de encontrar unos algoritmos generales se ve refrendada por la abundante literatura existente con trabajos específicos por órganos e incluso por tecnologías de adquisición para el mismo órgano. Así,

un algoritmo diseñado para segmentación de hígado, por ejemplo, no tiene por qué proporcionar necesariamente los mismos resultados si se trata de imágenes de CT o de MRI, dado que la calidad de las imágenes no es la misma, y los artefactos o problemas en la imagen tampoco. Las publicaciones siempre refieren métodos diseñados ad-hoc para imágenes captadas por una tecnología concreta y los algoritmos se prueban sobre un órgano determinado.

Entre los grupos de trabajo más relevantes, hay que destacar el equipo de desarrollo de ITK, financiado desde 1999 por la estadounidense National Library of Medicine of the National Institute of Health, que aglutina a un buen número de universidades y organismos importantes para el desarrollo de software avanzado con el objeto de realizar un atlas minucioso del cuerpo humano y así, impulsar la llegada de las nuevas tecnologías a la medicina. El resultado de este proyecto ha sido un éxito total y producto del mismo es un paquete de software libre para el procesamiento y visualización de imágenes médicas, ITK (www.itk.org) y VTK (www.vtk.org), respectivamente. Tanto ITK, centrado en la mejora de la calidad de la imagen, procesamiento, segmentación y extracción de información, como VTK, manejo y visualización de escenas, imágenes y volúmenes, constituyen un conjunto de clases en C++, que se han erigido como la base fundamental para la construcción de la mayor parte de los paquetes software existentes. Al ser libres, y contar con buenas herramientas de ayuda así como un foro nutrido de muchos desarrolladores internacionales, son utilizadas en muchos trabajos de investigación. Constituye una librería de funciones para el procesamiento y visualización de imagen muy potente, con un API bien definido. Los nuevos algoritmos desarrollados por la comunidad científica, como las técnicas de *level set*, los contornos activos...son implementados e incorporados a medida que los métodos publicados muestran buenos resultados y aceptación. De esta manera constituye una buena librería que consultar antes de hacer un desarrollo. Evidentemente, esta herramienta SW precisa de un elevado conocimiento por parte del usuario, que debe ser un programador experto y especialista en procesamiento de imagen, ya que el conjunto está pensado para ofrecer funciones que integrar en una capa de aplicación superior.

Además de este conjunto de métodos para procesamiento de imagen médica y, siguiendo la misma filosofía, han aparecido otros paquetes, como son IRTK, Image Registration Toolkit, para solucionar el problema de registro entre imágenes, y MITK, Medical Imaging Interaction Toolkit, para desarrollar aplicaciones médicas interactivas. Ambos dos módulos se basan en las versiones y métodos existentes de los ya mencionados ITK y VTK.

Un ejemplo de SW construido a partir del conjunto de métodos proporcionados por estos dos SW es ITK-Snap. Este SW permite el manejo de imágenes Dicom, y su visualización en los 3 planos así como la reconstrucción del volumen. Permite verificar el funcionamiento de algunos filtros y métodos existentes en ITK. En ITK – Snap los algoritmos disponibles están muy orientados a imágenes de cerebro. No es posible segmentar hígado desde este software.

Otro módulo software es Mazda, que ha sido desarrollado por el Institute of Electronics, Technical University of Lodz (Polonia), donde han ido generando diferentes versiones más

completas del software desde 1996. Su objetivo es realizar una segmentación de los elementos a partir de descriptores de textura. Entre las funcionalidades que ofrece, está la posibilidad de seleccionar regiones de interés (*Region Of Interest* o ROI) y extraer múltiples descriptores de textura en dichas regiones de interés (histograma, gradiente, variables de la matriz de co-ocurrencia, wavelet, etc). El detalle de estos descriptores se proporciona en el capítulo 3 de la presente tesis. Una vez descrita una región por un conjunto de características, se puede realizar un proceso de reducción de dimensionalidad, mediante análisis de componentes principales, por ejemplo, para posteriormente, segmentar la imagen por la clasificación de píxeles a través del método k-means, redes neuronales artificiales (*Artificial Neural Networks* o ANN) u otros métodos de clasificación.

VR-Render es un software desarrollado por el instituto IRCAD francés, admite datos provenientes de CT y MRI. VR-Render permite visualizar cada secuencia, mostrando los 3 planos posibles de cada corte del volumen, esto es, los planos axial, frontal y sagital, e informa del valor de nivel de gris de cada píxel. Se puede visualizar el volumen. Tiene unas funciones de transferencia prefijadas para determinadas secuencias, tanto para CT como MRI, que permite resaltar visualmente determinados órganos, pero no realiza una segmentación automática.

Otras herramientas en la literatura son MedINRIA, del centro INRIA francés, que permite la visualización de las secuencias de imágenes en diferentes mapas de colores. Sus funciones para procesamiento de imagen médica están centradas en procesamiento de imágenes de cerebro.

En el caso concreto que compete a esta tesis, se ha intentado cargar una secuencia de hígado captada en MRI para verificar la funcionalidad disponible en los software que aquí se han enumerado, y evaluar, en el caso que se encuentre disponible, la funcionalidad de segmentación automática o supervisada de hígado en el volumen abdominal. Como conclusión a esta revisión de los módulos software disponibles para el manejo y tratamiento de imagen médica, se concluye que, en los que se ha podido probar la carga de la secuencia de imágenes Dicom de MRI y revisado las funcionalidades disponibles para segmentación, no se ha podido realizar ninguna segmentación automática de hígado. Además, ninguno de los módulos software identificados como los más relevantes aprovechan el potencial que el uso de información multisequencia ofrece en las imágenes adquiridas mediante MRI y que aporta información complementaria de gran valor. Con ello, la propuesta de trabajo realizada en el marco de esta tesis tiene cabida en un área de interés tanto para la comunidad científica como para la práctica clínica.

En la literatura se encuentran muchos métodos que trabajan en esta área y que serán convenientemente referidos y explicados a lo largo de los capítulos 3, 4 y 5, según se incida en la obtención de descriptores, métodos de segmentación o los métodos que han inspirado el trabajo de la tesis, respectivamente.

2.5. Caso de aplicación: Segmentación hepática en MRI

2.5.1. Origen y Descripción del problema

El cáncer de hígado es una de las causas más importantes de mortalidad asociada al cáncer en todo el mundo. Únicamente, sólo es posible realizar una extirpación completa de los tumores a entre un 10 y un 20% de los pacientes, lo que constituye, por otro lado, el tratamiento más eficaz para esta enfermedad. La confirmación diagnóstica se realiza mediante biopsias quirúrgicas aunque en la actualidad estas técnicas invasivas pueden ser sustituidas en muchos casos por el diagnóstico por imagen, a través del análisis de imágenes de tomografía (CT) y resonancia magnética (MRI).

Muchas veces el éxito del tratamiento quirúrgico radica en la correcta localización de la lesión o lesiones hepáticas en una porción del hígado perfectamente determinada, de manera que su extirpación no sacrifique la supervivencia del órgano.

La cirugía hepática moderna tiene su base en la anatomía funcional hepática sistematizada (Couinaud, 1957), basada en la distribución en el interior del hígado de los pedículos portales y las venas suprahepáticas (derecha, media e inferior). La proyección vertical de las venas suprahepáticas divide al hígado en cuatro secciones: posterior derecha, anterior derecha, medial izquierda y lateral izquierda. Las fronteras anatómicas entre las cuatro secciones así definidas se denominan cisuras y no se corresponden con las cisuras de la anatomía morfológica clásica. Su importancia radica en que se trata de líneas fundamentales para la penetración en el interior del parénquima durante una resección reglada.

Si se traza un plano horizontal imaginario sobre el eje de la bifurcación portal, se observa cómo las cuatro secciones antes definidas se dividen en ocho segmentos que componen la base de la anatomía funcional hepática (figura 2.9). Esta división del hígado en 8 segmentos es fundamental para la planificación quirúrgica. Las lesiones identificadas deben ser ubicadas con la mayor precisión posible en uno o varios de dichos segmentos, de manera que la operación sea lo más rápida y eficiente posible, se extraiga todo el tejido y los problemas en la intervención sean mínimos. Por ello, un proceso de segmentación automático y preciso es necesario.

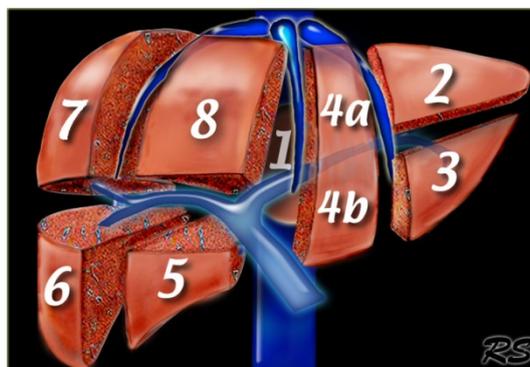


Figura 2.9. División del hígado según Couinaud (www.radiologyassistant.nl)

Las imágenes adquiridas tanto por CT como MRI se pueden interpretar como un volumen, o como una serie de N imágenes 2D, que es realmente como las proporciona el sistema de adquisición, obtenidas como una serie de cortes equi-espaciados a lo largo del eje axial, el que es atravesado por la columna vertebral, existiendo una resolución determinada en dicho eje axial, en la máquina de resonancia correspondiente.

La realización de una buena segmentación es por tanto valiosa para la práctica clínica y ello ha empujado a trabajar en los últimos años en algoritmos de procesamiento de imagen que hagan posible esta segmentación y determinación de qué es hígado y qué no en una imagen abdominal. Inicialmente, las imágenes más estudiadas fueron las adquiridas mediante CT, y a medida que la técnica de MRI se ha ido incorporando en los hospitales como herramienta diagnóstica, se ha comenzado a trabajar sobre las imágenes que proporciona. Las imágenes de MRI, como se ha indicado previamente, proporcionan mucha más información para el diagnóstico, son más ricas en matices, pero por el contrario, presentan una serie de artefactos y problemas que hacen que la segmentación de sus órganos sea más difícil, también en el caso del hígado.

2.5.2. Problemas y Limitaciones

La segmentación de hígado en MRI es un tema de investigación vigente que presenta una serie de problemas y limitaciones, y cuya resolución supone un reto importante.

En primer lugar, **los artefactos de MRI** mencionados previamente dificultan la segmentación en sí. El efecto del volumen parcial, el movimiento y desplazamiento de los órganos debido a la duración de las secuencias y los bordes difusos, principalmente, hacen que sea más compleja la segmentación en MRI que en CT.

Además, en determinadas secuencias, los píxeles que representan el hígado tienen unos **niveles de gris muy similares** a otros tejidos adyacentes, con lo cual, establecer el límite exacto de separación entre los dos órganos es difícil. Este problema es más acusado en la zona en la que el hígado limita con la pared abdominal, donde la franja en la que ambos tejidos aparecen pegados es muy grande. Este problema existe incluso para los radiólogos, a los que también les cuesta a veces determinar con precisión los límites del órgano.

El hecho de que los órganos adyacentes presenten los mismos valores de gris en determinadas secuencias pone de manifiesto dos efectos que pueden suceder como resultado del proceso de segmentación: la sobre-segmentación, ejemplo en figura 2.10, y la sub-segmentación, ejemplo en figura 2.11. La primera circunstancia se da cuando el algoritmo confunde ambas zonas, no es capaz de discriminarlas y da como resultado de la segmentación una región mayor que la del hígado, anexionando al hígado todo o parte de un órgano adyacente. Este fenómeno se conoce también como fuga.

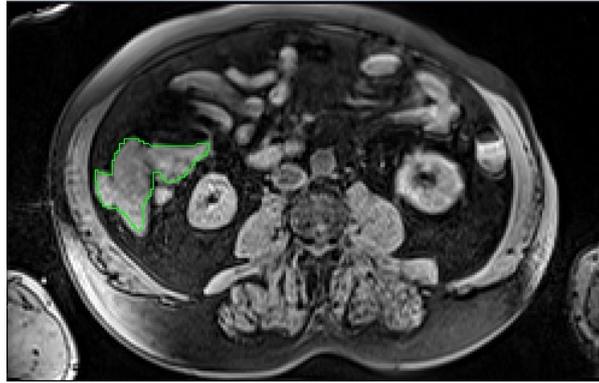


Figura 2.10. Ejemplo de fuga o sobre-segmentación en el caso de localización de hígado

El segundo fenómeno consiste en que la región segmentada es inferior a la región real y se aprecia en la figura 2.11.

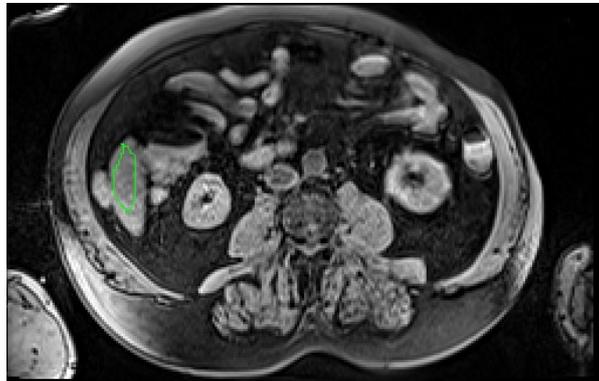


Figura 2.11. Ejemplo de sub-segmentación en el caso de localización de hígado

Las imágenes adquiridas mediante MRI presentan bordes muy difusos, pocos nítidos, y es muy difícil establecer cuál es el punto preciso de borde, ya que existe una zona de transición, como se ha visto previamente en la figura 2.4. Esto dificulta tanto el procesamiento automático como la propia segmentación manual.

Otra de las dificultades existentes en la segmentación de hígado, y en el procesamiento de imágenes de MRI en general, es que, si bien presentan mayor precisión para aportar más detalles que las de CT, dependiendo del objeto u órgano de estudio pueden presentar **poco contraste entre algunos elementos**. Es por ello conveniente la realización de una fase previa de procesamiento de imagen para intentar maximizar el contraste entre los elementos presentes, facilitando de esta manera el procesamiento y segmentación posterior. Se muestra un ejemplo de ecualización de la imagen en la figura 2.12.

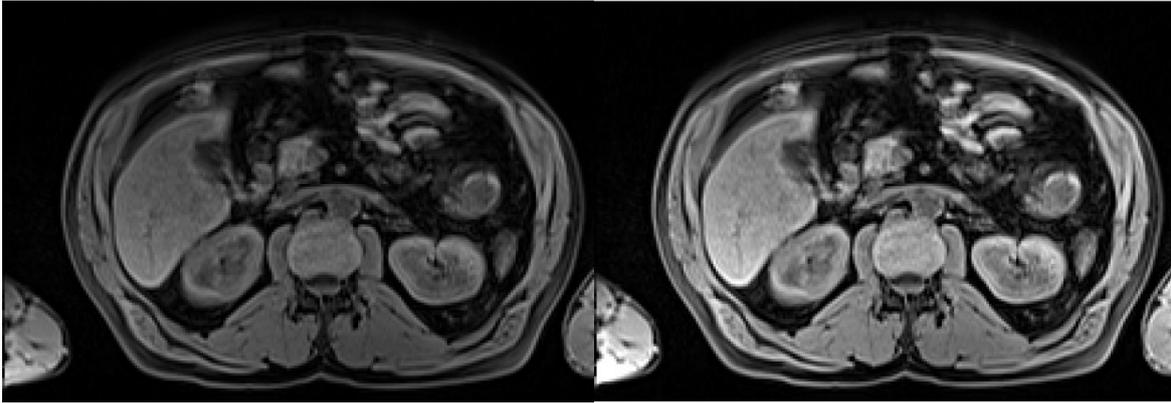


Figura 2.12. Ejemplo de ecualizado de imagen que aumenta el contraste entre órganos.

También es característico de las imágenes de MRI, aunque no exclusivo, la presencia de artefactos, como se ha visto previamente en la sección 2.3.2.2. Estos artefactos dificultan el procesamiento de la imagen. Un filtro de media puede reducir el problema, aunque existen algoritmos más complejos como son filtros anisotrópicos (Bao and Zhang, 2003) e incluso la aplicación de algoritmos de optimización (Boykov and Kolmogorov, 2004) y técnicas variacionales para eliminar el ruido (Chan et al., 2006)(Fang et al., 2010), entendiendo como óptima la imagen sin ruido.

Otro de los efectos que se pueden apreciar es la **iluminación no uniforme en la imagen**, debido a las características de adquisición de MRI. Este hecho dificulta la segmentación, ya que si se entiende segmentación como la separación de la imagen en regiones con características homogéneas, y se considera el nivel de gris como el descriptor más usado, diferentes píxeles de la misma región presentarán intensidades diferentes, lo cual dificulta la resolución del problema. Es recomendable aplicar técnicas de pre-procesamiento de imagen para intentar uniformizar la iluminación en la imagen. Se muestra en la figura 2.13 un ejemplo sencillo, donde se ha corregido parcialmente el problema de iluminación no homogénea convolucionando cada píxel de la imagen con la inversa a la curva lineal de intensidad obtenida en el eje de ordenadas.

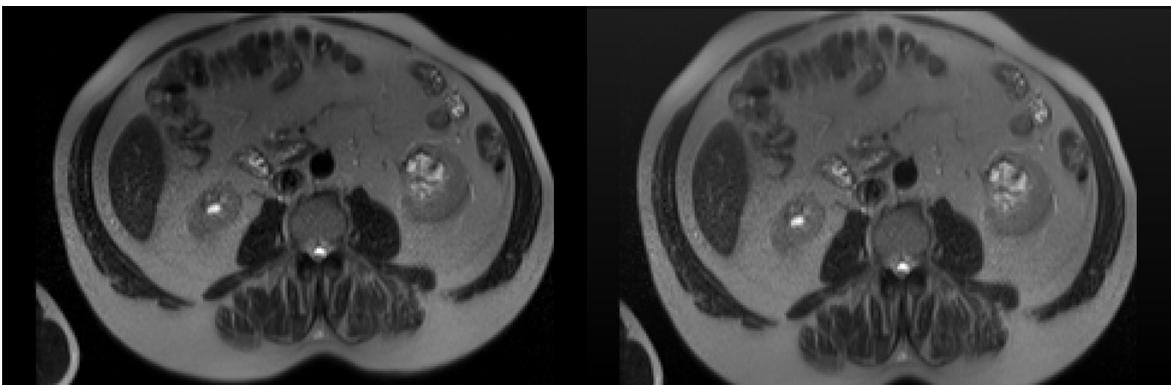


Figura 2.13. Corrección de iluminación no uniforme mediante función lineal inversa de intensidad

En algunas ocasiones, la imagen de MRI presenta también **efectos de saturación**. Se entiende por saturación el efecto de que una parte de la imagen es totalmente blanca, presenta su nivel máximo de nivel de gris, como si de un reflejo se tratase, y evita la percepción de detalles que están contenidos.

En el caso de emplear múltiples secuencias, existe un problema importante adicional a considerar y es el hecho de que las diferentes secuencias estén registradas. Dos imágenes están registradas si el mismo píxel (x,y) de cada imagen (o vóxel (x,y,z) del volumen) está referido al mismo punto físico. De esta manera tendrá valor la multisequencia, ya que se dispone de dos informaciones diferentes del mismo punto físico, es decir, cierto tejido se ha comportado de una manera determinada ante cierto campo electromagnético y de otra manera ante otro. El análisis de este comportamiento es una práctica clínica habitual.

El **registro de las imágenes** es un problema en sí mismo, y un área completa de estudio. A modo de resumen, existen dos grandes grupos de algoritmos para el registro dependiendo del tipo de órgano. Por un lado está el registro afín, que en principio, emplea operaciones de rotación, traslación y escalado y es suficiente (puede ser el caso de cerebro); y por otro lado, está el registro deformable, más complejo, donde las operaciones anteriores no son suficientes y es necesario aplicar técnicas deformables. El hígado constituye un ejemplo de registro deformable, ya que el movimiento respiratorio y a la imposibilidad de apnea en secuencias que duran más de 40 s, provoca movimientos indeseados e incontrolados en el hígado.

El problema de registro será abordado nuevamente en los próximos capítulos y se analizará su influencia en los resultados presentados en el capítulo 6 de esta tesis.

*Capítulo 3: ESTADO DEL ARTE:
Descriptores*

La interpretación de una imagen precisa de la identificación de las partes constituyentes. Esto es lo que se denomina segmentación y será tratado ampliamente en el capítulo 4. Sin embargo, para que un método de segmentación tenga éxito, es relevante la elección adecuada de los descriptores o características descriptivas de la imagen. Dependiendo del problema, unos son más adecuados que otros en cuanto a que aportan mucha más información.

En este capítulo se van a recoger los descriptores más extendidos y utilizados en procesamiento de imagen, y se va a incidir en su aplicación en procesamiento de imagen médica, y en concreto, en segmentación hepática, que es el objeto de esta tesis.

El capítulo comienza introduciendo el concepto de descriptor, y a continuación describe los diferentes tipos categorizados por descriptores de nivel de gris, de forma, de borde o gradiente y de textura. Asimismo, se muestra una serie de ejemplos de aplicación de dichos descriptores a imagen médica y a segmentación hepática. Finalmente, se exponen las conclusiones donde se identifican problemas y limitaciones de dichos descriptores en el marco de esta tesis y se indican las decisiones adoptadas.

3.1. Concepto de descriptor

Un descriptor en procesamiento de imagen es una característica que contribuye a describir el contenido visual de una imagen. Los descriptores que dan información básica a nivel de píxel se denominan de bajo nivel, y representan características de color, de forma, de textura, de gradiente. Los valores obtenidos para estos descriptores a partir de las imágenes constituyen la entrada de métodos de segmentación o clasificación, para localizar e identificar un determinado elemento como perteneciente a una clase concreta.

Los descriptores a veces se combinan con otras técnicas de inteligencia artificial para proporcionar información semántica, cubriendo lo que se denomina el *gap semántico* o *semantic gap*. No es objeto de esta tesis un estudio profundo sobre los descriptores, sino que se analizarán aquellos descriptores más extendidos en procesamiento de imagen para caracterizar estructuras y orientado a la realización de la segmentación hepática, objeto de esta tesis. Se abordarán en este capítulo los descriptores de nivel de gris (color de un solo canal), borde o gradiente, textura y forma.

3.2. Descriptor de Nivel de gris

El nivel de gris es el descriptor natural de una imagen monocroma. Si bien el color es representado en su manera más extendida por 3 canales: rojo, verde y azul (Red Green Blue (RGB)), cada una de estas componentes es un plano monocromo. Este plano consiste en un mapa de grises con un número de valores posibles para cada punto en la imagen que viene definido por la profundidad de píxel, es decir, por el número de bits que representan un píxel, lo que también se denomina resolución en niveles de gris. Así, una imagen con píxel de 8 bits, puede presentar 256 (2^8) valores de gris diferentes, mientras que una imagen con profundidad de píxel o *bitDepth* de 12 bits puede presentar 4.096 (2^{12}) niveles de gris

diferentes. Cuanta más profundidad tiene el píxel más precisión puede presentar la imagen para la caracterización de los elementos, de manera que se pueden apreciar los ligeros matices que puedan diferenciar los elementos presentes en la imagen.

En el marco de esta tesis se ha trabajado con imágenes DICOM obtenidas a partir de una máquina de resonancia Siemens Avanto de 1.5 T, con una profundidad de píxel elegida de 12 bits. A modo de ejemplo, se muestra a continuación una imagen abdominal tipo donde la información se representa correctamente con un píxel de 12 bits y la misma información representada con píxel de 8 bits. Se puede apreciar que en el segundo caso se pierden algunos detalles.

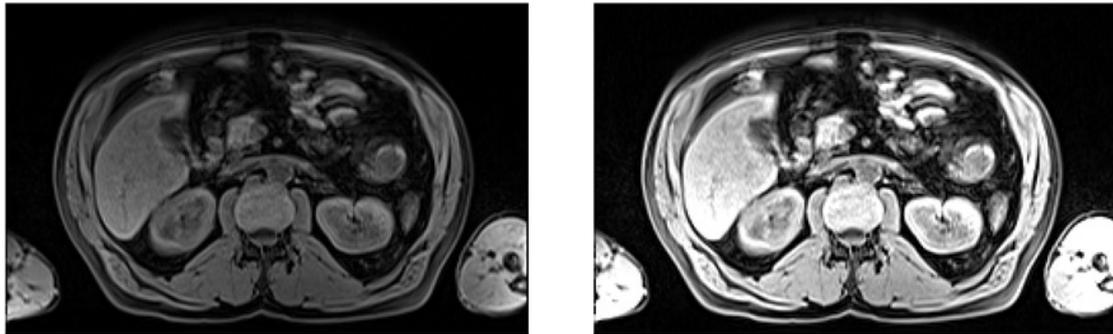


Figura 3.14. Imagen con profundidad de píxel de 12 bits (izda) y 8 bits (dcha)

Las imágenes médicas son todas monocromas. Son adquiridas por diferentes tecnologías, MRI, CT, PET, etc., donde cada máquina de adquisición permite configurar una determinada profundidad de píxel. Cuantos más valores posibles pueda presentar cada píxel, más posibilidad hay de caracterizar e identificar determinadas patologías, que, en muchos casos, se diferencian del tejido sano sólo por ligeros matices de tono.

El nivel de gris es la información de más bajo nivel que se puede obtener de una imagen, y de él el ojo humano extrae la información. A partir de los contrastes entre objetos, se detectan los bordes y los diferentes elementos, y el ojo humano es capaz de segmentar la imagen en milésimas de segundo, si bien no se sabe exactamente cómo. Es de suponer que a partir de dicho valor de gris, extrae la información de bordes y textura que precisa para dividir el resto de la escena. Si la información de bajo nivel de la que se dispone es el nivel de gris y los elementos de la imagen presentan un nivel de gris similar y poco contraste entre ellos, muchas veces el valor de gris no es suficiente para extraer información valiosa de la imagen. Esto es habitual en el caso de la imagen médica.

El histograma es una herramienta clave en procesamiento de imagen, especialmente para visualizar el contraste entre los elementos presentes en la imagen. Si los niveles de gris están concentrados en un cierto intervalo de valores, la imagen es de bajo contraste y bajo rango dinámico. Sin embargo, si el histograma es distribuido y existe separación entre los diferentes niveles de gris, el contraste entre los elementos presentes es elevado.

Sobre una imagen de grises se puede realizar una serie de operaciones que mejoran las condiciones de contraste entre los objetos, lo que facilita mucho la segmentación y

caracterización posterior. Estas técnicas modifican el valor original del nivel de gris. Algunas de ellas son:

- Ecuilizado: método para aumentar el contraste en la imagen mediante la transformación del histograma. Consiste en expandir el histograma de manera que todos los píxeles estén distribuidos en todos los valores de niveles de gris, y no agrupados en unos pocos valores.
- Filtros: matriz que mediante su convolución con la imagen original produce la eliminación de una característica indeseada o la potenciación y enfatizado de otras. Pueden ser lineales, no lineales, anisotrópico u otros.

Se muestra a continuación, en las figuras 3.2 y 3.3, respectivamente, el efecto de la aplicación de la técnica de ecualizado para el aumento de contraste entre los elementos de la imagen, y en el segundo caso, el resultado de aplicar un filtro de media para eliminar el ruido. Se puede apreciar que este filtro de media uniformiza el interior de la región pero que a su vez suaviza los contornos, y provoca que los bordes estén menos definidos.

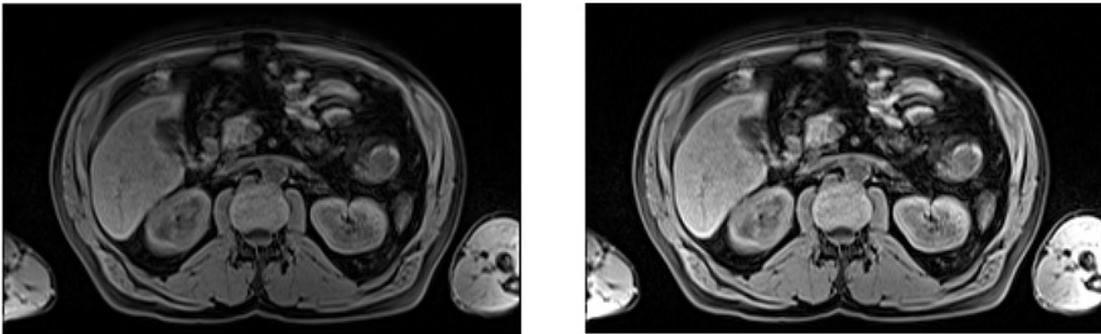


Figura 3.2. Efecto de la ecualización de la imagen

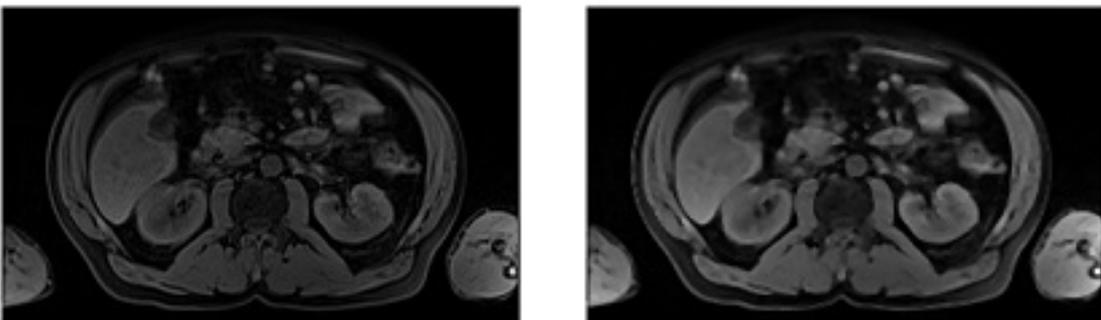


Figura 3.3. Filtro de media para eliminar ruido (3 x 3)

La aplicación de estas técnicas, y otras más complejas, conocidas como técnicas de pre-procesamiento, es habitual y permiten que la información obtenida de la imagen, bien sea el mismo valor de nivel gris directamente u otro descriptor de textura o forma que se obtenga a partir de los valores de gris de un píxel y su vecindario, aporte más información discriminante entre los diferentes objetos presentes. Se aumenta la calidad de la imagen.

Mejorar la calidad de una imagen se traduce en reducir el nivel de ruido, aumentar el contraste entre los elementos constituyentes y resaltar los bordes eliminando zonas borrosas o artefactos relacionado con la adquisición (en el caso de imagen médica).

Una vez se ha aplicado un pre-procesamiento a la imagen y se ha conseguido una calidad aceptable, se pretende obtener una serie de descriptores, esto es, características asociadas a cada píxel, que va a permitir definirlo e identificarlo. Al conjunto de descriptores de una imagen se le denomina vector de características. Estas características serán las que permitirán a las técnicas de segmentación discriminar correctamente las diferentes zonas de una imagen, en base al criterio de separación establecido. Este vector constituye la entrada del método de segmentación o clasificación posterior permitirá extraer la información final deseada.

El nivel de gris constituye el descriptor más directo que se puede obtener de una imagen, si bien puede ser modificado para que tenga más valor y poder descriptivo.

3.3.Descriptor de Forma

Otro tipo de características que pueden describir un objeto son los descriptores de forma. Para que éstos sean unos descriptores válidos y aplicables a la caracterización deben cumplir las siguientes premisas:

- Ser identificables y fiables: distintos ejemplos del mismo objeto deben tener formas similares y claramente diferentes de otro conjunto de elementos, de manera que las características que puedan extraerse de las imágenes les representen.
- Invariantes a traslación, rotación y escala: los descriptores de forma deben cumplir estos condicionantes para describir propiamente un objeto. Este hecho está ligado con la denominada “invarianza afín”. La transformada afín se construye mediante la aplicación de una secuencia de operaciones de traslación, escala y/o rotación. Las características extraídas deben ser invariantes a dichas operaciones. Estas operaciones permiten ubicar en el mismo espacio de referencia objetos de diferentes tamaños y posiciones, para identificar los parámetros de dicho conjunto de elementos.
- Resistentes al ruido: los descriptores deberían ser lo más robustos posible frente a la presencia de ruido.
- Invariante a la ocultación: si en una imagen una parte de un objeto se encuentra oculta, las características que describen la forma del resto del objeto no deben cambiar.

Los descriptores de forma pueden ser de representación estructural o representación estadística, pero en cualquier caso, deben permitir reconstruir el objeto completo a partir de los valores característicos.

Algunos descriptores de forma que representan la estructura de los objetos son el centro de gravedad/centroide, eje de menor inercia, excentricidad, ratios de circularidad, varianza elíptica, rectangularidad, convexidad, número Euler (diferencia entre el número de partes contiguas y el número de agujeros de una forma), ratio de área de agujero, etc. Los momentos de inercia también se utilizan para describir la forma de un objeto. El esqueleto de un objeto también puede ser descrito y es considerado una característica estructural.

Los descriptores de representación estadística se basan en vóxeles, momentos, descripción de puntos a través de perfiles, etc. y se refieren estrictamente a la geometría y no a la estructura del objeto. El volumen de un objeto se puede representar mediante el contorno que delimita la forma tanto en 2D como la superficie en 3D; o mediante nubes de puntos distribuidos de determinada manera.

Un método de segmentación muy extendido, denominado Modelos de Forma Activos, o más conocido por sus siglas en inglés, *Active Shape Models* (ASM) introducido por (Cootes et al., 1995) liga descriptores estadísticos de forma con contornos activos. Este método es descrito en detalle en el capítulo 4, pero en líneas generales, se basa en lo siguiente. Existe un conjunto de puntos característicos (*landmarks*) capaces de describir el objeto de interés. Estos puntos deben poder identificarse en cada nuevo ejemplo de dicho objeto. Pueden ser localizados mediante interacción manual del usuario o automáticamente. De cada uno de esos puntos se puede obtener una caracterización de su línea de perfil, que es la normal al contorno o superficie que pasa por dicho punto. Los valores estadísticos de los perfiles de cada punto de interés caracterizan la forma. Esta nube de puntos es representada en otro sistema de referencia único mediante una transformación, en lo que se puede denominar “espacio de forma” donde el objeto de estudio es una clase claramente separable de otros objetos en dicho espacio de forma. Mediante este análisis, también puede compararse la forma entre objetos pertenecientes al mismo tipo.

Por otra parte, se puede calcular lo que se denomina histograma de forma (Ankerst et al., 1999), que se obtiene fijando el origen en el centro de masas y estableciendo distintos radios (que abarca el objeto) se computa el número de píxeles que representan el objeto en cada radio.

La representación de un objeto mediante descripción de forma ha sido ampliamente estudiada, y principalmente aplicada en la recuperación de imagen por contenido, en el desarrollo de los denominados sistemas CBIR (*Content Based Image Retrieval systems*).

Se puede encontrar amplia información sobre descriptores de forma para caracterización de objetos en las siguientes referencias (Loncaric, 1998), (Kazhdan et al., 2003), (Zhang and Lu, 2004), (Heimann and Meinzer, 2009).

Los descriptores de forma de base estadística mediante nubes de puntos son los más utilizados en imagen médica para segmentación de estructuras. No se va a profundizar más en los descriptores de forma dado que no se aplicarán en el contexto de esta tesis. Esta tesis propone un nuevo método de segmentación de hígado automático. Los modelos de forma

basan la identificación de los *landmarks* o conjuntos de puntos característicos en su identificación manual o semiautomática, lo cual no es el objeto de este trabajo. Además, la obtención automática de *landmarks* en el hígado es compleja, debido a que si bien la forma puede ser reconocible, existen variaciones entre individuos.

La interacción manual puede precisar de una alta dedicación por parte del usuario y aquí se pretende una rápida segmentación. Además, se trata de un órgano con muchas formas curvas y pocas esquinas, vértices o puntos claramente diferenciables, de manera que la definición y obtención del conjunto de *landmarks* es un problema en sí mismo.

3.4.Descriptor de Borde o Gradiente

La caracterización y distribución de los bordes mediante filtros de gradiente es otra manera de describir una imagen u objeto. Existen varios filtros que permiten destacar características específicas en cada caso. De la aplicación de un determinado número de filtros N, se obtiene el mismo número N de imágenes filtradas de las que obtener la información. A partir de dichas imágenes se puede representar cada píxel con los N valores de la imagen filtrada en una coordenada concreta de la imagen (x,y).

Se muestran a continuación algunos de estos filtros de borde. Los operadores de Prewitt (Prewitt, 1970) o de Laplace (Marr and Hildreth, 1980) son también muy conocidos, pero se ha optado por describir el de Sobel.

3.4.1. Sobel

El operador o filtro de Sobel (Sobel and Feldman, 1973) calcula el gradiente de la intensidad de la imagen en cada punto, proporcionando la dirección del mayor incremento en el gradiente de claro a oscuro y la tasa de cambio en dicha dirección. El resultado muestra cómo de abrupta o suavemente cambia la imagen en ese punto, y por tanto, la probabilidad de que dicho punto se trate de un borde junto con la posible orientación del mismo. El operador emplea 2 núcleos o *kernel* de tamaño 3x3 que convolucionan con la imagen original, para calcular la aproximación a las derivadas horizontal y vertical.

$$G_x = \begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix} \quad G_y = \begin{vmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{vmatrix}$$

Figura 3.4. Operador de Sobel en ambas direcciones horizontal y vertical

Los valores obtenidos para el gradiente vertical y horizontal pueden ser combinados para obtener la magnitud del gradiente global, mientras la dirección puede ser obtenida de los valores del gradiente en ambas direcciones. El resultado de la aplicación de este operador sobre una imagen es un mapa bidimensional del gradiente, que indica los posibles bordes.

3.4.2. Kirsch

Este operador (Kirsch, 1971) presenta 8 núcleos o *kernels* diferentes para computar la magnitud del gradiente en diferentes direcciones, resultado de rotar 45° la máscara inicial para cubrir todas las direcciones. Algunas de estas máscaras se muestran a continuación:

$$N = \begin{vmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{vmatrix} \quad NE = \begin{vmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{vmatrix} \quad E = \begin{vmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{vmatrix} \dots$$

Figura 3.5. Operador Kirsch con orientación N (norte), NE (noreste), E (este)

Se aplica la máscara sobre cada píxel de la imagen. La magnitud del borde se define por la máscara que produce el máximo valor y la dirección es la de la máscara que produce la mayor magnitud de borde.

3.4.3. Canny

El operador de Canny (Canny, 1986) se define como un algoritmo multifase, ya que la detección de bordes se obtiene después de realizar varios pasos. Pretende proporcionar resultados óptimos mediante esta secuencia. Inicialmente, aplica un filtro de gaussiana que suaviza la imagen y elimina los posibles ruidos. A continuación, calcula las magnitudes del gradiente a la manera del filtro de Sobel y destaca las regiones con mayor magnitud en la derivada, esto es, mayor gradiente. Una vez obtenidos estos valores, el algoritmo rastrea estas regiones y elimina aquellos píxeles que no presentan estos valores máximos de gradiente. El conjunto con los valores de gradiente que quedan es reducido de nuevo mediante un proceso denominado histéresis. Este proceso emplea dos umbrales. Si el valor de un píxel está por debajo del primer umbral, se clasifica como no borde (etiqueta de 0 en imagen binaria). Aquellos píxeles cuyos valores superan el umbral más alto son clasificados como borde (etiqueta 1 en imagen binaria); y en el caso de que el valor de píxel se encuentre entre el umbral bajo y el alto, éste no es asignado como borde salvo que exista un camino entre este píxel y otro píxel cuyo gradiente sobrepase dicho umbral más alto.

3.4.4. SIFT

El operador *Scale Invariant Feature Transform* (SIFT) es un algoritmo muy popular que detecta puntos de interés invariantes en un objeto de la imagen y calcula sus características descriptoras correspondientes. Estos descriptores pueden ser empleados para reconocer este objeto específico en otras imágenes. Hay que asegurarse de que las características extraídas en dichos puntos de interés son identificables en condiciones cambiantes, como son la escala, la iluminación o la presencia de cierto ruido. Habitualmente estos puntos de interés son ubicados en zonas de la imagen con alto contraste, como son los bordes. El número de características asociadas que extrae el operador SIFT es enorme (hasta 128) para minimizar

los posibles errores causados por las variaciones en escala o posición del mismo objeto en diferentes imágenes.

Por esta razón, este tipo de descriptores no son empleados a menudo en imagen médica, sobre todo en MRI, dado que estas imágenes se caracterizan por el bajo contraste existente entre los diferentes órganos y la dificultad para determinar los bordes.

Los pasos para generar automáticamente este set de características son los siguientes:

- Se aplica un filtro gaussiano a diferentes escalas para identificar los potenciales puntos de interés que son invariantes a escala y orientación
- Se analizan dichos puntos candidatos y se seleccionan aquellos más estables
- Se asigna una orientación a cada punto característico acorde al valor de su gradiente
- Se mide el gradiente en la escala seleccionada y se transforman en una representación que describe perfectamente cada punto

3.4.5. DSIFT

El operador *Dense SIFT*, conocido como DSIFT, es una variación del SIFT que se basa en que un conjunto de descriptores de imagen amplio proporciona en un mallado denso mucha más información que un conjunto de descriptores obtenidos de un conjunto más disperso de puntos. Se establece el área de interés de la imagen a analizar y de la que obtener los puntos de interés y la distancia específica entre puntos de interés. Este operador ha demostrado presentar mejores resultados de clasificación que la versión inicial SIFT, y es el que más se suele emplear actualmente para reconocimiento de objetos.

Los filtros de detección de bordes, como el de Sobel, son empleados como elemento de detección de bordes en un paso concreto del algoritmo, previo a la identificación de regiones por otros métodos posteriores. Este es el caso del trabajo realizado en (Siewert and Schnapauff, 2010)

La segmentación de imagen médica, si bien emplea la detección de bordes como paso previo a la aplicación del método de segmentación elegido, no se basa exclusivamente en la detección de bordes o cálculo de gradiente para segmentar, debido a la enorme dificultad en la detección del borde por el bajo contraste existente en las imágenes, y en consecuencia, la falta de precisión en estimarlo. En el caso de MRI esto es más acusado por los efectos del movimiento sobre la imagen recogida.

3.5. Descriptor de Textura

La textura es la propiedad de las superficies externas de los objetos, que se puede percibir mediante la vista o el tacto. En procesamiento de imagen, se puede definir la textura como la relación espacial de los niveles de gris en la imagen. Esta puede ser regular, si presenta un patrón definido, o aleatoria, si se trata de estructuras visuales no organizadas.

El análisis de la textura mediante técnicas de procesamiento de imagen es complejo, debido principalmente a la alta variabilidad de texturas presentes. Existen muchas técnicas para extraer los descriptores de textura.

El caso concreto de aplicación de información de textura en la segmentación de imagen médica presenta algunas particularidades con respecto a la información que se maneja. El modo habitual de proceder consiste en obtener las características de textura de la imagen, y segmentar mediante la elección de aquella región que presenta las características del elemento a estudiar. También se suelen emplear clasificadores que proponen a qué clase pertenecen las regiones. En el caso de imagen médica se aplica, por ejemplo, a detección y clasificación de lesiones.

Se ha trabajado mucho en el análisis por textura. Las técnicas de extracción de características de textura se pueden agrupar de manera diversa. Se muestra a continuación una propuesta de categorización de dichos descriptores de textura (Tuceryan and Jain, 1998):

- Basadas en estadísticos: representan la textura mediante el análisis de las diferentes relaciones estadísticas entre sus píxeles. Estos valores estadísticos pueden ser de primer orden, segundo orden, lo que se conoce como matrices de concurrencia, o de orden superior.
- Geométricas o estructurales: se definen las texturas mediante la relación entre varias primitivas que definen una microtextura y su distribución en el espacio (macrotextura)
- Basadas en modelo: se emplea un modelo generativo de la textura combinado con modelos probabilísticos. Estos métodos consisten en la localización de los parámetros de dichos modelos.
- Basadas en procesamiento de señal: se aplican diferentes técnicas de procesamiento de señal, tales como Fourier, filtros de Gabor y Wavelets. Se pueden catalogar en:
 - Filtrado en el dominio espacial
 - Filtrado en el dominio de la frecuencia: transformada de Fourier
 - Combinando información espacial y espectral: filtros de Gabor y transformada Wavelet

Estas características permiten asociar cada uno de los píxeles de la imagen con un vector, que representa las características de dicho píxel, así como las relaciones con sus píxeles vecinos y las relaciones entre las diferentes secuencias de la propia imagen (en la resonancia magnética multicanal). Esto puede ser denominado textura espectro espacial entendiendo el espectro como el conjunto discreto de bandas y siendo cada banda una de las secuencias adquiridas mediante resonancia magnética. En el marco de realización de esta tesis se dispone de varias secuencias adquiridas sobre el mismo espacio físico.

3.5.1. Textura basada en estadísticos

Se parte de la base de que la distribución espacial de los valores de gris es una característica que define la textura. Se han identificado varios casos de aplicación en imágenes médicas obtenidas mediante resonancia magnética, como se mostrará más adelante. Se emplean valores estadísticos de 1^{er} orden, 2^o y superiores, que se detallan a continuación.

3.5.1.1. Estadísticos de 1^{er} orden: media y varianza

Momento de una variable respecto del origen

Dada una variable aleatoria x con función de probabilidad o densidad $f(x)$ se puede definir una función de x que sea igual a la variable elevada a un exponente entero no negativo.

$$z(x) = x^k, \text{ siendo } k \in \mathbb{Z}, k \geq 0 \quad (3.1)$$

El valor esperado de $z(x)$ es el k -ésimo momento de la variable x respecto a su origen, se denomina μ y se llama de diferente manera dependiendo del valor de k . El valor de k indica el orden del estadístico. De esta manera, $k = 0, \mu = 1$; y para $k = 1$ (orden 1), $\mu = \text{media}$. El primer momento de una variable con respecto del origen es la media aritmética.

El resto de los momentos respecto al origen tienen escaso interés en la mayoría de los casos.

Momento de una variable respecto de la media

Dada una variable aleatoria x con función de probabilidad o densidad $f(x)$ se puede definir una función de x que sea igual a la diferencia entre la variable y su media aritmética elevada a un exponente entero no negativo.

$$z(x) = (x - \mu)^k \text{ con } k \in \mathbb{Z}, k \geq 0 \quad (3.2)$$

El valor esperado de $z(x)$ es el k -ésimo momento de la variable x respecto a la media y se llama μ_k . Se expresa de la manera recogida en la ecuación (3.3):

$$\mu_k = E[(x - \mu)^k] = \begin{cases} \sum_{x \in X} (x - \mu)^k f(x), & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx, & \text{si } X \text{ es continua} \end{cases} \quad (3.3)$$

Esto es, en cualquier variable aleatoria su primer momento respecto de la media es igual a 0. Esta propiedad se utilizar reiteradamente en las demostraciones estadísticas, para $k = 2, \mu_2 = E[(x - \mu)^2] = \sigma_x^2 = \sigma^2$

Este segundo momento respecto de la media se denomina varianza. La varianza de una variable mide la dispersión de sus valores respecto al valor central μ .

El principal problema de la varianza es que se expresa en unidades cuadráticas que no tienen siempre una interpretación clara. Para obviar este problema se define otra medida de la dispersión que es la desviación típica, σ_x , o simplemente σ , que se calcula como la raíz cuadrada positiva de la varianza. La desviación típica se mide en las mismas unidades que la variable, lo cual es más sencillo de interpretar.

La aplicación de estos conceptos a procesamiento de imagen, conduce a indicar que los estadísticos de 1^{er} orden están basados en el histograma de los niveles de gris y en valores que pueden obtenerse directamente a partir del valor de gris de los píxeles, sin tener en cuenta la relación espacial entre ellos, como es el caso de la media, la varianza o desviación típica.

Estos valores por si solos, en general, no aportan gran valor discriminante, pero muchas veces son empleados no a nivel global, sino en modo local, sobre los que se aplican métodos de clasificación y segmentación potentes. Representan muy bien la información que el ojo aprecia.

3.5.1.2. Estadísticos de 2º orden: Matrices de co-ocurrencia y LBP

3.5.1.2.1. Matrices de co-ocurrencia o GLCM

(Haralick, 1973) propuso las denominadas matrices de Haralick o matrices de co-ocurrencia. El método se denominó GLCM (Gray Level Cooccurrence Matrix) y fue aplicado para caracterizar diferentes tipos de terrenos en imágenes obtenidas por satélite. Rápidamente se extendió la propuesta convirtiéndose en una técnica muy extendida para obtener características de textura en otros tipos de imágenes.

La matriz de co-ocurrencia es una matriz que recoge la frecuencia con la que un píxel con un nivel de gris (i) aparece en una relación de espacio específica con otro píxel de nivel de gris (j). Las matrices de concurrencia son medidas de 2º orden porque consideran parejas de píxeles vecinos, separados una distancia d y en un determinado ángulo. Por tanto, las matrices de concurrencia pueden revelar ciertas propiedades sobre la distribución espacial de los valores de grises.

El método propuesto consiste en la aplicación de los siguientes pasos:

- se dispone de la imagen en niveles de gris
- se definen varias distancias ($d = 1, 5, 10\dots$) y varios ángulos ($0^\circ, 45^\circ, 90^\circ, 135^\circ, \dots$)
- se genera una nueva imagen, donde cada fila son los valores de niveles de gris de la línea de píxeles que comienza en el píxel de interés en la dirección del ángulo correspondiente y la distancia fijada. Hay tantas filas como ángulos.

- la matriz de concurrencia $n \times n$ indica el número de veces que se repite cada relación de niveles de gris. Es decir, punto (1,1) de la matriz, se buscan (siempre en horizontal, por línea), el número de veces que está presente esta combinación, y se escribe. Ejemplo,

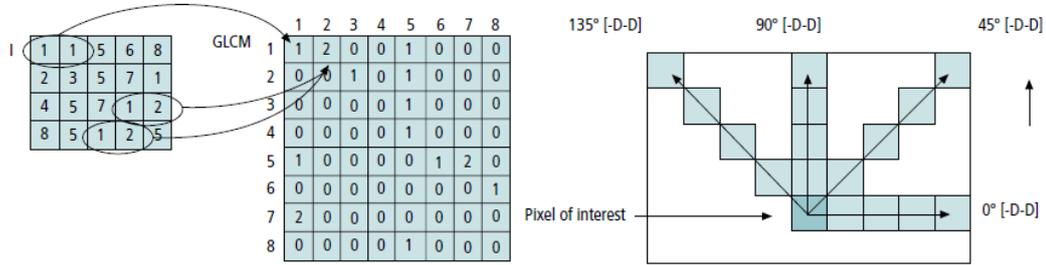


Figura 3.6. Imagen de análisis textural mediante cálculo de matrices de concurrencia

Una vez se dispone de la GLCM (construida con parámetros de segundo orden, ya que incluyen relaciones de niveles de gris entre píxeles a determinada distancia y con cierto ángulo), se pueden obtener algunas variables a partir de ella que son las que se emplearán como descriptores de textura (para clasificación posterior). Algunos de estos estadísticos de 2º orden son:

- **Segundo Momento Angular o Angular Second Moment (ASM)** o uniformidad, también llamado **energía**. Mide la uniformidad de la imagen y cuantifica la repetición de pares de grises. A mayor ASM, mayor uniformidad (menos variación en los niveles de gris). $ASM = 1$, imagen completamente uniforme. P_{ij} es cada elemento de la matriz; siendo i el nivel de gris de un píxel y siendo j el nivel de gris de otro píxel; es decir, sobre GLCM

$$ASM = \sum_{i,j=0}^N P_{i,j}^2 = \sum_{i=0}^N \sum_{j=0}^N P_{ij}^2 \quad (3.4)$$

- **Contraste:** es la cantidad de variaciones locales en los niveles de gris de la imagen. A mayor variación en los tonos de gris, mayor contraste. Si el contraste es 0, los tonos de gris son constantes.

$$contraste = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2 \quad (3.5)$$

- **Correlación:** es la dependencia lineal de los niveles de gris de los píxeles en la dirección del vector de desplazamiento. Si la correlación es igual a 0, no existe correlación lineal.

$$correlacion = \sum_{i,j=0}^{N-1} P_{i,j} \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i \sigma_j)}} \quad (3.6)$$

- **Entropía: medida** de la inhomogeneidad en la textura de la imagen. A mayor entropía, más inhomogénea es la imagen en cuanto a características de textura.

$$entropy = \sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j}) \quad (3.7)$$

- **Momento de Diferencia Inversa o Inverse Difference Moment (ISM)** u homogeneidad local. Indica la homogeneidad de una textura. El ISM aumenta cuando el contraste entre las parejas de píxeles desciende.

$$ISM = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2} \quad (3.8)$$

- **Máxima probabilidad:** El descriptor de máxima probabilidad suministra información sobre la respuesta más fuerte de un par de píxeles en la matriz de coocurrencia.

$$\max = \max(P_{ij}) \quad (3.9)$$

El valor de este descriptor será mayor si en la matriz de coocurrencia se observa un pico de intensidad, el mayor de estos elementos dará el valor máximo de probabilidad. Este descriptor es útil para percibir la mayor ocurrencia o cómo de predominante es un valor en la textura.

La desventaja de emplear las matrices de concurrencia es que en caso de imágenes de elevada resolución, supone un coste computacional excesivo, y disminuir la resolución puede implicar la pérdida de información valiosa de textura.

3.5.1.2.2. Local Binary Patterns (LBP)

Los patrones binarios locales, más conocidos por sus siglas en inglés LBP proveniente de *Local Binary Pattern* se pueden entender como un operador que aúna en un solo descriptor los métodos estadísticos y los modelos estructurales en el análisis de la textura.

El LBP es un descriptor de textura invariante en intensidad con baja complejidad de cómputo. Se ha mostrado que es robusto frente a algunos artefactos en MRI, como son el movimiento y la iluminación no homogénea (bias field)(Unay et al., 2007).

El descriptor LBP fue inicialmente propuesto por (Ojala et al., 1996) a partir del método de (Wang and He, 1990) en el que definía unidades de textura de 8 bits. Dicho LBP en su manera más simple se construye de la siguiente manera:

- se divide la ventana a examinar en celdas
- para cada píxel de la celda, se compara el píxel con sus 8 vecinos, se siguen los píxeles a lo largo de un círculo, en el sentido de las agujas del reloj o en el contrario

- en los casos en que el valor del píxel central es mayor que el vecino, se pone 1, en el caso contrario, se pone 0. Esto proporciona un número binario de 8 dígitos, el cual se suele convertir a decimal para que pueda manejarse. Cada píxel (1/0) se multiplica por su peso ($2^0, 2^1, 2^2, 2^3, \dots$) de arriba abajo y de izquierda a derecha, se suman todos los valores de cada píxel y se obtiene un número para cada celda.
- se obtiene el histograma sobre la celda, junto con la frecuencia que presenta cada número
- se puede normalizar el histograma (o no, opcional)
- se pueden concatenar histogramas normalizados de todas las celdas. Esto proporciona el vector de características LBP de una ventana o zona de la imagen.

Se muestra en la figura 3.7 un ejemplo de cálculo del descriptor LBP para un píxel.

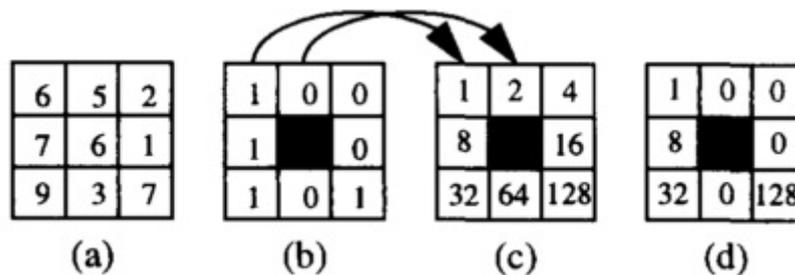


Figura 3.7. Ejemplo cálculo LBP mostrado en (Ojala et al., 1996)

En este caso el valor del descriptor para el píxel central sería $LBP = 1+8+32+128 = 169$.

Este LBP o vector de características puede ser ahora empleado como entrada a un clasificador, por ejemplo, para realizar una buena categorización.

Existen variaciones del LBP, lo que lo ha convertido en un descriptor muy robusto para medida de textura. Una de sus grandes ventajas es que se ha verificado su robustez frente a cambios en los tonos de gris, como, por ejemplo, los causados por variaciones en la iluminación.

Una de sus variaciones más destacables es LBP multiescala, que propone un operador invariante a escala y rotación, donde se barren los píxeles del vecindario de modo circular (*rotation invariant*) y se definen el número de píxeles vecinos y el radio a considerar para identificar los píxeles vecinos (*scale invariant*). Se suelen computar varios valores LBP obtenidos con diferente número de vecinos y a diferente distancia, lo que le convierte en un descriptor de textura robusto.

3.5.1.3. Estadísticos de 3^{er} y 4^o orden y superiores

Los métodos estadísticos, como ya se ha indicado, proponen el análisis de las distribuciones espaciales de los niveles de gris.

- de primer orden: analiza los niveles de gris de cada píxel por separado (media, desviación típica, varianza) a partir del histograma
- de segundo orden: analiza la relación de los niveles de gris entre 2 píxeles (entropía, correlación, contraste a partir de las matrices de co-ocurrencia)
- de orden n : analiza la relación entre n píxeles.

Se ha analizado el comportamiento de los glioblastomas (Jackson et al., 2002). Aunque es necesario hacer más pruebas para certificarlo, parece posible discriminar texturas cuyos estadísticos de 1^o y 2^o orden son visualmente idénticos pero cuyo momento de 3^{er} orden, *skewness*, difiere. Los autores enunciaron que a pesar de que la media y varianza presentaban valores similares, los valores de *skewness* de las distribuciones de píxeles que representan la recirculación de sangre demuestran que hay una diferencia clara entre tumores de grado III y IV.

En aquellos casos donde los datos (valores de gris, por ejemplo) agrupados en el histograma normalizado de los valores de gris, no presenten una distribución simétrica ni se ajusten a una gaussiana, sino a otro tipo de distribuciones, es posible que pueda aparecer un gradiente, un borde debido a un estadístico de 3^{er} orden.

Los valores estadísticos de 3^{er} y 4^o orden se denominan *skewness* y *kurtosis*, respectivamente.

Skewness (3^{er} orden) es el coeficiente de simetría. El tercer momento respecto de la media mide la asimetría de la distribución, es decir, si existen o no observaciones muy extremas en algún sentido con frecuencias razonablemente altas. Si la asimetría es negativa, la variable toma valores muy bajos con mayor frecuencia que valores muy altos y se dice que tiene una cola izquierda pesada o que es asimétrica hacia la izquierda. Si la asimetría es positiva, la variable toma valores muy altos con mayor frecuencia que valores muy bajos y se dice que tiene una cola derecha pesada o que es asimétrica hacia la derecha. Si la asimetría es cero, los valores bajos y altos de la variable tienen probabilidades iguales (el ejemplo más típico de variable simétrica es la variable normal)

La asimetría tiene el mismo problema que la varianza y la covarianza en cuanto a sus unidades de medida y, por ello, normalmente se utiliza una medida adimensional de la asimetría que es el coeficiente de asimetría, g_1 , que se calcula como el cociente entre el tercer momento y el cubo de la desviación típica.

Curtosis (4^o orden). El cuarto momento respecto de la media mide la curtosis de la distribución, es decir, la forma de la distribución de probabilidad. Al representar gráficamente variables con curtosis pequeña, se observan curvas o histogramas con colas

cortas y aspecto aplanado o en meseta; si la variable tiene curtosis grande, su gráfica es alta y estilizada, con colas largas y pesadas.

Todos estos momentos están relacionados con las variaciones de intensidad en la región, y han de ser usados con prudencia porque no siempre son coherentes con la percepción visual. En cualquier caso, pueden ser usadas como características extraídas en aplicaciones de reconocimiento de patrones cuyo valor potencial se explora en un análisis cuantitativo.

De manera general, los n momentos en una imagen pueden ser calculados de la siguiente manera:

$$m_n = \frac{1}{K} \sum_{k=0}^{K-1} (k - \mu)^n h(k) \quad (3.10)$$

sobre una imagen con K niveles de gris, en una región de valor medio μ , histograma $h(k)$ con k en el rango de 0 a $K-1$.

3.5.2. Textura basada en métodos geométricos o estructurales

Estos métodos consideran la textura compuesta por elementos o primitivas e intentan describir estas primitivas y las reglas que gobiernan la organización espacial.

La textura puede ser definida como una disposición espacial de primitivas de textura, organizadas de una manera más o menos periódica, donde la primitiva de textura es un grupo de píxeles que representan el subpatrón más simple o básico. El elemento más pequeño de la textura es el píxel en sí, y es denominado “texton” o “texel” (*texture element* en inglés).

Una vez identificadas dichas primitivas, la textura puede analizarse de dos maneras distintas: se computan las propiedades estadísticas de los elementos extraídos, o bien se extrae la regla de colocación que describe la textura, lo que conlleva estos métodos geométricos. Algunos de estos métodos estructurales se pueden encontrar en estas referencias (Sonka et al., 1993) (Tuceryan and Jain, 1998).

El inconveniente de este tipo de métodos es que no todas las texturas tienen un elemento que se repite. Este tipo de métodos no es usado en procesamiento y segmentación de imagen médica, con lo que no se profundizará en ellos. No aplica en el campo de estudio de esta tesis.

3.5.3. Textura basada en modelos

Se basan en la construcción de un modelo, cuyos parámetros estimados son los que describe la textura. Los campos aleatorios de Markov y los fractales son los más extendidos.

3.5.3.1. Campos aleatorios de Markov

Los campos aleatorios de Markov (*Markov Random Fields* o MRF) son a veces identificados como métodos de segmentación pero realmente son modelos estadísticos que permiten esta segmentación. Los MRF se clasifican a veces como métodos de segmentación de clustering como es el caso de k-means (ver capítulo 4). Una gran ventaja que presentan los MRF es que son bastante robustos frente al ruido. Son empleados en procesamiento de imagen médica, entre otras aplicaciones.

Los píxeles únicos no se asumen que puedan constituir una clase por sí mismos, sino que siempre tienen que ir asociados a una región. Los métodos que utilizan los campos aleatorios de Markov modelan la relación estadística entre un píxel y sus vecinos. Es decir, este método se basa en la hipótesis de que la intensidad de un píxel depende de las intensidades de los píxeles vecinos. Obtienen modelos locales de textura.

El modelo estocástico resultante, es decir, la manera en que se produce la sucesión de variables aleatorias en función de otra variable, generalmente el tiempo, consta de un número de parámetros igual al tamaño de la máscara de la vecindad. En un modelo estocástico cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad y, entre ellas, pueden estar correlacionadas o no.

La máscara se forma tomando una serie de píxeles vecinos, siendo las más sencillas la de primer orden, formada por los cuatro píxeles directamente conectados a él, y la de segundo orden, formada por los ocho píxeles que lo rodean. Los parámetros del modelo se pueden estimar mediante un algoritmo de mínimos cuadrados sobre cada posición de la máscara en la imagen.

Para la obtención del modelo es necesario el entrenamiento con N muestras de las que obtener los parámetros.

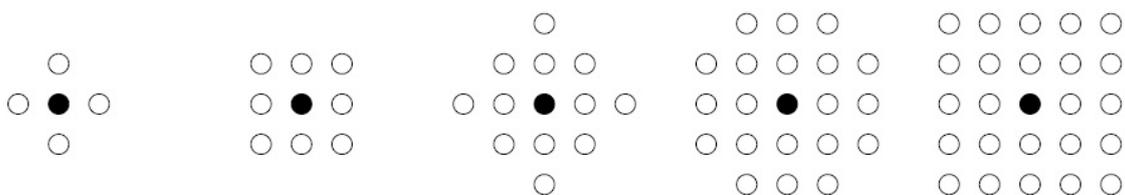


Figura 3.8. Máscaras de orden 1, 2, 4, 5 y 8 que definen la vecindad para el modelo de textura

3.5.3.2. Fractales

Las primitivas son los patrones más pequeños que componen las texturas. Estas primitivas definen la aspereza de dicha textura. En general, primitivas pequeñas dan lugar a texturas finas mientras que primitivas más grandes generan texturas ásperas. A partir de estas

premisas surgió un nuevo enfoque del análisis de textura basado en los denominados fractales.

El análisis de textura por fractales muestra la correlación existente entre la dimensión fractal de una textura y su “aspereza” (*coarseness*). La descripción fractal de texturas se basa en la determinación de la dimensión fractal.

El análisis de texturas mediante fractales fue introducido en 1984 (Pentland, 1984) basándose en una definición de fractales previamente propuesta (Mandelbrot, 1982) para describir elementos de dimensión anómala, siendo el análisis de la estructura de un copo de nieve el primer ejemplo que acuñó el término. Un fractal puede ser descrito como una forma de repetición geométrica, inspirada en la naturaleza, en la cual copias más y más pequeñas del patrón se anidan unas en otras, y la misma forma continúa independientemente de cuánto zoom se realice sobre la imagen.

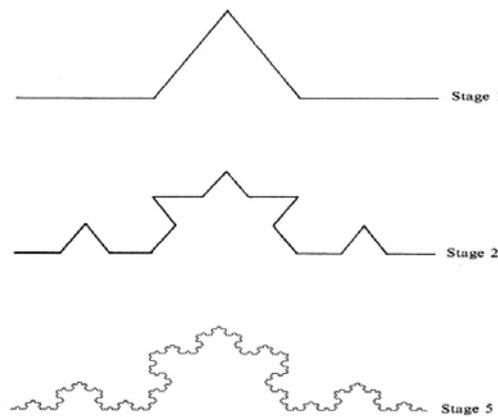


Figura 3.9. Primer ejemplo de fractal: copo de nieve descrito en (Mandelbrot, 1982)

La propiedad de autosimilaridad conlleva que la dimensión fractal de una imagen es independiente de la escala: un conjunto acotado A se dice que es autosimilar, si está constituido por la unión de N copias no solapadas y distintas de sí misma, cada una de ellas escaladas con un factor r . La dimensión fractal correspondiente a dicha región viene dada por la expresión:

$$D = \frac{\log(N_r)}{\log(r^{-1})} \quad (3.11)$$

La dimensión fractal D informa de la rugosidad de una superficie, de modo que cuanto mayor sea D más rugosa será la superficie.

3.5.4. Textura basada en tratamiento de señal

En este tipo de métodos, la imagen se trata como si fuera una señal, y se aplican sobre ella una serie de filtros para extraer características. Estas características se convierten en la

entrada de métodos avanzados de clasificación o para fines de cuantificación y medida, según la aplicación.

3.5.4.1. Filtros en el dominio de Fourier

Los descriptores de frecuencia se basan principalmente en el análisis del espectro de Fourier. Se han desarrollado filtros para análisis de texturas en este dominio de Fourier. Cada filtro permite seleccionar una frecuencia y orientación. De su aplicación se extrae una serie de características de textura de la imagen.

El espectro de Fourier está especialmente indicado para describir la direccionalidad de patrones bidimensionales periódicos de una imagen, ya que estos patrones de textura son fácilmente distinguibles como concentraciones altas de energía en el espectro. Las características más importantes vienen dadas por:

- Magnitud de “picos” prominentes en frecuencia: proporciona información sobre la direccionalidad de los patrones de textura.
- Localización de los “picos”: proporciona información sobre el periodo espacial fundamental de los patrones.

Sin embargo, su implementación se ve restringida por la poca información que se puede extraer cuando las texturas son heterogéneas o de carácter aleatorio.

3.5.4.2. Filtros de Gabor

El filtrado de una imagen con funciones de Gabor parece estar relacionado con los procesos en la corteza visual, y es similar a la interpretación visual humana en cuanto a representación de frecuencia y orientación (Daugman, 1980).

Los filtros de Gabor en procesamiento de imagen se emplean en dos áreas principales: 1) reconstrucción de imagen, dado que el fundamento es similar al de Fourier, es decir, existen unas frecuencias que pueden caracterizar una imagen y a partir de ellas se puede componer y descomponer; 2) caracterización de textura en diferentes regiones de una imagen para su segmentación.

El filtro de Gabor está ampliamente extendido en procesamiento de señal y de imagen. Se trata de un filtro lineal cuya respuesta de impulso es una función sinusoidal multiplicada por una función gaussiana (Gabor, 1946). Cada filtro de Gabor es un filtro paso banda elíptico y orientado, que definen un muestreo del espacio frecuencial. Como salida de la aplicación de dichos bancos de filtros se generan N imágenes a partir de las cuales se puede generar un vector de características de textura para cada píxel. Gabor es un filtro paso banda en 2D, de manera que si se le asigna una determinada frecuencia y dirección, se puede obtener una reducción del ruido a la vez que se preserva una dirección de la imagen original.

Al igual que para el resto de los métodos, dicho vector de características puede constituir la entrada a cualquier clasificador, y también se pueden abordar métodos de reconocimiento de patrones mediante estimaciones estadísticas.

La principal ventaja que se obtiene al introducir la envolvente gaussiana es que las funciones de Gabor están localizadas tanto en el dominio espacial como en el de la frecuencia, a diferencia de lo que ocurre con las funciones sinusoidales, que están perfectamente localizadas en el dominio frecuencial y completamente deslocalizadas en el espacial (las funciones sinusoidales cubren todo el espacio). Por tanto, son funciones adecuadas para representar una señal conjuntamente en ambos dominios. Este es el fundamento teórico propuesto por Gabor.

La transformada de Fourier de un filtro de Gabor son gaussianas centradas en la frecuencia de la función sinusoidal (siendo estas gaussianas la transformada de Fourier de la gaussiana temporal o espacial). Se puede llegar a este resultado empleando la propiedad de convolución de la Transformada de Fourier, que transforma los productos en convoluciones. Así, la transformada de la respuesta de impulso de Gabor es la convolución de la transformada de la función sinusoidal y de la transformada de la función gaussiana.

Los filtros de Gabor pueden diseñarse como un banco de filtros con diferentes dilataciones y rotaciones. Por ejemplo, se pueden descomponer imágenes en bancos de Gabor y volver a reconstruir la imagen original simplemente sumando los diferentes canales. Esto sólo se podría hacer en modo estricto si los canales de Gabor fuesen ortogonales, pero los errores que se introducen muchas veces no son perceptibles en la inspección visual.

El filtro tiene una componente real y otra imaginaria que representan direcciones ortogonales. Ambas partes se combinan para constituir el filtro de Gabor. A continuación dicho filtro se aplica sobre la imagen modificando la frecuencia y la orientación para extraer un conjunto de imágenes resultado de filtrar en esas condiciones determinadas. De una manera general, se suelen aplicar con 3 frecuencias diferentes y 8 orientaciones (se divide el espacio entre todas ellas).

Se muestran 6 filtros de Gabor bidimensionales con diferente orientación y misma frecuencia en la figura 3.10.

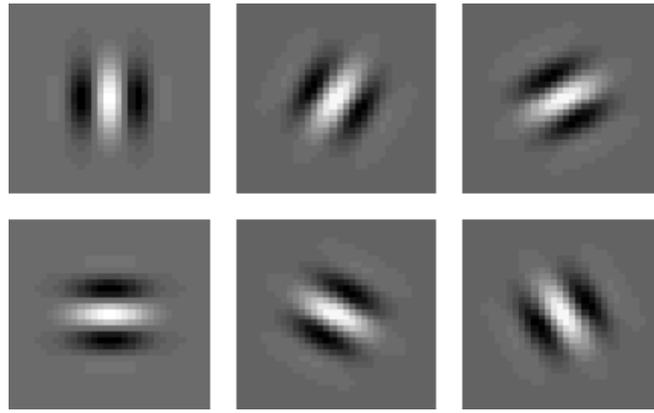


Figura 3.10. Ejemplo de filtro de Gabor bidimensional en 6 orientaciones diferentes

Gabor se aplica localmente, es decir procesando el filtro en ventanas pequeñas, por ejemplo, se pueden generar ventanas de 16×16 para cada una de las m direcciones, obteniendo un vector de m características, una por ventana y dirección. Una vez se dispone de las N imágenes filtradas resultantes de aplicar los N filtros, se pueden obtener valores adicionales, como calcular la energía de la imagen filtrada, calcular la media y varianza de un vecindario de cada píxel analizado, y/o calcular el histograma de la imagen filtrada en determinada zona. El conjunto de características se puede analizar y discriminar para cualquier propósito, y pueden constituir la entrada de un sistema de clasificación.

La elección de los parámetros de sigma, frecuencia y direcciones es un problema en sí mismo y depende mucho de la imagen que se vaya a analizar, de manera que unos parámetros puedan derivar en descriptores exitosos en un caso y no en otros. Existen algunos trabajos en la literatura que evalúan cómo determinar dichos parámetros (Bovik et al., 1990).

3.5.4.3. Wavelets

El estudio de los problemas de convergencia de las series de Fourier, del significado de las funciones y de los sistemas ortogonales, llevó a los matemáticos del concepto de análisis en el dominio tiempo-frecuencia al de análisis en el dominio tiempo-escala. Este desarrollo contribuyó a la posterior definición de wavelets. La primera mención de wavelets se remonta a Haar en 1909.

Las Transformadas de Wavelets (WT) comprenden la Transformada Wavelet Continua (CWT) y la Transformada Wavelet Discreta (DWT). Estas son dos herramientas matemáticas que permiten el análisis de señales de manera muy similar a la Transformada de Fourier. La diferencia fundamental está en que la WT puede entregar información temporal y frecuencial en forma cuasi-simultánea, mientras que la transformada de Fourier sólo da una representación frecuencial. Esta transformada examina la señal a diferentes frecuencias con diferentes resoluciones. Como la WT es capaz de proporcionar información del tiempo y la frecuencia, ofrece buena resolución temporal y baja resolución frecuencial en eventos de altas frecuencias. Por el contrario, ofrece buena resolución frecuencial y baja resolución temporal en eventos de bajas frecuencias.

Este tipo de transformadas están siendo cada vez más empleadas en un amplio campo de especialidades, a menudo sustituyendo a la transformada de Fourier, por su ventaja para el análisis de señales en el dominio del tiempo y la frecuencia.

El análisis multiresolución permite estudiar una imagen en distintas resoluciones y extraer las características propias de cada resolución. El análisis multiresolución se puede realizar a través de la transformada wavelet, ya que en la definición de la misma se puede ver que, a partir de la constante de dilatación a , el análisis multiresolución se encuentra de forma implícita, donde $\Psi(x)$ es la wavelet madre.

La transformada wavelet se expresa según esta expresión:

$$w(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(x) \Psi\left(\frac{x-b}{a}\right) dx \quad (3.12)$$

Los valores de las transformadas son funciones de dos parámetros a y b ; siendo a el factor de escala que fija el nivel de resolución en el análisis de cada imagen, y b el parámetro que indica el lugar de donde proviene el valor transformado. Los $w(a, b)$ son también llamados coeficientes de detalles, cada uno de estos coeficientes indican los detalles o frecuencias que se pueden hallar en la imagen y los lugares de donde provienen dichos detalles o frecuencias.

Uno de los métodos para detectar diferentes texturas consiste en realizar primero la transformada wavelet en una resolución determinada, según qué frecuencias se estén buscando asociadas con esas texturas en la imagen; luego se aplica un filtro pasa altos a estos coeficientes (Gulsrud, T.O., Kjode, 1996) y se les multiplica por un factor a a los miembros que quedan para que al realizar la transformada wavelet inversa resalte los detalles en forma más pronunciada, y por último, se resta la imagen original de la filtrada, obteniéndose puntos en la imagen que indican las fronteras de separación de las diferentes texturas o zonas de igual frecuencia.

Las transformadas wavelets tienen un fundamento similar a los filtros de Gabor, ya que aúnan la componente espacial y la frecuencial. La diferencia radica en qué tipo de función representa cada parte. En los filtros de Gabor la parte frecuencial es una función sinusoidal y la parte espacial viene representada por una gaussiana, que le aporta propiedades de simetría, y demás se comporta casi como paso banda. Sin embargo, en los wavelets, la parte frecuencial puede estar representada por una función sinusoidal u otras, y la parte espacial puede ser cualquier función que se desee. Con lo que no se puede conocer a priori el comportamiento, ni si existe simetría o no.

La idea detrás de las wavelets es poder representar un función cualquiera como la superposición de un conjunto de wavelets o funciones base.

3.5.4.4. Transformada discreta del coseno

El modelo visual humano (Daugman, 1980) emplea la parte frecuencial y espacial, con lo que, al igual que los filtros de Gabor y las wavelets, transformadas como la transformada discreta del coseno, parece adecuada para extraer información de textura de las imágenes.

La transformada discreta del coseno, *Discrete Cosine Transform* o DCT representa una imagen como una suma de funciones sinusoidales de amplitudes variables y distintas frecuencias. Es similar a la transformada discreta de Fourier pero sólo con la parte real. Los coeficientes de DCT se emplean para comprensión de imagen (por ejemplo, en formato JPEG). Los coeficientes de DCT se almacenan como un vector de características. Posteriormente, se puede reconstruir la imagen a partir de dichos coeficientes. Sin embargo, puede perderse algo de nitidez (altas frecuencias, falta de definición y detalle) y aparecer cierto ruido.

Una variación de dicha transformada es la DCT-II muy usada en procesamiento de señal y de imagen dado que tiene gran capacidad de compactación, debido fundamentalmente a que la mayor parte de la información se encuentra siempre en componentes de baja frecuencia, y quedando los detalles para las altas. Esta DCT – II es una DCT aplicada en una matriz y se ejecuta primero en dirección filas y luego en dirección columnas, obteniéndose las frecuencias en dirección vertical y horizontal.

En la siguiente figura se muestran 8x8 funciones básicas de la DCT. En cada paso de izquierda a derecha y de arriba abajo la frecuencia aumenta $\frac{1}{2}$ ciclo. La imagen original es entonces expresada como una combinación lineal de esas 64 frecuencias que aparecen en la imagen. Dichos coeficientes son los que se almacenan y pueden emplearse para compresión o para caracterización de una imagen o estructura que luego constituya el vector de características de entrada a un clasificador o método de segmentación.

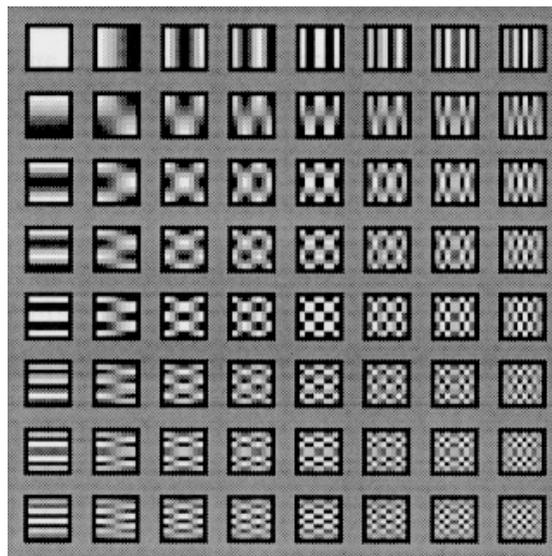


Figura 3.11. Funciones básicas de transformación 2D de la DCT

Existen otros métodos de análisis de textura para la obtención de descriptores, como la matriz de autocorrelación o la matriz *run-length*. Se puede revisar un compendio de estos métodos en estas referencias (Haralick, 1979) (Tuceryan and Jain, 1998) (Zhang and Tan, 2002).

3.6. Aplicación en imagen médica

Las imágenes médicas, en general, y sobre todo, las obtenidas mediante resonancia magnética presentan numerosos problemas que dificultan la segmentación y caracterización automática, como ya se ha indicado en el capítulo 2. Algunos de estos problemas son la baja nitidez y definición de bordes, ruidos, y el hecho de que tejidos diferentes presenten niveles de gris similares. Por ejemplo, en la siguiente imagen se aprecia que el hígado (flecha roja) presenta un nivel de gris similar a la propia pared abdominal (flecha verde) en determinada secuencia, lo cual claramente dificulta el proceso de segmentación automática.

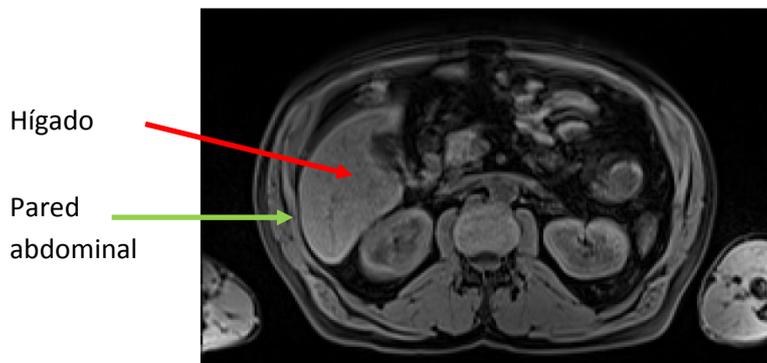


Figura 3.12. Órganos con valor de nivel de gris similar en una secuencia

Para abordar cualquier problema de segmentación es necesario identificar qué descriptores se van a emplear y qué método de segmentación. En la medida en que los descriptores sean más complejos, y representen el contenido visual de la imagen de manera certera, eficiente y mediante características linealmente independientes, el problema podrá ser resuelto con un método de segmentación relativamente sencillo que se base en dichos descriptores. Pero en caso de que el descriptor aporte poco valor, habrá que incidir fuertemente en el método de segmentación a usar, debiendo abordarse métodos avanzados.

Los estadísticos de primer orden (media, varianza) no aportan apenas valor en situaciones complejas si el método posterior de segmentación o clasificación es sencillo, y ha de recurrirse a los de 2º orden (matrices de concurrencia), órdenes superiores u otros descriptores, como los expuestos a lo largo de este capítulo. Se encuentra literatura en la que se ha recurrido a estadísticos de 3º orden o superiores para la detección de límites de tumores, en estructuras donde el ojo humano es incapaz de observar diferencia en 1º y 2º orden, por ejemplo, se ha usado para analizar glioblastomas (Jackson et al., 2002). En este caso, Jackson et al. muestran resultados de cómo es posible diferenciar tejidos aparentemente iguales con el gradiente del momento de 3º orden. En dicho trabajo, la metodología empleada para obtener los datos emplea una ventana $n \times n \times n$ que barre todo el

volumen 3D, se calculan los momentos, se hace un histograma y se miran los gradientes para ver los picos, lo cual indicaría los bordes, es decir, los límites de dichas regiones.

El análisis de textura ha sido explorado en diversos campos de aplicación y por supuesto, también en el procesamiento de imagen médica, principalmente para la localización y segmentación de órganos y para la caracterización de patologías.

Se encuentran en la literatura multitud de enfoques para la obtención de descriptores de textura sobre una imagen médica. Por ejemplo, se ha utilizado la expansión de la serie de Taylor, sus coeficientes de 1^{er} y 2^o orden para obtener las características de textura que controlan los movimientos de un modelo activo de forma (*Active Shape Model* o ASM que será descrito en el capítulo 4) (Ginneken and Frangi, 2002). Este ASM es aplicado para segmentar un área concreta del cerebro. La serie de Taylor de una función $f(x)$ infinitamente derivable (real o compleja) definida en un intervalo abierto $(a - r, a + r)$ se define como la siguiente suma:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(a)}{n!} (x - a)^n \quad (3.13)$$

En aplicaciones médicas, también está muy extendido el uso de las matrices de co-ocurrencia o GLCM, descritas en el apartado 3.5.1.2.1 a partir de las cuales se obtienen una serie de descriptores que constituirán el vector de características. Sobre él se suele hacer análisis de correlación, reducción de componentes u otras técnicas para que el vector de características contenga aquellas que son linealmente independientes. Existen muchos ejemplos de aplicación, como es el caso de clasificación de tejidos de próstata en ultrasonidos (Basset et al., 1993), donde consiguieron clasificar el 78% de los tejidos mediante matrices de concurrencia, tejidos que no podían distinguirse con el ojo humano; imágenes de hígado de ultrasonidos (Wu et al., 1992) donde se realizó un análisis de textura para diferenciar y clasificar (mediante clasificador de Bayes) los distintos hígados como normales, cirróticos o hepatoma. Las características de textura incluidas en el vector de entrada también se componían, en este caso, de coeficientes de las frecuencias de Fourier, y estadísticos obtenidos de las matrices de coocurrencia. Con dicho vector de entrada, los autores consiguieron clasificar correctamente hasta un 90% de los tejidos de hígado en estas imágenes de ultrasonidos.

El análisis de textura también ha sido empleado en el análisis de mama para la caracterización de los posibles tumores benignos o malignos contenidos. Chen et al. realizan una clasificación de tejidos de mama a partir de coeficientes de wavelets como vector de características de entrada y clasificador mediante redes neuronales. Distinguen entre tumores benignos y malignos, consiguiendo una sensibilidad entorno al 90% (Chen et al., 2002).

Los filtros de Gabor también han sido testados para validar si aportan información discriminatoria de la textura de los tejidos presentes en las imágenes médicas como, por ejemplo, en el análisis de rodilla por ultrasonidos (Xie et al., 2005).

3.7. Aplicación a segmentación hepática

En el área de patología hepática se han identificado muchos trabajos con diversos tipos de descriptores que, unidos a métodos de segmentación o clasificación, han permitido caracterizar convenientemente el hígado y algunas de las patologías presentes en él. Estos estudios han sido realizados tanto para imágenes obtenidas por CT como por MRI. El objeto de estudio de esta tesis es sobre imágenes de MRI, más difíciles de procesar, siendo los estudios sobre CT más frecuentes. Al ser la respuesta de los tejidos del cuerpo diferente ante ambas tecnologías de adquisición, es posible que los mismos descriptores no representen la misma información en ambos tipos de imágenes, y que permitan discriminar más elementos en una que en otra. En principio, MRI ofrece mayor riqueza de detalles en los diferentes tejidos.

El nivel de gris y los estadísticos de primer orden constituyen descriptores ampliamente utilizados en la segmentación hepática. La mayor parte de los métodos de segmentación que se exponen en el capítulo 4, como son los contornos activos, los métodos estadísticos basados en forma, y el resto de los modelos deformables, utilizan nivel de gris (Gloger et al., 2011) (Masoumi et al., 2012) (Göçeri et al., 2014). Algunos de ellos también incorporan descriptores de textura que pueden aportar un valor diferencial (Zhang et al., 2005) (Gambino et al., 2010) (Luo et al., 2013). Esta incorporación se realiza principalmente cuando el objetivo principal no es segmentar el hígado, sino caracterizar las lesiones presentes en él, y es un trabajo más de clasificación que de segmentación. Para clasificar, en general, se opta por extraer el mayor número posible de características, identificar las más relevantes, mediante análisis de correlación y otros métodos discriminantes y a continuación, aplicar una técnica de clasificación, normalmente supervisada, que categorice la entrada en una de las clases establecidas, cuyas características específicas han sido calculadas mediante entrenamiento previo.

En segmentación hepática se pueden encontrar algunos ejemplos basados en descriptores de forma, como son (Heimann et al., 2007), (Seghers et al., 2007), (Kainmüller et al., 2007), (Zhang et al., 2010).

Existe un ejemplo de segmentación de hígado en imagen abdominal donde caracterizan hígados cirróticos o sanos a partir de análisis de imágenes laparoscópicas, y los puntos de interés son obtenidos mediante el operador SIFT (Sinha et al., 2014). Las características de dichos puntos son la entrada de un clasificador lineal, y obtiene resultados con tasa de acierto en torno al 90%. En otro trabajo se emplea el operador SIFT para caracterización de movimiento 4D tanto en hígado en MRI o en pulmón y aparece como un extractor de características robusto para describir el movimiento de dichos órganos (Paganelli et al., 2012).

En relación a la descripción por textura, se propuso un método para identificar la fibrosis hepática (Cao et al., 2005). Para ello los autores combinaron 2 tipos de descriptores de textura, uniendo la dimensión fractal y el descriptor entropía de la matriz de concurrencia de borde (independiente del nivel de iluminación y ganancia en la adquisición). Para la

clasificación se empleó el discriminante de Fisher lineal (Fisher, 1936) y el método de Support Vector Machine (SVM) (Cortes and Vapnik, 1995). Los descriptores de entropía resultaron mucho más relevantes que los fractales para la clasificación, en la que tanto con Fisher como con SVM se lograron ratios de alrededor del 90%. Las imágenes de estudio tenían ventanas de 64x64 cuidadosamente elegidas, que no fueron automáticamente extraídas de la imagen.

En otra de las experiencias identificadas en la literatura se emplearon variables estadísticas sobre tejidos de hígado para diferenciar hígados sanos de aquellos que sufrían cirrosis (Nishiura, 2005). En otra aproximación al mismo problema, se emplearon los descriptores de Haralick sobre matrices de coocurrencia de dimensiones de celda de 32x32 como vector de entrada a una red neuronal, capaz de predecir el grado de cirrosis padecido (Zhang et al., 2005).

Se ha identificado una aproximación multimodal sobre imágenes de MRI, CT y ultrasonido en hígado combinando características de textura extraídas de ellas para la detección de hígado sano, quistes fibrosis y grasa (Chung and Logeswaran, 2007). Han conseguido un buen ratio de acierto con matrices de concurrencia de nivel de gris y estadísticos de primer orden.

Más recientemente, se ha propuesto una caracterización del hígado y algunas de sus patologías mediante un extenso conjunto de características de textura de diferentes tipos (Mayerhoefer et al., 2010). Para ello, emplean el histograma de niveles de gris, las variables de la matriz de coocurrencia y matriz *run-length*, gradiente y la transformada wavelet. De aquí se extrae el conjunto de las características más relevantes. Mediante este conjunto y métodos de agrupamiento *k* vecino más cercano (*k nearest neighbour* o *k-NN*) y *k-means* se clasifican las lesiones contenidas en las imágenes y discriminan tejidos de fibrosis y hemangioma. Dichos métodos de agrupamiento serán descritos en el capítulo 4.

También se trabaja con primitivas de textura, siendo estas primitivas generadas sintéticamente, que luego son clasificadas mediante redes neuronales para la categorización de tejidos (Sharma and Ray, 2009), aunque en este caso, se hicieron pruebas exclusivamente con imágenes sintéticas, con texturas de Markov. Posteriormente, se genera un vector de características que contiene 3 valores estadísticos de 1^{er} orden (media, desviación estándar, entropía) y hasta 14 características derivadas de las matrices GLCM, es decir, estadísticos de 2^o orden. Este método se ha probado sobre imágenes de hígado obtenidas por CT y los resultados en la caracterización de tejidos son aceptables.

En otro de los trabajos identificados, se pretende realizar la clasificación de 3 tipos de hígados: sanos, con cirrosis y hematoma (Wu et al., 1992). Para ello, se emplea inicialmente un vector de características que contiene el espectro de Fourier y la relación de niveles de gris, entre otros, con poco éxito inicial. A continuación, se emplea un vector de características de multiresolución fractal para obtener una localización difusa de las enfermedades, asegurando un ratio de la clasificación del 90%.

Existen más ejemplos de aplicación de descriptores de frecuencia para la localización de hígado. En (Luo et al., 2009) se emplearon los coeficientes de Wavelet y a continuación se propuso aplicar Support Vector Machine (SVM) para clasificar los píxeles como pertenecientes a la clase hígado u a otras. En una propuesta posterior del mismo autor, se propuso la combinación de estadísticos de 2º orden, como son las matrices de co-ocurrencia de Haralick y los coeficientes Wavelet para combinar información local y de los órganos adyacentes y se clasificó cada píxel posteriormente a través de un SVM junto con un procesamiento morfológico posterior para conseguir una segmentación del hígado adecuada sobre imágenes de CT (Luo et al., 2013). En este trabajo se muestra que se obtienen mejores resultados con la combinación de descriptores que solamente con las características de las matrices de Haralick. La gran baza de este tipo de enfoques es que al haberse desarrollado técnicas de inteligencia artificial muy potentes, como SVM, las redes neuronales (ANN), técnicas fuzzy, neuro-fuzzy, que pueden encontrar cierta separabilidad entre clases fuertemente no lineales, se traslada el problema de la interpretación o manejo de los valores de dichas características a estos clasificadores no lineales.

Otro ejemplo de uso de SVM sobre características de textura propone el uso de los coeficientes de la transformada discreta del coseno (DCT) como las características discriminantes empleadas (Danciu et al., 2012).

El análisis de texturas mediante los campos aleatorios de Markov (Markov Random Fields, MRF) también ha sido aplicado a la segmentación de hígado junto con contornos activos, donde los contornos activos han refinado la segmentación inicial burda realizada mediante MRF (Alomari et al., 2008).

Los ejemplos referidos pueden ser incluidos tanto en este capítulo como en el siguiente, ya que la aplicación de descriptores y métodos de segmentación es insoluble. No se han repetido los ejemplos en ambos capítulos.

3.8. Conclusiones

La localización de un elemento determinado en una imagen necesita de la aplicación de un método de segmentación que cuente con una entrada, un vector capaz de describir el contenido visual presente en la imagen, y a partir del cual, el método de segmentación proporcione una buena separación entre regiones.

A lo largo de este capítulo se han expuesto los descriptores más habituales de color (nivel de gris en este caso), forma, borde o gradiente y textura. En el capítulo 4 se exponen los métodos de segmentación más frecuentes. Se han orientado ambos capítulos a su aplicación a imagen médica, y en concreto a la segmentación de hígado en imagen abdominal obtenida mediante tecnología MRI.

Los operadores o filtros para detección de bordes o gradientes son muy utilizados, pero no es habitual que sean los únicos descriptores a partir de los cuales realizar una segmentación valiosa y robusta de la imagen, y menos en imagen de MRI donde los artefactos, el ruido, el

bajo contraste y la baja definición de los bordes es un hecho. Los operadores de gradiente se emplean para la identificación de puntos de interés o como parte del proceso en la cuantificación de la energía de un contorno activo, level set o modelo deformable.

En general, los métodos basados en descriptores de textura asemejan la manera en la que el ser humano localiza las regiones, sus bordes y los segmenta, acorde a texturas y regiones de apariencia uniforme. Existen multitud de descriptores de textura y muchos clasificadores no lineales muy potentes. Esto permite explorar muchas vías para la mejora en la precisión de la segmentación de órganos o patologías que el ojo humano no puede distinguir. Este enfoque presenta algunos inconvenientes. Por una parte, no es intuitivo decidir cuáles son las características de textura que aportarán más valor en la segmentación, con lo que el problema puede estar sobredimensionado. Tampoco se sabe a priori qué método de clasificación va a ser el más potente. Si bien un SVM necesita pocos ejemplos para estar entrenado y ser muy eficiente, tampoco funciona siempre. Se parte de la base de que las clases son separables por textura pero no tiene por qué ser así. En MRI principalmente, la extracción de características relevantes de textura es compleja porque las imágenes son de menor claridad, presentan ruidos y artefactos que alteran las características de textura que son susceptibles de caracterizar adecuadamente cada órgano presente en la imagen.

En el caso de aplicación del análisis de textura en imágenes de MRI y segmentación de hígado, cabe decir que la resolución espacial (volumen) depende del resonador, de la máquina de adquisición MRI. De esta manera las máquinas de 1 - 1.5 teslas, proporcionan bajas resoluciones y las texturas no son distinguibles. En los nuevos resonadores de 3 teslas, la resolución es mayor y se pueden observar las texturas mucho mejor. La mayor parte de los hospitales disponen de las de 1.5 teslas. Las imágenes sobre las que se ha trabajado en esta tesis han sido obtenidas mediante máquina de resonancia de 1.5 teslas. Además la presencia de artefactos en MRI implica un solapamiento en las intensidades de las clases de tejidos, de manera que la segmentación no es trivial.

Tras analizar el estado del arte y para abordar el problema de segmentación de hígado expuesto en esta tesis, se puede optar por dos enfoques diferentes: 1) trabajar el descriptor o, 2) trabajar el método de segmentación.

La primera posibilidad consiste en trabajar en un descriptor avanzado, que combine varios de los descriptores que se han presentado en este capítulo o incluso crear un descriptor nuevo, capaz de caracterizar con detalle todos los elementos, de manera que un clasificador sencillo permita discriminar el hígado del resto de tejidos presentes en la imagen. Se podrá abordar esta línea de trabajo en la medida en que las imágenes presenten más resolución de píxel, lo cual puede aportar matices y diferencias para la caracterización de tejidos. Esta mayor precisión en la información obtenida es el objetivo de las máquinas de MRI de 3 Teslas, que ya están implantándose en los hospitales.

La segunda posibilidad consiste en emplear el valor de nivel de gris como descriptor. El hígado en sí no muestra, al menos de manera visual, una textura en imágenes MRI claramente diferenciada de otros órganos o elementos presentes en la imagen, como son los

riñones, intestino o grasa abdominal, sino que la diferenciación con estos órganos es simplemente por intensidad de gris. Lo mismo sucede con las patologías asociadas, como así nos han indicado los radiólogos de referencia. Las imágenes han sido captadas con escáner de 1.5 T. De hecho, se han realizado algunas pruebas preliminares de obtención de descriptores de textura sobre varios órganos presentes en la imagen, y la firma visual de los mismos es muy similar, con lo que la caracterización por textura no se ha podido realizar. La otra posibilidad es emplear la característica discriminante que ahora mismo usa el radiólogo, es decir, el nivel de gris, e incidir en el método de segmentación. En el marco de esta tesis se ha optado por esta segunda vía. La propuesta de método segmentación se expondrá en el capítulo 5 y los resultados obtenidos para el conjunto de imágenes de prueba disponibles serán recogidos en el capítulo 6.

*Capítulo 4: ESTADO DEL ARTE:
Segmentación*

La segmentación de una imagen es uno de los problemas más importantes de la visión artificial. Los seres humanos somos capaces de adecuarnos a diferentes entornos, y extraer información de ellos de manera rápida y eficaz. Sin embargo, la realización de esta acción de una manera automática es un problema complejo. En la actualidad, no existen métodos o algoritmos que puedan realizar la segmentación universal de imágenes, sino que debe realizarse un algoritmo *ex profeso* para cada caso, siendo muy relevante la iluminación y la información previa de lo que va a aparecer en la escena.

En el caso concreto de análisis de imágenes médicas, el objetivo es ofrecer un análisis cuantificado de características existentes en un determinado órgano o lesión. Esto requiere la separación de la imagen en regiones que presenten un conjunto de características similares mediante un procedimiento de segmentación. Este proceso presenta una gran dependencia del problema concreto, no siendo lo mismo localizar tumores en el hígado que cuantificar la materia gris del cerebro o localizar una fractura ósea. Cada caso exige un procesamiento específico. La segmentación de imagen médica es un caso concreto del problema general de segmentación de imagen en visión por computador. Muchos grupos de investigación están trabajando en esta línea para conseguir la segmentación automática de la escena, con la menor intervención posible por parte del usuario, para su interpretación posterior.

En este capítulo se muestra el estado del arte de los métodos de segmentación más populares presentes en la literatura, que han sido desarrollados y continúan siendo mejorados por la comunidad científica que trabaja en el procesamiento de imagen. Se ha enfocado el estado del arte en su aplicación a imagen médica. De entre todos los métodos expuestos, se hace un especial análisis de los que han constituido la base de esta tesis, esto es, los contornos y superficies activas, que forman parte del conjunto de métodos de los modelos deformables.

Dado que el objeto de esta tesis es la segmentación del hígado sobre imagen abdominal obtenida mediante resonancia magnética, se va a hacer una revisión de los métodos desarrollados para segmentación de hígado, de cuyo análisis ha surgido la propuesta del algoritmo de segmentación aportado en esta tesis, y cuya aproximación teórica se detalla en el capítulo 5, para posteriormente describir los resultados de las pruebas pertinentes en el capítulo 6. Las conclusiones obtenidas del estudio y las posibles líneas de trabajo futuro se exponen en el capítulo 7.

Junto con los métodos de segmentación, en este capítulo 4 se añaden algunos apartados en relación con los diferentes puntos que aborda la tesis. Por una parte, se describen las técnicas variacionales como método de minimización, que forma parte de algunos procesos de segmentación, como es el caso de los modelos deformables. De los múltiples métodos para segmentación hepática recogidos en la literatura, se analizan especialmente los trabajos de segmentación de hígado propuestos en 3D. También se estudian en detalle aquellos métodos realizados sobre múltiples canales o secuencias de MRI. Finalmente, se exponen las limitaciones de los métodos existentes, algunas de las cuales son solucionadas con el método propuesto en esta tesis. Se acaba el capítulo con unas conclusiones.

4.1. En qué consiste la segmentación

La segmentación de una imagen consiste en la división o separación de la imagen en regiones de atributos similares. Es decir, la segmentación subdivide una imagen en sus partes constituyentes u objetos. El grado de subdivisiones depende del problema a resolver, por lo que la segmentación se habrá realizado correctamente cuando los objetos de interés del estudio hayan sido aislados. De manera sencilla, la segmentación consiste en que a cada píxel de una imagen se le asigna una etiqueta correspondiente a cada región. Todos los píxeles con la misma etiqueta (región) son similares en una serie de características, que pueden ser de intensidad, color o textura.

En el caso de la imagen médica, las técnicas de segmentación se aplican a la delimitación de órganos o tejidos. El cerebro es uno de los más estudiados, pero se encuentran también multitud de estudios para localizar riñón, pulmón, corazón, colon, próstata, hígado, así como vasos sanguíneos, detección de tumores o patologías concretas. Una vez localizados e identificados, se suele obtener información cuantitativa resultante del análisis de la región segmentada. En los casos de planificación quirúrgica, las regiones segmentadas se reconstruyen en un volumen 3D. Las técnicas de segmentación se pueden aplicar sobre imágenes obtenidas mediante distintas tecnologías de adquisición, como MRI, CT o PET, que normalmente requieren de técnicas de pre-procesado o mejora en la calidad de la imagen. Estas tecnologías de adquisición se han descrito brevemente en el capítulo 2.

La segmentación de imágenes médicas implica abordar imágenes complejas que presentan muchos problemas adicionales como son ruido, volúmenes parciales, movimientos en la imagen, variabilidad entre órganos, artefactos derivados de las distintas técnicas de adquisición, límites imprecisos entre órganos y necesidad de conocimiento experto para la identificación de ciertas zonas.

Los algoritmos clásicos para segmentar imágenes monocromáticas generalmente se basan en una de las dos propiedades básicas de los valores del nivel de gris: discontinuidad y similitud. En la primera categoría, la aproximación es realizar una partición de la imagen basándose en cambios abruptos en el nivel de gris. Las principales áreas de interés en esta primera categoría son la detección de puntos aislados, de líneas y de bordes en una imagen. La segunda categoría se basa en segmentar por similitud acorde a niveles de umbral, generando, por ejemplo, regiones de crecimiento.

4.2. Introducción al problema de segmentación de hígado en imagen médica

La segmentación de hígado en imagen abdominal es un tema de interés que ha dado lugar a múltiples trabajos, principalmente sobre imágenes obtenidas mediante tomografía axial computerizada (TAC o CT, siglas en inglés), y también, aunque en menor proporción, sobre imágenes de resonancia magnética (RMN o MRI, siglas en inglés). Estas dos tecnologías ya se han descrito en el capítulo 2. Además, la utilización de resonancia magnética multiseccional permite una descripción matemática mucho más rica y precisa de las estructuras anatómicas contenidas en la imagen al poder ser éstas descritas por sus patrones

de comportamiento en las diferentes secuencias, lo cual facilita el diagnóstico. El inconveniente de las imágenes de MRI es que son más difíciles de procesar que las de CT ya que existe menor contraste entre los órganos y los bordes son más borrosos debido, entre otros, a la propia tecnología de adquisición y a efectos indeseados como los movimientos respiratorios del paciente. La dificultad inherente a las imágenes obtenidas mediante MRI unido al hecho de que las máquinas de CT tenían un precio más competitivo y por ello, estaban más extendidas hasta hace unos años en los hospitales, explica que las imágenes abdominales obtenidas mediante CT hayan sido más trabajadas hasta el momento para la localización del hígado.

La segmentación de hígado en imagen abdominal presenta los siguientes problemas principales:

1. El hígado es un órgano de forma y tamaño variable. Si bien existe un patrón más o menos regular en el caso de hígados sanos, existen variaciones en la forma y el tamaño en cada individuo.
2. El hígado es un órgano de tamaño apreciable que está colocado al lado de otra serie de órganos y elementos, como son los riñones, pulmones, vesícula, intestinos o pared abdominal, que, evidentemente, también aparecen en la imagen. La proximidad y la apariencia semejante con varios de ellos hace muy difícil incluso para radiólogos expertos determinar el límite que separa el hígado de los otros órganos en algunos casos. En consecuencia, el procesamiento automático para localizarlo correctamente se complica.
3. El hígado sufre ligeros desplazamientos debidos al movimiento respiratorio. Esto es relevante principalmente en las imágenes de MRI (en comparación con las de CT), que por la larga duración de la secuencia completa no puede ser adquirida en apnea. Estos movimientos en la imagen se añaden a los artefactos propios de MRI. Es necesario gestionar adecuadamente estos desplazamientos en el caso de emplear varias secuencias de MRI para su procesamiento conjunto, lo que se denomina tarea de registro.

Junto con estos problemas técnicos aparece otra serie de problemas relativos a la validación de la bondad de los métodos de segmentación. No existe ninguna base de datos pública para realizar unas pruebas con respecto a una segmentación de referencia (*gold standard* o *ground-truth*), que sea utilizada por la comunidad científica. Esto implica que cada trabajo expuesto en la literatura aporte su propio conjunto de imágenes de pruebas con su propia segmentación de referencia. Lo cual es un problema para la comparación de resultados de los diferentes métodos. De hecho, los radiólogos suelen diferir en sus segmentaciones manuales, pero además, incluso el mismo radiólogo en dos momentos temporales distintos puede realizar una segmentación manual diferente sobre la misma imagen (Göçeri et al., 2015).

Otro problema importante en el caso de la segmentación hepática, y la segmentación de imagen médica en general, son las métricas usadas para la evaluación. No existen unas métricas comunes aceptadas por la comunidad científica para determinar de manera

indiscutible cómo de bueno es un método y compararlo convenientemente con otros. Cada autor aplica el criterio que le parece más adecuado acorde a su propósito. La dispersión en las métricas y en los procedimientos de valoración ha sido objeto de estudio (Campadelli et al., 2009).

Como conclusión, la segmentación automática de hígado constituye un reto, principalmente en imágenes de resonancia magnética. Si, además, se deciden emplear varias secuencias del mismo estudio para la segmentación conjunta de las mismas, el problema de registro y combinación de la información adquiere especial relevancia. Este conjunto de problemas ha sido abordado en el marco de la presente tesis.

En los siguientes apartados se van a presentar los métodos de segmentación más populares y extendidos junto con una selección de los trabajos más representativos que los adaptan para segmentación hepática. Dichos métodos podrían ser extendidos a cualquier otro órgano de estudio.

4.3. Métodos de segmentación

Existen multitud de técnicas de segmentación en la literatura cuyo uso depende mucho del objeto de la segmentación y la tecnología de adquisición, con complejidad dispar, desde sencillas umbralizaciones hasta técnicas variacionales para minimizar funcionales de energía mediante los que se pueden representar regiones de la imagen. Existen bastantes referencias que presentan un compendio de los métodos y técnicas más usados (Pal and Pal, 1993), (Clarke et al., 1995), (Pham et al., 1998), (Sharma and Aggarwal, 2010).

No existe un método único para resolver el problema de segmentación sino que la elección correcta del mismo dependerá de cada caso. De hecho, en base a la complejidad del problema, muchas veces es necesaria la combinación de varias técnicas.

(Luo et al., 2014) recoge recientemente una breve clasificación de los principales tipos de métodos automáticos para segmentación de hígado propuestos en los últimos años. No hay una manera única de categorizar los métodos de segmentación, sino que dependiendo del criterio se puede agrupar como propone este trabajo, como métodos basados en nivel de gris, en textura o en estructuras; o mediante otros conceptos, como supervisados / no supervisados; basados en píxel o en región; basados en características locales o globales; de machine learning, etc. Realmente es difícil encontrar una clasificación única. En (Göçeri et al., 2015) también se muestra una descripción actualizada de los métodos más extendidos de segmentación aplicados a segmentación hepática con el valor añadido de que dichos métodos han sido implementados y se presenta una comparativa de resultados acorde a la metodología de evaluación y métricas establecidas en dicho trabajo.

En este capítulo, se describen diversos métodos de segmentación que han sido categorizados de la siguiente manera:

- Basados en umbralización o *thresholding*

- De detección de bordes
- Crecimiento de regiones
- Métodos estadísticos
- Basados en modelos a partir de atlas
- De clasificación o *clustering*
- Basados en conocimiento a priori
- Particionamiento gráfico (*graph cuts*)
- Modelos deformables

4.3.1. Umbralización o *thresholding*

Aproximación teórica

Se trata de uno de los métodos de segmentación más populares, antiguos y relativamente sencillos en procesamiento de imagen.

Los métodos de umbralización o *thresholding* son basados en histograma. El conjunto de valores de gris de todos los píxeles agrupados por la cuantificación del número de píxeles que poseen el mismo valor de nivel de gris es lo que se conoce como histograma. Los picos y valles de dichos conjuntos de valores se usan para determinar los grupos o clústeres en los que los píxeles podrían clasificarse y son los denominados umbrales. Los píxeles cuyo valor de gris presente un valor comprendido entre dichos valores se asocian a su grupo correspondiente. A veces, se realiza un proceso recursivo para volver a separar cada clúster en sub-clúster, y distinguir picos y valles en dicha subregión. Es decir, pueden basarse en características globales de la imagen (histograma de la imagen completa) o en características locales (si se selecciona un umbral por región o vecindad determinada). Se muestra un ejemplo de histograma en la figura 4.1.

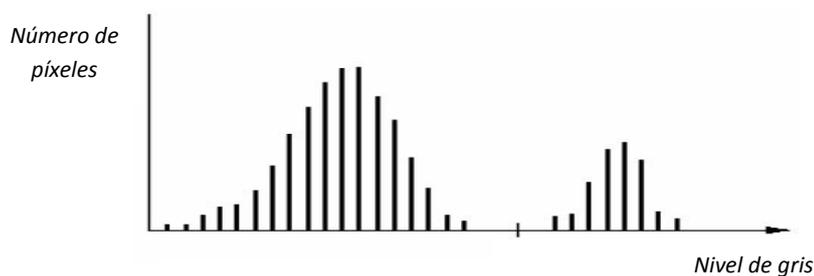


Figura 4.1. Ejemplo de histograma con un umbral claro

Las aproximaciones más sencillas determinan exclusivamente un nivel de umbral que separa los píxeles en dos grupos, como pertenecientes o no a la clase, y transforma la imagen inicial en niveles de gris en una imagen binaria. Hay variaciones en las que se pueden indicar 2 o más umbrales (determinados por los picos o valles del histograma) para separar en varias clases. Algunos métodos son adaptativos y los umbrales pueden variar en

las diferentes zonas de la imagen en base al análisis de histograma local que el algoritmo hace de las inmediaciones de un píxel, pudiendo determinarse el tamaño de esta vecindad. Un método conocido es el de Otsu, en el que el algoritmo propone un valor de umbral óptimo para maximizar la separación entre clases (Otsu, 1979). Una mejora del mismo requiere de la interacción del usuario que indica un punto o semilla perteneciente al objeto que se desea segmentar y otro punto que no pertenece. Es interesante una revisión de las diferentes variaciones propuestas sobre los métodos de umbralización (Pal and Pal, 1993).

Normalmente este tipo de métodos no son muy robustos en imágenes complejas o con poco contraste entre los elementos presentes, ya que la determinación de los umbrales de separación no es clara y es decisiva para la correcta segmentación de la imagen. Para usar este tipo de métodos es fundamental la aplicación de una fase de pre-procesamiento que pueda ecualizar, uniformizar la iluminación, eliminar ruidos (filtros de media inicial), eliminar algunos de los artefactos para aumentar el contraste entre los órganos adyacentes y disminuir la zona más borrosa o el efecto de volumen parcial.

Las imágenes médicas son bastante complejas, como se ha explicado en el capítulo 2, es por ello que este tipo de métodos se presentan como insuficientes para realizar una buena segmentación de la imagen y extraer los elementos relevantes. Se pueden emplear como métodos auxiliares para la localización de puntos semillas o segmentaciones muy burdas combinados con métodos de segmentación más potentes. Computacionalmente hablando, estos métodos son bastante rápidos en cuanto a que sólo necesitan un análisis sobre los valores de gris de los píxeles que constituyen la imagen, sobre todo en las versiones más sencillas.

Aplicación a segmentación hepática

Son pocos los casos de segmentación de imagen médica, en general, y hepática, en particular, en los que se haya empleado el histograma como único elemento para discriminar las regiones, dadas las limitaciones de estos métodos en cuanto a bondad de resultados. Se ha identificado un método que se antoja extremadamente sencillo, empleando segmentación basada en umbralización adaptativa, operadores morfológicos y un ligero procesamiento previo consistente en un sencillo filtro de media para eliminar el ruido (Antonidoss and Kaliyamurthie, 2014). En base a las imágenes que se muestran en el trabajo, todas de CT, la segmentación parece razonable, pero no se describe ningún resultado de unas pruebas metodológicas con una segmentación de referencia y unas métricas establecidas.

Es conocida, como ya se ha explicado previamente en el capítulo 2, la dificultad en la segmentación de imágenes de MRI en comparación con las de CT. Las imágenes abdominales, y principalmente las de MRI, presentan una iluminación nada uniforme, además de muy poco contraste entre los órganos adyacentes. De esta manera la determinación del umbral óptimo capaz de discriminar entre dos elementos diferentes es una tarea difícil y es necesaria la aplicación de técnicas avanzadas. Se muestra en la figura 4.2 una imagen abdominal de MRI binarizada.



Figura 4.2. Ejemplo imagen abdominal binarizada

En ella se pueden observar varias de las limitaciones que presenta la segmentación por técnicas de umbralización en las imágenes de MRI. Dado que se trata de una imagen en nivel de gris, se impone un valor de corte, aquellos píxeles que superan dicho valor se etiquetan a '1' (blanco), mientras que los que tienen valor inferior se etiquetan a '0' (negro). Los tejidos de niveles de gris similares no son capaces de ser distinguidos, como es el caso del hígado y el otro elemento adyacente que están enmarcados en el círculo rojo. Además el efecto de iluminación no uniforme hace que no sea fácil establecer un umbral que permita que sólo todos los píxeles pertenecientes al hígado lo superen sin que vaya aumentando el número de elementos no deseados en la imagen.

El análisis del histograma para determinación de umbrales se suele emplear para la localización de semillas iniciales que sean el punto de partida de técnicas de clusterización (*clustering*), de crecimiento de regiones, o de grafos (*graph cuts*), o que puedan llevar a una segmentación burda que posteriormente sea refinada por técnicas más complejas (Rusko and Bekes, 2007).

4.3.2. Detección de bordes

Aproximación teórica

Los bordes son entendidos como los límites o zonas de transición entre los elementos presentes en las imágenes. Si se detectan los bordes de los objetos, se considera como región segmentada la ubicada dentro del contorno cerrado que delimita el borde. Los bordes identificados en ocasiones están desconectados. Para segmentar un objeto a partir de una imagen es necesario que los bordes formen figuras cerradas.

En una imagen de nivel de gris, las regiones a segmentar se caracterizan por la uniformidad en el nivel de intensidad, de manera que los bordes son aquellos puntos en los que hay abruptas transiciones en dicho nivel de gris, lo que se traduce en un gradiente de valor elevado, y menor energía (este concepto se explicará con la descripción de los contornos activos).

Los operadores o filtros más populares para la detección de los bordes son los de Sobel (Sobel and Feldman, 1973), Prewitt (Prewitt, 1970) o el operador de Laplace (Marr and Hildreth, 1980). Pueden detectar bordes reales o cambios considerados bordes en zonas de ruido, lo que no es estrictamente correcto. No todos ellos responden igual a la presencia de bordes en una u otra dirección o al efecto del ruido. El operador de Laplace, por ejemplo, es muy sensible al ruido. Otra variación muy popular por sus buenos resultados es el detector de bordes de Canny (Canny, 1986). Se puede apreciar un ejemplo tipo en la figura 4.3.

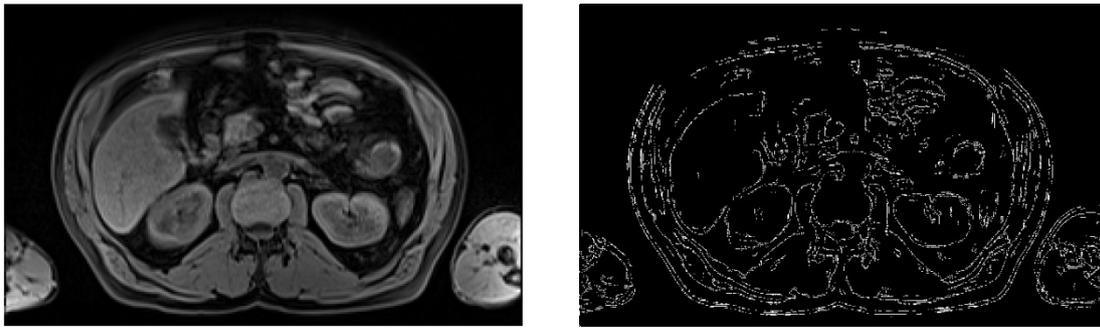


Figura 4.3. Detección de bordes mediante el método de Canny

Los operadores buenos para detectar bordes en la imagen debieran ser invariantes a la rotación, al mismo tiempo que pudieran ser configurados en varias escalas, y que actuaran igualmente bien en bordes difusos o en bordes muy precisos, dependiendo de la imagen. En el caso de imágenes borrosas, con amplia zona de transición entre una región y otra, un buen método debe indicar el borde en la zona intermedia que corresponde a la zona de transición, minimizando la confusión y la múltiple detección.

Este tipo de métodos de segmentación no son válidos en imágenes complejas. En imágenes relativamente sencillas con pocos elementos a localizar y con condiciones de iluminación uniformes, con contraste entre los elementos, la detección de bordes mediante uno de los métodos descritos es razonable. A continuación, se deben aplicar técnicas de procesamiento morfológico que permitan unir y cerrar contornos, eliminar pequeños partículas remanentes y asumir como región segmentada la contenida en esos límites.

En el caso de imágenes complejas, como es el caso de las imágenes médicas, la técnica de detección de borde no es lo suficientemente robusta para delimitar regiones, ya que, probablemente, muchos bordes quedarán abiertos y sin unir y algunos elementos de ruido serán identificados como falsos bordes. Estos métodos deben complementarse con otros para poder proporcionar buenos resultados en la segmentación. Sin embargo, la detección de bordes es una técnica muy extendida y empleada en muchos otros métodos, de manera indirecta. A modo de ejemplo, los contornos activos o *snakes* emplean funciones detectoras de bordes para estimar las denominadas fuerzas de la imagen, y evolucionar acorde a dichos bordes y al resto de fuerzas identificadas. En este caso se localizan los límites de las regiones mediante una función detectora de bordes basada en gradiente. Más detalles sobre estos métodos en el apartado 4.2.9 sobre los modelos deformables.

Un método que se puede considerar una variación compleja sobre los métodos de detección por borde, son las técnicas de watershed. Estas técnicas pretenden generar una especie de mapa de niveles (tipo mapa cartográfico) que emularían el flujo de agua en una cuenca hidrográfica representada por niveles de gris. Sobre dicha imagen de gris se calculan gradientes, esquinas y otros elementos representativos que contribuyen a delimitar las zonas resultantes de la segmentación.

Aplicación a segmentación hepática

No se han encontrado trabajos en los que se haya empleado de manera exclusiva la detección de bordes para la localización del hígado en la imagen abdominal, bien en CT o en MRI. Se han identificado estas técnicas como elemento complementario en otros métodos, donde se realiza la segmentación en la secuencia T1 con contraste, y se combinan varios métodos de segmentación, como los mencionados de detección de bordes, análisis de histograma, operadores morfológicos, y un contorno activo que evoluciona hasta llegar a la segmentación final (Platero et al., 2008b).

Se ha localizado un trabajo reciente donde se emplea la técnica de watershed para segmentación automática de hígado sobre imágenes de MRI. En dicho trabajo emplea transformada watershed y particiones estocásticas. Modifica el algoritmo con puntos de control para evitar sobre-segmentación en la segmentación con watershed. Las regiones finales son sometidas a una fase de procesamiento final mediante un clasificador basado en umbralizado de intensidad por regiones, lo que proporciona el hígado final segmentado. Este trabajo presenta buenos resultados tras una validación detallada y mediante métricas establecidas. (López-Mir et al., 2014). Sus resultados serán comparados con los obtenidos por los métodos propuestos en esta tesis.

4.3.3. Crecimiento de regiones (*region growing*)

Aproximación teórica

Es una técnica muy utilizada, aunque puede ser lenta dependiendo del objeto de la segmentación ya que es bastante costosa computacionalmente hablando. Consiste en la identificación (manual en la mayoría de los casos) de una serie de semillas, constituyendo cada semilla un objeto a segmentar. Las regiones van creciendo alrededor de las semillas como resultado de la comparación de los valores de gris de los píxeles vecinos sin clasificar con alguna de las regiones que van creciendo. La medida de similitud es la diferencia entre la intensidad del píxel con la media de esa región. Cada píxel es asignado a la región con la que tiene más similitud. El proceso se repite iterativamente hasta que todos los píxeles han sido clasificados. Es altamente dependiente de la elección inicial de las semillas (Adams and Bischof, 1994).

Los métodos basados en crecimiento de regiones consisten en, una vez identificada la semilla inicial o conjunto de las mismas, evaluar ciertas características de los píxeles vecinos y en base a unos criterios establecidos decidir si se anexionan o no a la región a

reconstruir. Los criterios que evalúan la pertenencia o no de cada píxel a la región pueden ser diversos, siendo el valor de nivel de gris uno de los más usados. El problema de los métodos automáticos, que son los que están siendo evaluados aquí, es la dificultad en la estimación automática de la ubicación de los puntos semillas a partir de los cuales comenzar a iterar para realizar la reconstrucción. Por ello, los autores plantean todo tipo de métodos para localizar con seguridad esta semilla ya que el éxito final depende de este primer paso. En los métodos interactivos, si el usuario selecciona la semilla, se están eliminando riesgos en la segmentación, al menos, el origen ha sido el adecuado.

Otro de los problemas de estos métodos es que hay que establecer una condición de parada, que bien puede ser cuando ya no haya ningún píxel adyacente que verifique las condiciones establecidas y la distribución permanece en estado estacionario después de un número de iteraciones determinado. Estas técnicas se encuentran limitadas cuando la región no es homogénea y existen discontinuidades en la misma que impiden que zonas del mismo elemento queden conectadas. En aquellas imágenes con iluminación no uniforme debida a los artefactos de la captura u otras causas, requiere de alguna etapa de pre-procesado que uniformice las regiones.

Existen muchas variaciones del método. Con esta técnica se tiene en cuenta la proximidad espacial así como la probabilidad de que pertenezca a una clase por aproximación en valores de intensidad.

Aplicación a segmentación hepática

Como ejemplo de aplicación de estas técnicas a segmentación hepática, se ha identificado un método de crecimiento de regiones en 3 dimensiones que realiza posteriormente un post-procesado que permite seleccionar el hígado en imágenes de CT (Rusko and Bekes, 2007). Se extrae el histograma de la imagen y se analiza para obtener la distribución de intensidades del hígado y en base a los límites superior e inferior estimados localiza la región semilla a partir de la cual va a realizar el crecimiento de la región. Este crecimiento se hace en base al valor de intensidad. Previamente se ha identificado la región que une las zonas inferiores de los pulmones (conocimiento a priori), y actúa de bloqueo para que la región del hígado no se siga extendiendo donde no le corresponde, derivando en una sobre-segmentación. Se emplean filtros posteriores para suavizar el contorno, eliminar alguna impureza y asumir los vasos como pertenecientes al hígado.

Como se ha indicado, es posible realizar un crecimiento de regiones en base a la evaluación de otros criterios que no sean el valor de intensidad. Existen trabajos donde se aplican varios métodos conjuntamente, también sobre imágenes de CT. Susomboon et al. propone la siguiente secuencia. Primero se asigna a cada píxel un valor de probabilidad de pertenencia a una de las clases especificadas que se corresponden con los elementos presentes en la imagen, y son preliminarmente clasificados como tejido o no tejido. A continuación se emplean como descriptores las características extraídas de las matrices de co-ocurrencia o de Haralick (Haralick, 1979) de las que se ha hablado en el capítulo 3, y a través de un árbol de clasificación y regresión se obtiene la probabilidad de que se trate de

hígado. Aquella región que presenta la mayor probabilidad de ser hígado es la seleccionada como semilla, a la que se van anexionando los píxeles vecinos que cumplen los criterios de pertenencia. Se coteja con la segmentación realizada en los cortes adyacentes de la secuencia, entendiéndose que entre dos cortes en el eje z debe haber al menos un 85% de solape ya que no hay demasiado desplazamiento entre ellos (Susomboon, 2007).

En otro trabajo localizado en la literatura se desarrolla una técnica de crecimiento de regiones basada en textura y en 3 dimensiones sobre imágenes de CT. Como elemento de evaluación de pertenencia o no a la clase hígado se utilizan características de textura obtenidas de la matriz de co-ocurrencia. Tanto la semilla para comenzar la evaluación como la condición de parada se obtienen automáticamente (Gambino et al., 2010).

4.3.4. Métodos estadísticos

Aproximación teórica

Este conjunto de métodos de segmentación consiste en utilizar técnicas estadísticas para clasificar cada píxel como perteneciente a una clase determinada, esto es, una región diferente en la imagen, sea tipo de tejido, órgano o lesión. La aproximación estadística se puede aplicar a cualquier característica extraída de la imagen, no sólo al nivel de intensidad de cada píxel, sino también, a otras características, como las basadas en textura o forma, lo que constituye el conjunto de métodos de modelos estadísticos de forma, de los que se habla en el apartado 4.3.9.3, dentro de los modelos deformables.

Los **estadísticos de primer orden** son la media y la varianza, es decir, el valor medio de una característica de un conjunto de píxeles y el cuadrado de la desviación que existe entre esa media y el valor de cada característica asociada al píxel. Para poder calcular estos valores es necesario establecer un vecindario de píxel sobre el que se hará el cálculo, es decir, la ventana $N \times N$ que rodea al píxel analizado.

Los modelos de **campos aleatorios de Markov** (MRF – Markov Random Fields) se han incluido brevemente en el capítulo 3 ya que pueden ser interpretados como descriptores. Sin embargo también puede ser interpretado como un método de segmentación, o siendo más precisos aún, son un modelo estadístico que puede ser usado dentro de otros métodos de segmentación. Los MRF modelan las interacciones espaciales entre vecinos o píxeles cercanos. Los modelos de Markov son modelos probabilísticos que emplean la correlación entre los píxeles en un vecindario para decidir la región objeto. De esta manera, no se emplea únicamente el valor de un píxel para determinar si es objeto o fondo, sino que se emplea su valor y el de su vecindario.

Estas correlaciones locales proveen un mecanismo para modelar una variedad de propiedades de la imagen. En el tratamiento de imágenes médicas, se utilizan frecuentemente para considerar el hecho de que la mayoría de los píxeles pertenecen a la misma clase a la que pertenecen sus píxeles vecinos. Esto implica que bajo la asunción del

MRF, cualquier estructura anatómica que consista de un solo píxel tiene una probabilidad muy baja de ser tal.

Los MRF son incorporados frecuentemente en los algoritmos de segmentación por agrupamiento o *clustering*, como k-means. La segmentación es obtenida maximizando la probabilidad *a posteriori* de la segmentación a partir de los datos de la imagen. A un k-means para clasificación no supervisada (se describe en apartado 4.3.5.1) le añade robustez frente al ruido.

Una dificultad asociada con los modelos MRF es la selección apropiada de los parámetros que controlan la fuerza de las interacciones espaciales. Unos valores muy altos pueden resultar en una segmentación excesivamente suave y una pérdida de los detalles estructurales. Además, los métodos MRF normalmente requieren muchos recursos para ser ejecutados. A pesar de estas desventajas, los MRF son ampliamente utilizados.

Aplicación a segmentación hepática

En (Zhang et al., 2010) se presenta un método de segmentación de hígado sobre imágenes de CT basado en modelo estadístico de forma integrado con una estrategia de detección de superficie óptima. La inicialización se realiza a través de perfiles de intensidad y gradiente. El modelo de forma se adapta al contorno del hígado en cada caso mediante teoría de grafos para detección de la superficie óptima. Presenta buenos resultados.

Recientemente, se ha publicado un trabajo que aplica un nuevo método de segmentación basado en modelo estadístico de forma sobre imágenes de CT. Las características condicionantes del modelo de forma son más estrictas o no en base a un modelo de error generado. El algoritmo tiene tres fases. Inicialmente extrae las características de forma condicionante mediante el método probabilístico de máximo a posteriori. A continuación, realiza la estimación previa de forma a través del modelo de forma estadístico basado en level set con el modelo de error integrado; y a continuación, se realiza una segmentación basada en grafos. Sobre imágenes de CT los resultados presentados son buenos. (Tomoshigea et al., 2014)

4.3.5. Agrupamiento o Clustering

Aproximación teórica

Se entiende como agrupamiento, clusterización o *clustering* (su conocido término en inglés) el proceso de asignación de cada uno de los píxeles presentes en una imagen a una de las estructuras con respecto a la que presenta mayor grado de similitud. En este sentido se puede interpretar que los métodos de *clustering* además de segmentar (localizar las estructuras) también son capaces de identificarlas, lo que se interpreta como clasificación.

El fino matiz que separa las técnicas de *clustering* de las de clasificación es que en las primeras no se sabe a priori las características de los grupos que se constituirán (a veces no se conoce ni el número de grupos), mientras que en los métodos de clasificación en sí se

conocen con certeza las propiedades discriminantes de las clases. Los métodos de *clustering* se conocen como no supervisados, en general, y los métodos de clasificación son, en principio, supervisados.

Se recogen en este apartado los dos tipos, ya que ambos son considerados técnicas de segmentación.

4.3.5.1. Clustering no supervisado

Algunos **métodos no supervisados** no conocen a priori los rangos de valores de las características que van a definir una región, sino que dadas unas semillas o puntos iniciales, asumidos como centroides de una región se irá asignando cada píxel de la imagen a una región en base al mayor grado de similitud que presenta dicho píxel con el centroide de la misma (en una serie de características determinadas, que pueden ser nivel de gris, textura...). La medida de similitud suele ser estimada calculando una distancia, siendo la más usada la distancia Euclídea. A medida que nuevos píxeles se van anexionando a las regiones, los centroides que representan dichas regiones se van actualizando dinámicamente y el proceso de asignación sigue iterativamente hasta que ya no se produce ninguna asignación nueva y los píxeles permanecen inamovibles en sus regiones. Este es el conocido algoritmo **k-means** (Lloyd, 1982). El valor de k se define inicialmente y representa el número de grupos. Este valor puede ser elegido aleatoriamente o en base a algún criterio conocido. La calidad de los resultados depende fuertemente del número de grupos indicado en que se desea dividir la imagen, por lo que, si se conoce a priori, es conveniente fijarlo. Los centroides iniciales o puntos semillas son elegidos aleatoriamente por la mayoría de las implementaciones software aunque pueden ser asignados por otros métodos. Es de suponer que puntos de inicio diferentes ocasionan, en muchos casos, resultados diferentes. La solución puede quedarse estancada en mínimos locales.

Han surgido variantes que intentan minimizar estas limitaciones. Una variación extendida de este algoritmo es el **Fuzzy C-means**, la variación en lógica fuzzy del algoritmo, donde a cada píxel se le asigna un grado de pertenencia a cada grupo en vez de ser asignado únicamente a una clase (Bezdek et al., 1984). El resto de la operativa es idéntica.

Los **Gaussian Mixture Models (GMM)** (Reynolds and Rose, 1995) o modelos de mezclas gaussianas también se pueden considerar dentro del grupo de clustering con una base de funcionamiento estadística. Estos métodos parten de la premisa de que en la imagen existen N distribuciones normales gaussianas que pueden ser separables con su media y su desviación estándar. En base a estos datos, cada uno de los píxeles de la imagen es asignado a una de las distribuciones, es decir, a una de las clases. Por lo tanto, las N clases habrían sido segmentadas. Se muestra en la figura 4.4 un ejemplo de una representación de GMM.

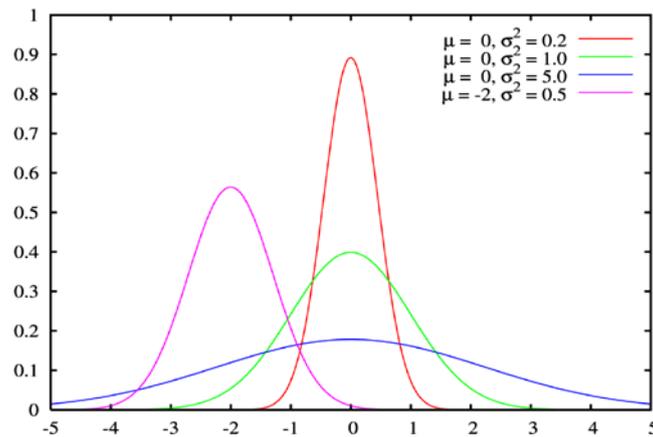


Figura 4.4. Ejemplo de GMM

Otro método de agrupamiento muy utilizado es el **algoritmo EM** (Expectación – Maximización o *Expectation – Maximization*) (Dempster et al., 1977). El algoritmo EM es un método general para encontrar el estimador de máxima verosimilitud de los parámetros de una distribución de probabilidad. Realiza dos tareas principales: *expectation* y *maximization*.

El problema ante un planteamiento de clustering o agrupamiento no supervisado es que no se sabe de qué distribución viene cada dato y no se conocen los parámetros de las distribuciones existentes (las clases o grupos).

El algoritmo EM empieza estimando los parámetros de las distribuciones y los emplea para calcular las probabilidades de que cada objeto pertenezca a un clúster. Dichas probabilidades las usa para estimar de nuevo los parámetros de las distribuciones, hasta converger.

El cálculo de las probabilidades de las clases o los valores esperados de las clases es la parte de *expectation*. El paso de calcular los valores de los parámetros de las distribuciones es *maximization* y consiste en maximizar la verosimilitud de las distribuciones dados los datos. Para estimar los parámetros, hay que considerar que únicamente se dispone de las probabilidades de pertenecer a cada clúster y no de los clústeres en sí.

Se repiten las dos acciones, *expectation* y *maximization*, iterativamente hasta alcanzar convergencia. El algoritmo tiende a converger aunque no se puede garantizar que sea a un mínimo global y no local.

Si se compara el método *k-means* con el método EM, el método EM puede proporcionar mejores resultados. El método *k-means* presenta dos limitaciones principales: a) la métrica de asignación de cada píxel es la distancia euclídea; y b) la medida de dispersión entre grupos es la varianza. Además, es necesario definir el parámetro de entrada *k*, que es el número de grupos. Una indicación errónea de este valor (que a veces no se puede conocer a priori con exactitud) redundará en una segmentación inadecuada. La tendencia del *k-means*

a producir grupos con tamaños parecidos lleva a obtener malos resultados, mientras que EM se beneficia de la distribución gaussiana presente en el conjunto de datos.

Los métodos de clustering no supervisados requieren a veces de procesamiento posterior, ya que pueden contener píxeles erróneamente asignados.

4.3.5.2. Clustering supervisado o clasificación

Los **métodos de clustering supervisado o clasificación** son aquellos que aprenden las características de similitud de las clases en una fase previa de entrenamiento en base a ejemplos proporcionados de cada clase, y son capaces de generalizar a otros nuevos casos y clasificarlos correctamente. Este grupo de métodos son desarrollados en las disciplinas de *machine learning* en inteligencia artificial, y se trata de un campo de trabajo inmenso.

Se muestra a continuación un breve resumen de los métodos de clasificación más extendidos para segmentación de imagen médica, no siendo objeto de esta tesis trabajar con clasificadores. Los desarrollos propuestos en esta tesis se basan en otras técnicas de segmentación, en contornos activos, que serán descritos con gran detalle en el apartado 4.2.9. Bien es cierto que el uso de clasificadores es un método muy válido para la segmentación del hígado, combinando una selección de características extraídas de las imágenes y un clasificador adecuado, como se mostrará en los ejemplos citados.

El **método kNN** (Fix and Hodges, 1951), como todos los métodos de clasificación supervisados, precisa de acceso a parejas de ejemplo y clase conocidas y de otra serie de imágenes para prueba y clasificación. En la fase de entrenamiento se proporcionan una serie de ejemplos pertenecientes a diferentes clases, de manera que se pueden visualizar los agrupamientos de clases en el espacio de características n -dimensional, siendo n el número de elementos del vector de características descriptoras. Ante una nueva entrada de la que se extraen dichas características, el clasificador le asigna la clase más común a la que pertenecen sus k vecinos más cercanos, siendo k un valor típicamente pequeño. Los vecinos más cercanos se estiman mediante el cálculo de una distancia, habitualmente la distancia Euclídea, y estos vecinos más cercanos son aquellos con los que se presenta menor distancia. El funcionamiento se puede apreciar en la figura 4.5:

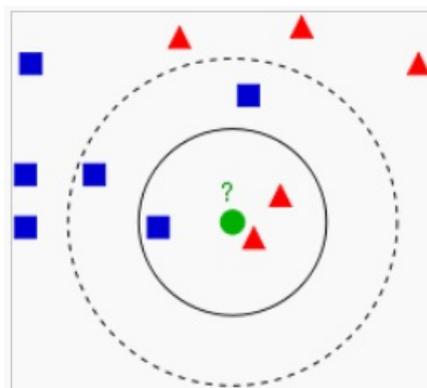


Figura 4.5. Ejemplo de asignación de kNN

En este caso, se pretende clasificar el caso verde. Para ello, se calcula la distancia euclídea a los casos de entrenamiento, y en este caso con $k=3$, los 3 vecinos más cercanos son dos rojos y uno azul. La clase más frecuente de estos tres vecinos es la roja, con lo que el caso de entrada es clasificado como perteneciente a la clase roja.

Este algoritmo puede verse limitado por la calidad de los casos de entrenamiento que además deben de ser representados por un vector de características relevantes. Demasiados elementos en el vector de características, sin reducir la dimensionalidad de alguna manera, puede implicar que algunos sean irrelevantes, y puede ocasionar la determinación errónea de los vecinos más cercanos y por tanto, su clasificación.

Las **redes neuronales artificiales** (Artificial Neural Networks o **ANN**) basan su esquema de funcionamiento en las redes neuronales humanas y se representan como un conjunto de neuronas conectadas. La activación de uno o más elementos y la transmisión de la señal a través de las conexiones hace que ante determinada entrada, la respuesta sea una concreta, y diferente si se cambia el tipo de respuesta de las neuronas (lineal, sigmoideal) conectadas con diferentes pesos. Además, se pueden diseñar conjuntos de ANN con muchas capas, siendo obligatorias la de entrada y salida, y pudiendo añadirse tantas capas ocultas como se desee, de manera que la complejidad puede ser enorme. Algunos tipos de red neuronal pueden ser: perceptron multicapa, backpropagation, redes Hopfield, y una larga lista de diseños diferentes.

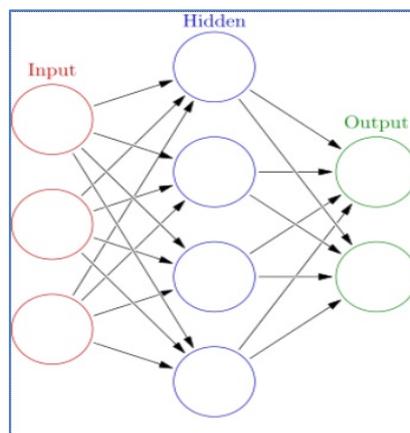


Figura 4.6. Red neuronal con capa de entrada, capa de salida y una capa oculta

Una desventaja de los clasificadores o métodos supervisados es la necesidad de la interacción manual para obtener los datos de entrenamiento. Los conjuntos de entrenamiento pueden ser adquiridos para cada imagen que requiera segmentación, pero esto puede requerir de mucho tiempo y esfuerzo. Además, el uso del mismo conjunto de entrenamiento debe ser adecuado, amplio y variado en todo el rango de clases ya que de lo contrario, puede provocar resultados incorrectos. Esta dificultad en el modelado de problemas difícilmente separables ha sido revisitado en estudios recientes, como **Deep Learning** (Bengio et al., 2013), método de aprendizaje profundo para clasificación que está teniendo una buena acogida en la comunidad científica.

Otro ejemplo de clasificador no lineal ampliamente extendido son los **Support Vector Machine** (SVM) (Cortes and Vapnik, 1995). Estos métodos están relacionados con problemas de clasificación y regresión. Se trata de un método de clasificación supervisado, ya que es necesario proporcionar una serie de ejemplos de entrenamiento etiquetados como pertenecientes a una clase u otra. Se entrena el SVM y se genera el modelo de clases en el hiperespacio. Intuitivamente, una SVM es un modelo que representa a los puntos de muestra en el espacio, separando las clases en un espacio lo más amplio posible. Cuando se introduce un nuevo ejemplo, son clasificados en una u otra clase en función de su proximidad. En la figura 4.7 se muestra el principio de funcionamiento de los SVM.

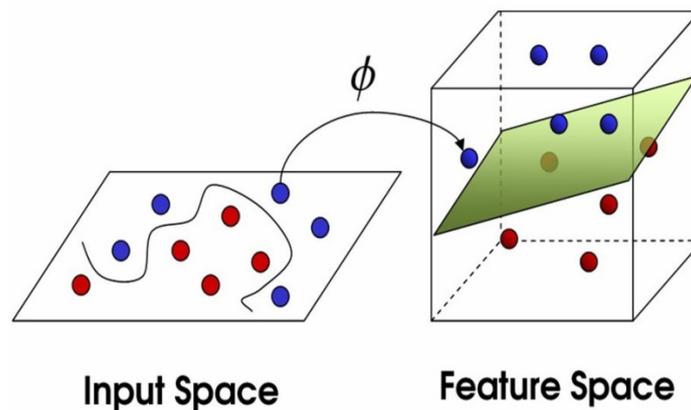


Figura 4.7. Principio de funcionamiento de los SVM

Un SVM construye un hiperplano o conjunto de hiperplanos en un espacio de dimensionalidad muy alta que puede ser utilizado en problemas de clasificación o regresión. Se emplea sobre todo en casos complejos, casos no lineales, con varias clases y donde no se ve una clara separabilidad entre ellas.

Son muchos más los métodos de clasificación existentes. Sin embargo, dado que el objeto de esta tesis no es la investigación ni la aplicación de métodos de clasificación, no se va a hacer aquí una amplia descripción de todos. Por mencionar otros métodos de clasificación supervisada muy conocidos están los **árboles de decisión** (Rokach and Maimon, 2005), o los **métodos de decisión de Bayes** (Bayes, 1984). El teorema de Bayes es un método estadístico que vincula la probabilidad de A dado B con la probabilidad de B dado A.

Un árbol de decisión es un modelo de predicción y clasificación. Dado un conjunto de ejemplos se construyen diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema. De esta manera nuevos ejemplos introducidos como entrada son clasificados.

Aplicación a segmentación hepática

Los **métodos basados en agrupamiento o clustering (no supervisados)** propuestos en la literatura para segmentación de hígado son principalmente de dos tipos, k-means y Fuzzy c-means (FCM).

En (Zhao et al., 2010) se describe una segmentación de hígado con una propuesta inicial de agrupamiento mediante FCM seguido de un procesamiento morfológico capaz de suavizar y corregir algunas regiones. A continuación, para clasificar cada región segmentada como hígado u otro se diseña una red neuronal. Las regiones que han sido clasificadas como hígado se extraen de la imagen para la reconstrucción posterior. Las técnicas de k-means suelen dar resultados aleatorios y en el caso de la segmentación del hígado en imágenes de CT, donde el nivel de gris con los órganos adyacentes es similar, a veces no es capaz de separarlos. Estos métodos de agrupamiento se suelen usar para una segmentación burda, que sirve como inicialización de un proceso de regularización más fino, como pueden ser los Contornos Geodésicos Activos o *Geodesic Active Contours* (ver sección 4.3.9.6 de este capítulo), como propone Foruzan en su trabajo, obteniendo unos resultados aceptables (Foruzan and Yen-Wei, 2013).

Dichos métodos no supervisados no son empleados como método principal de segmentación. Sin embargo, como excepción, el método FCM, al asignar a cada píxel un valor de pertenencia a una clase y dado que este valor es susceptible de verse modificado acorde a la evolución de los grupos, da lugar a un agrupamiento más correcto, y por ende, una segmentación más precisa. En (Yuan et al., 2010), se propone sobre imágenes de resonancia magnética (MRI) un método de segmentación que combina *Fast Marching*, técnica derivada de los level set (ver sección 4.3.9.2 de este capítulo) para obtener una segmentación burda, que a continuación es refinada por FCM. El principal problema de este método es que el resultado es fuertemente dependiente de la inicialización.

En cuanto a los **métodos de clasificación supervisada**, existe una amplia muestra de trabajos en los que se han usado para análisis de imagen médica sobre cualquier órgano, incluido el hígado en imagen abdominal. La casuística es variada ya que cada trabajo elige la combinación de descriptor, o características visuales de la imagen, y de clasificador, acorde al caso concreto de aplicación. En el caso de aplicación hepática, cabe mencionar a (Danciu et al., 2012) que propone el uso de los coeficientes de la Transformada Discreta del Coseno (DCT) como características de la imagen que conforman la entrada a un *Support Vector Machine*. Posteriormente se suavizan dichos resultados con operadores morfológicos.

En otro trabajo, se propone un software de ayuda al diagnóstico para localizar tumores basada en características de textura, de primer y de segundo orden, obtenidas a partir de la matrices de coocurrencia, y realiza una clasificación mediante ANN, consiguiendo clasificar 4 patologías diferentes (Zhang et al., 2008).

Con el mismo propósito que esta tesis se ha identificado un trabajo en el que emplean redes neuronales para segmentar el hígado en una imagen abdominal (Masoumi et al., 2012). En primer lugar se aplica una fase de pre-procesamiento para reducir el ruido preservando bordes; a continuación, se utilizan redes neuronales para la extracción de características de hígado, que luego son empleadas en una técnica watershed. Las redes neuronales y las características extraídas también se usan para evaluar la calidad de la segmentación y se van ajustando los parámetros automáticamente.

Finalmente como ejemplo ilustrativo de aplicación con *Support Vector Machines* se ha identificado un trabajo donde se segmenta el hígado mediante dicho SVM, siendo la entrada un vector de características de textura, definido por los coeficientes de Wavelets, descritos en el apartado 3.5.4.3 del capítulo de descriptores (Luo et al., 2009). Se recoge la aplicación de algunas técnicas de *machine learning* al caso particular de segmentación de hígado sobre imágenes tanto de CT como algunas de MRI en este reciente trabajo (Punia and Singh, 2013).

4.3.6. Conocimiento a priori

Aproximación teórica

Los métodos basados en conocimiento a priori son aquellos en los que se hace uso de información conocida previamente del órgano o elemento a segmentar. Esta información puede ser diversa, como por ejemplo, que determinado órgano se encuentre siempre en la misma zona aproximada de la imagen (lo cual puede ayudar a eliminar por posición otros resultados de la segmentación), que se encuentre flanqueado por determinados elementos, lo cual puede ser útil para evitar fugas como en el caso de métodos de crecimiento de regiones, u otra información existente. De hecho, la generación de modelos se puede considerar como un conocimiento a priori.

Si dicha información está disponible, es habitual usarla, ya que en los métodos automáticos, en los que se persigue la ausencia de interacción por parte del usuario, puede ser muy útil para, por ejemplo, elegir las semillas iniciales en métodos de crecimiento de regiones, para inicialización de contornos, u otros elementos críticos para el éxito de la segmentación final.

Aplicación a segmentación hepática

Schmidt propone un método donde se emplea conocimiento a priori. En él aplica una serie de reglas a través de las cuales se extraen las diferentes estructuras presentes en la imagen abdominal, entre ellas el hígado. Este conocimiento a priori es múltiple, referido tanto a nivel de intensidad como a características geométricas y a las relaciones existentes con los órganos adyacentes. Finalmente, se selecciona una semilla del hígado y mediante crecimiento de región se segmenta, quedando los límites de este crecimiento previamente establecidos por algunos de los órganos que han sido detectados con anterioridad (Schmidt, 2007).

Muchos trabajos emplean conocimiento a priori aunque no sea referido como tal. El conocimiento a priori se puede incluir de muchas maneras. Por ejemplo, Rusko y Bekes emplean la línea que une la parte inferior de los pulmones como elemento bloqueador de fugas en el crecimiento de regiones (Rusko and Bekes, 2007).

En otro trabajo tras la segmentación del hígado en una secuencia 2D, el contorno de la región resultante se convierte en la inicialización del siguiente corte del estudio (Oliveira et al., 2011). Esto se puede hacer ya que se conoce a priori la resolución en el eje z de todos los cortes del estudio y ésta es pequeña, entre 2 y 3 mm. Además se sabe que la forma del hígado es suave y redondeada, no existiendo cambios abruptos entre cortes consecutivos. Este hecho permite suponer que entre dos cortes próximos las segmentaciones sufren variaciones graduales, y que, por tanto, la solución de un corte está razonablemente próxima a la solución del siguiente corte. Adicionalmente, la ubicación de la inicialización próxima a la solución contribuye a que se minimicen los errores de convergencia en muchas de las técnicas que se puedan emplear.

4.3.7. Basada en modelos (a partir de atlas)

Aproximación teórica

La segmentación basada en modelos surge por varias limitaciones experimentadas del uso de otras técnicas anteriores. La localización de elementos en imágenes con mucho ruido o donde éstos estén parcialmente ocultos es difícil y son necesarios métodos que, de alguna manera, extrapolen datos en base a determinada información. Esta información disponible puede ser la forma global del objeto a segmentar. De esta manera, se puede detectar el objeto de interés en su totalidad, aunque no todos los píxeles se hayan segmentado correctamente. Se dispone de un patrón de forma o plantilla, siendo ésta más o menos rígida.

Esta técnica de segmentación es muy usada en medicina, dado que los modelos pueden absorber la variabilidad existente entre los órganos de distintos individuos. Se basa en que las estructuras u órganos de interés tienen una forma geométrica repetitiva, y se recoge formalmente en los denominados atlas. Se puede establecer un modelo probabilístico que explique la variación en la forma del órgano (y que absorba la variabilidad de la población). Todo el conocimiento a priori de que se disponga es relevante en la generación de los modelos.

La aproximación basada en atlas es aquella en la que un atlas de previa construcción representa la geometría y la relación topológica de los objetos en la región corporal de interés. Un atlas, en el mundo médico, ayuda a entender y comprender la anatomía, las relaciones entre los diferentes elementos anatómicos que componen el cuerpo humano, y visualizarlos. Los atlas deben ser alimentados continuamente en base a los estudios clínicos realizados.

Existen atlas anatómicos de diferentes órganos orientados a la actividad médica como tal, que informan sobre la forma, tejidos y vasos de los diferentes órganos. Sin embargo, a efectos del desarrollo de un software automático que determine qué es hígado, qué es un vaso y qué es un tumor, esa información no resulta directamente aplicable. Es por tanto necesaria la generación de un atlas anatómico sobre las imágenes de MRI de otra manera. En este caso, se debe generar la información necesaria para determinar si cada punto en la imagen de MRI es hígado, vaso u otro órgano (no hígado).

Los atlas anatómicos valiosos para el desarrollo de métodos automáticos de segmentación son generados a partir de información recogida del equipo médico, que va anotando imágenes de MRI. Es decir, debe existir una herramienta que permita la anotación manual, la selección por parte del usuario de ciertas regiones de interés y la introducción de etiquetas que la caractericen. El atlas es una base de imágenes médicas anotadas a partir de un protocolo médico diseñado a tal efecto, necesaria para la generación de los modelos computacionales. Estos modelos se generan en base a técnicas avanzadas, principalmente clasificadores estadísticos. Se necesita un conjunto elevado de ejemplos para poder efectuar una generalización adecuada. Esto permite generar modelos tridimensionales a partir de las imágenes médicas.

Una vez identificadas las regiones, se definen vectores de características de la región segmentada con descriptores diversos, como los mencionados en el capítulo 3, pero con la salvedad de que se recoge la descripción de la región completa. En base a este vector de características, se pueden generar modelos matemáticos que permitan inferir las características de la región segmentada.

Estas técnicas de clasificación permiten crear, en base a ejemplos, modelos para cada tejido, patologías y lesiones. Estas técnicas son capaces de absorber la variabilidad entre individuos de una misma clase, así como de maximizar las diferencias entre elementos pertenecientes a diferentes clases.

Para la obtención de estos modelos se pueden usar diferentes enfoques. El más habitual es la generación de modelos de forma y apariencia basados en niveles de gris y gradiente, que se obtienen de la evaluación de los perfiles generados al trazar la perpendicular a la superficie desde cada punto de interés o *landmark* en determinado órgano. Existen otros modelos descriptores de forma, pero son menos usados (Rousson and Cremers, 2005).

Cuando los modelos se basan en características de forma y presentan la filosofía de los contornos activos en cuanto a flexibilidad y evolución, se denominan ASM o modelos activos de forma, y serán objeto de estudio del apartado 4.3.9.3.

Aplicación a segmentación hepática

Existen muchos ejemplos en la literatura que emplean este tipo de técnicas en segmentación de imagen médica, y en particular en segmentación de hígado. Kainmüller propone una

segmentación de hígado mediante ajuste de la región a un modelo estadístico de forma; y posteriormente refina los resultados con una malla deformable. Para la construcción de este modelo estadístico de forma utiliza 7.000 puntos de interés. En cada uno de estos puntos, el perfil generado en las perpendiculares a la superficie recogen las características de apariencia útiles para la generación de los modelos de hígado. Se trata un método semiautomático ya que precisa de la interacción del usuario para la identificación de algunos puntos de interés del hígado. Es un método muy costoso computacionalmente hablando, necesitando hasta 15 minutos por imagen (Kainmüller et al., 2007).

Heimann et al. propusieron un método de segmentación hepática bastante complejo que emplea 35 ejemplos de entrenamiento con 2.500 puntos de interés sobre cada uno de ellos con las que modela el hígado. Para inicializar la región emplea algoritmos genéticos que van evolucionando acorde a las características del modelo generado. Una vez se alcanza una solución estacionaria, se inicializa una malla deformable que evoluciona acorde al modelo estadístico de forma y fuerzas de la imagen representadas por los gradientes. Son métodos muy lentos, en los que se emplean unos 7 minutos por imagen (Heimann et al., 2007).

En 2009 se publicó un compendio de métodos de segmentación basados en modelos estadísticos de forma en el que, para su realización, se revisaron las publicaciones de IEEE Transactions on Medical Imaging and Medical Image Analysis de los últimos 10 años, junto con búsquedas en la web de todas las segmentaciones, analizándose más de 400 artículos (Heimann and Meinzer, 2009). En dicho estudio, se explica cómo representar la forma (puntos de interés o *landmarks* y mallas); cómo generar los modelos (alineando, pasándolos al mismo espacio de referencia, reduciendo la dimensionalidad); cómo hacer la correspondencia de forma (son métodos que relacionan la malla con la distribución de puntos y con el volumen en cada elemento del entrenamiento, hay varios métodos que también se usan para registro); habla de los Appearance Models (como los AAM o ASM, que se expondrán más adelante); se muestran 3 experiencias de segmentación de hígado en CT y habla de los inconvenientes, principalmente la recopilación de los casos para el entrenamiento y la malla de puntos que debe corresponderse bien entre todos. A veces necesita la interacción manual.

En otro trabajo se genera un atlas probabilístico a partir de una serie de imágenes de entrenamiento que son registradas mediante métodos afines y no lineales, y finalmente re-escaladas. Las transformaciones afines son aquellas que comprenden operaciones de traslación, rotación y escala, son transformaciones lineales y de naturaleza global. Las no afines son “deformables” o “no rígidas” y permiten deformar la imagen objetivo para adecuarla al modelo. Para la segmentación del hígado se superpone en cada corte el modelo aproximado del atlas y mediante métodos de registro deformable se ajusta para extraer la región definitiva del hígado. Se reconstruye el hígado tras el procesamiento corte a corte en 2D (Linguraru et al., 2009).

Platero presenta un trabajo donde se combinan varios métodos. El conocimiento de forma es una restricción global y local en modelos basados en borde y en región. El conjunto de nivel o *level set* (será explicado más adelante) se mueve en una banda estrecha, donde cada

píxel evoluciona según su nivel de gris, vecindario y sus propiedades geométricas que han sido establecidas por los modelos definidos. El alineamiento de forma se hace mediante transformación afín y procedimiento de normalización de imagen (Platero et al., 2011).

Este tipo de métodos basados en modelo a partir de atlas tienen varios inconvenientes. Por una parte, la obtención de los *landmarks* o puntos de interés necesarios para caracterizar los órganos y generar los modelos correspondientes no siempre es sencilla. En el caso del hígado, con formas muy redondeadas, es difícil identificar puntos claros y reconocibles en todas las muestras. Es por ello, que la obtención de los puntos de interés de manera automática se hace compleja y es necesaria la interacción con el usuario. Además, una vez generado el modelo, la segmentación corte a corte requiere un gran tiempo de cómputo ya que, la práctica habitual incluye evaluar el valor de gris (y el gradiente) de cada píxel de cada línea del perfil que se genera de manera perpendicular a la superficie en cada punto de interés, o evaluar características de textura u otros descriptores. En general, son métodos lentos y que requieren interacción por parte del usuario.

4.3.8. Cortes de grafos (graph cuts)

Aproximación teórica

El método de segmentación basado en cortes de grafos, o *graph cut* en inglés, pertenece al tipo de métodos de segmentación definidos como la resolución de un problema de minimización de energía en espacios discretos. Se usará el término en inglés en la descripción.

Este método fue propuesto en 2001 y desde entonces ha sido empleado en problemas de segmentación y otros problemas de procesamiento y análisis de imagen (Boykov and Jolly, 2001).

La idea principal de los *graph cuts* consiste en representar la imagen mediante un grafo con pesos. Cada nodo del grafo representa un píxel de la imagen, y cada borde o arista conecta un par de píxeles adyacentes. El peso del borde o arista representa el grado de similitud de cada par de píxeles en características de nivel de gris, textura o color que se hayan definido. La segmentación es un corte del grafo. Cada región representa lo que se denomina sub-grafo. El corte óptimo es lo que hace la similitud máxima en un sub-grafo y la similitud entre los distintos sub-grafos mínima. Se ha extendido su aplicación en imagen médica en los últimos años y puede alcanzar el mínimo global con algunas funciones de energía.

Los elementos que se manejan en el método son los siguientes:

$$X = (x_1, x_2, \dots, x_p, \dots, x_{|P|})$$

Donde X es el conjunto de píxeles de la imagen a segmentar con su valor de nivel de gris

$$P = (1, 2, \dots, p, \dots, |P|)$$

Siendo P el conjunto de índices de la imagen

Existe un conjunto N de pares no ordenados $\{p, q\}$ en un sistema con una vecindad 4 u 8, donde p y q implica que dos píxeles son vecinos.

$$L = (L_1, L_2, \dots, L_p, \dots, L_{|P|})$$

Donde L es un vector de valores que indican la asignación de cada píxel de la imagen, es decir, su etiqueta o clase a la que pertenecen. En el caso de una segmentación de un órgano determinado (por ejemplo, el hígado) sobre el fondo, estas etiquetas contendrían el valor de '0' para el fondo (*background*) y el valor de '1' para aquellos píxeles que finalmente hayan sido clasificados como hígado (región de interés o *foreground*).

En este método de segmentación la energía a minimizar se define como:

$$E(L) = \lambda R(L) + B(L) \quad (4.1)$$

Donde $R(L)$ es el término unario y R_p la penalización que se le da al píxel p por pertenecer a fondo o al órgano segmentado. Las penalizaciones se asignan en base a contexto de vecindario y según dicho píxel pertenezca a objeto o fondo.

$$R(L) = \sum_{p \in P} R_p(L_p), \begin{cases} R_p, & \text{"foreground"} \\ R_p, & \text{"background"} \end{cases} \quad (4.2)$$

Y $B(L)$ es el término de frontera, de vecindad, con la siguiente expresión:

$$B(L) = \sum_{\{p,q\} \in N} B_{\{p,q\}} \delta(L_p, L_q), \text{ con } \delta(L_p, L_q) \begin{cases} 1, L_p \neq L_q \\ 0, \text{ en otro caso} \end{cases} \quad (4.3)$$

Y el coeficiente $\delta \geq 0$ indica la importancia de $B(L)$ con respecto a $R(L)$.

El objetivo del método *graph cut* es identificar la segmentación que minimiza la energía $E(L)$ de todas las posibles segmentaciones L . La energía se minimiza sumando los pesos de los bordes o aristas del grafo. Normalmente se añaden algunas restricciones o condicionantes en cada problema concreto de segmentación.

Para calcular este mínimo global que proporcionará la segmentación óptima se representa en modo de grafo, siendo cada nodo del grafo un píxel. En el grafo, los nodos denominados "T" son interpretados como el fondo, y los denominados "S" son interpretados como el elemento segmentado. Las aristas que unen cada píxel a T y S tienen el valor de los términos unarios de $R(L)$ y las aristas entre píxeles vecinos tienen los términos de frontera $B(L)$.

En este método los píxeles p y q son las semillas, que pueden ser escogidas de diversas maneras, una semilla perteneciente a la clase fondo y la otra perteneciente a la clase del

objeto a segmentar. El resultado óptimo que implica que se ha localizado el mínimo global y por tanto, la mejor segmentación, es aquel en el que se obtiene el coste mínimo de corte en el grafo. Para realizar este corte que minimiza la energía existen varios algoritmos, como el de max-flow o min-cut, donde su eficiencia es comparada por el mismo promotor de los graph cuts en un trabajo (Boykov and Kolmogorov, 2004). Otras variaciones incluyen modelos de Markov, o el algoritmo GrabCut, (Rother et al., 2004), bastante utilizado recientemente, que propone un método de optimización iterativa que minimiza la necesidad de interacción por parte del usuario para la segmentación eficiente de fondo y de la región de interés. Este método combina tanto la información de bordes como de textura y presenta resultados robustos sobre las imágenes que muestra en el trabajo.

Boykov sigue apostando por estas técnicas de segmentación y en un nuevo trabajo se centra en la aplicación más directa de los *graph cuts*, que es la segmentación de objetos en las imágenes, y se mencionan las conexiones interesantes con otros métodos de segmentación, como son los contornos activos, *level set* o contornos geodésicos activos (Boykov and Funka-Lea, 2006). También se relaciona con el funcional de Mumford-Shah (Mumford and Shah, 1989) (se explican todas estas técnicas en el apartado 4.3.9 de modelos deformables) y derivados que combinan regularización de bordes y propiedades de las regiones. Desde que este método de *graph cuts* fue propuesto en 2001 ha habido muchas variaciones basadas en cálculo iterativo de los parámetros, aprendizaje, extensión multiescala, banda estrecha, etc. Los enfoques más usados para segmentación de imagen basada en grafos están recogidos en la referencia (Peng et al., 2013)

Existen algunos trabajos interesantes donde se propone el uso de *graph cuts* iterativos. Es decir, en vez de separar el fondo del objeto de interés de manera global, comenzar de manera local en sub-grafos del vecindario localizando los cortes óptimos para seguir expandiéndose por el resto del grafo (Peng et al., 2010). Este trabajo también propone que, en vez de que los nodos del grafo sean los píxeles de la imagen, los nodos sean unas regiones, agrupaciones de píxeles homogéneas (por la característica descriptora que sea) previamente generadas, lo que demuestra que el método es más robusto y eficiente que empleando píxeles en los nodos, como es la aproximación inicial. Parece dar buenos resultados en comparación con otros métodos más estándar como el GrabCut mencionado previamente.

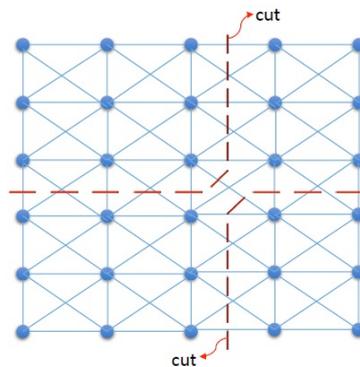


Figura 4.8. Ejemplo de grafo generado y cortes posibles

La principal desventaja que presenta este tipo de métodos es que es necesario conocer dos semillas en la imagen, una que pertenezca al objeto a segmentar y otra al fondo. Esto hace que en muchos trabajos se precise de la interacción por parte del usuario, aunque también se trabaja en la obtención de dichas semillas de manera automática, lo que añade complejidad al problema. Acorde a los resultados ofrecidos por los trabajos mencionados en este apartado, parece funcionar bien si el fondo es homogéneo con respecto al objeto a segmentar. La dificultad comienza cuando el fondo no siempre es uniforme, debido a alguna de las causas descritas en el capítulo 2.

Aplicación a segmentación hepática

Se han identificado en la literatura algunos trabajos que emplean *graph cut* para segmentación de hígado en imagen abdominal. En uno de ellos, se segmenta el hígado a través de técnicas de *graph cuts*, combinadas con otras, y se aplica a imágenes de CT y MRI (Massoptier and Casciaro, 2007). El algoritmo propuesto comprende 3 fases. Primero, sobre la imagen original se realiza un pre-procesado para eliminar el ruido, preservando los bordes. Para ello, se aplica un filtro no lineal (Comaniciu and Meer, 2002). A continuación, sobre esta imagen mejorada, y con el objetivo de poder proporcionar al método de *graph cut* un ejemplo de la región de interés y un ejemplo de fondo, se han establecido unos valores de umbral adaptativo basado en estadísticos de primer orden (media y varianza) entre los cuales se encuentra el hígado, de manera que se puede realizar esta inicialización burda frente al fondo. Esta región es sobre la que se aplica el método *graph cut* que realiza la segmentación final del hígado. Los resultados se evalúan con respecto a una segmentación de referencia realizada manualmente y mediante 4 métricas: coeficiente de similitud de Dice (DSC), ratio de falsos positivos (FPR), ratio de falsos negativos (FNR) y la medida de superficie simétrica, que penaliza fugas o segmentaciones deficientes. Los resultados parecen razonables.

Chen et al. proponen un método automático sobre imágenes de MRI con poco contraste, donde se emplea k-means para obtener las semillas tanto del objeto de interés (hígado) como del fondo. Estas semillas constituyen la información de entrada de un método de *graph cut*, que realiza la segmentación posterior (Chen et al., 2010). Como último ejemplo, se indica un método en el que se aplican también las técnicas de *graph cuts* con algoritmo de max-flox para segmentación hepática unido con modelo de apariencia y restricciones estrictas de ubicación anatómica. También se añade a la expresión de la energía una medida de conectividad espacial para robustecer la segmentación (Y. Chen et al., 2012). Los resultados de ambos dos trabajos son aceptables.

4.3.9. Modelos Deformables

Este apartado del estado del arte referido a los modelos deformables va a ser el más extenso de todos, dado que el método propuesto en la presente tesis para segmentación hepática se apoya en ellos y surge de la evolución y combinación de varios de estos elementos. Por

tanto, se van a describir en detalle los contornos activos, desde sus orígenes hasta su evolución como funcionales de energía resolubles por técnicas variacionales.

A modo de resumen, este conjunto de métodos tiene gran aceptación en la comunidad científica. Se basan en que un contorno inicial se encoge progresivamente hasta ajustarse a la forma del objeto a segmentar. El movimiento de dicho contorno es regido por una combinación de fuerzas externas e internas condicionadas por las características de la imagen. Son altamente dependientes de la inicialización del contorno (mejores resultados en un punto cercano a la solución), y de los parámetros que regulan la distribución uniforme de puntos y la flexibilidad de la curvatura para definir irregularidades más o menos acentuadas. Entre los modelos de este tipo más conocidos y con mayor repercusión se encuentran los Geodesic Active Contours (GAC) de Caselles (Caselles et al., 1997) y los Active Contours Without Edges de Chan y Vese (ACWE) (Chan et al., 2000), que han sido fuente de trabajos y variantes muy prometedoras.

4.3.9.1. Contornos activos o *snakes*

Aproximación teórica

La localización de contornos no rectangulares o lineales en las imágenes es habitual, pudiendo identificarse multitud de contornos curvilíneos que delimitan regiones. Esto ha originado el desarrollo de técnicas para localizarlos, englobadas en el concepto de los contornos activos (*active contours*).

El objetivo de los contornos activos es la segmentación de elementos que pueden ser difíciles de caracterizar con un patrón fijo predefinido dado que presentan gran variabilidad. También son válidos en el caso de que se pretenda seguir un objeto cuya forma varíe ligeramente en el tiempo. Es necesario, por tanto, el uso de técnicas capaces de evolucionar para adaptarse al objeto cambiante y de esta manera, continuar segmentándolo correctamente. Son formas flexibles. En definitiva, son conjuntos de puntos (contornos) que se mueven y actualizan en cada iteración (activos). El apelativo de *snakes* fue asignado por su creador, ya que los contornos activos en su evolución recuerdan el movimiento de una serpiente. Estos métodos persiguen una regularización de la segmentación ante el ruido.

Las *snakes* (Kass et al., 1988) son unas curvas de minimización de energía que evolucionan hacia características de la imagen específicas como pueden ser los bordes de los objetos. La curva evoluciona modificando las posiciones de sus puntos en cada iteración, y en el caso de las *snakes*, normalmente encogiéndose hasta pegarse a los bordes (si ésta es la característica delimitante, que suele serlo en la mayor parte de casos). En el momento en que dicha curva queda adherida a los bordes de una región, su energía es mínima, y óptimamente, muy próxima al valor 0. Las *snakes* suelen encoger, y por tanto es mejor inicializar la primera región fuera del contorno real a describir. La primera propuesta de *snake* se planteó como un conjunto discreto de puntos cuyas posiciones deben determinarse para que la energía que representan sea mínima. También podría expandirse, aunque los trabajos que lo abordaron obtuvieron peores resultados. De manera frecuente, se ubica el

contorno inicial fuera de la región y éste va encogiéndose hasta pegarse a los bordes del objeto a localizar. ¿Cuándo finaliza el proceso? En alguna de estas dos circunstancias: se han realizado todas las iteraciones previstas o se ha alcanzado el mínimo de energía.

La característica objetivo es un mínimo de un funcional de energía convenientemente formulado. Este funcional de energía contiene además de la información de bordes, información acerca de cómo el contorno puede ajustarse y curvarse. El movimiento de una *snake* es un compromiso entre sus propias características (cómo se curva y ajusta) y las propiedades de la imagen (como el valor del borde). En definitiva, el funcional de energía es función de la energía del contorno interno, su energía limitante y la energía de la imagen. Dichas funciones de energía son funciones del contorno (*snake*), cuyo conjunto de puntos (x, y) constituyentes irán evolucionando hacia el contorno mínimo o solución.

Se pueden considerar dos enfoques según se trate del espacio discreto (n° puntos establecido) o continuo (n° puntos infinito). Dado que la solución en el espacio continuo puede ser más difícil de comprender, es preferible comenzar con el algoritmo de Greedy en el espacio discreto (Williams and Shah, 1992). Es más sencilla la implementación en el espacio discreto. En él, el contorno se compone de un conjunto de puntos $p_1 \dots p_n$, donde $p_i = (x_i, y_i)$. El contorno evoluciona gestionado por su energía, que tiene 2 componentes:

- *Energía interna*: depende de la forma del contorno y de la relación entre sus puntos constituyentes. Se buscan contornos de curvas suaves y regulares.
- *Energía externa*: depende de la imagen, y persigue que el contorno se acerque a la forma buscada, por ejemplo, pegándose a los bordes del objeto.

Se puede expresar la energía total del contorno de la siguiente manera, siempre en el espacio discreto.

$$E = \sum_{i=1}^N (\alpha_i E_i^{elas} + \beta_i E_i^{curv} + \gamma_i E_i^{image}) \quad (4.4)$$

Cada uno de los términos tiene un parámetro de peso que regula su influencia en la energía final, en base a que se decida dar más relevancia a una u otra componente. Los dos primeros términos: elasticidad y curvatura son energías internas del propio contorno; la energía externa, depende de la imagen, por ejemplo, de los bordes presentes en la misma.

Conceptualmente, la elasticidad de la curva mide el espaciado de los puntos del contorno. Cuanto más regularmente estén espaciados, el valor de la energía es menor. Numéricamente,

$$E_i^{elas} = (\| p_i - p_{i-1} \| - \bar{d})^2 \quad (4.5)$$

siendo d la distancia entre puntos.

La energía de curvatura persigue que el contorno sea suave, con pocas protuberancias y zonas de curvatura pronunciada. Cuanto más suave sea el contorno, el valor de la energía es menor.

$$E_i^{curv} = \| p_{i+1} - 2p_i + p_{i-1} \|^2 \quad (4.6)$$

La energía de la imagen, en este caso de los bordes, mide cuánto se pega el contorno a los bordes del objeto. En los bordes, el gradiente es máximo y la energía, mínima.

$$E_i^{image} = -\|\nabla I\| \quad (4.7)$$

siendo I la imagen.

La ubicación definitiva de los puntos del contorno en cada iteración es aquella que minimiza el conjunto de las 3 energías. Parece lógico pensar que para poder comparar energías es necesario que éstas estén normalizadas. Se pueden normalizar restándoles el mínimo y dividiéndolas entre el rango de valor.

El algoritmo de búsqueda del contorno óptimo seguiría los siguientes pasos:

1. Se inicializa el contorno.
2. Se evalúa la energía total (suma de las 3 energías) de todos los puntos del contorno en el punto actual y en todos los puntos de una vecindad de paso establecido.
3. Se desplaza cada punto del contorno a la posición que hace que la energía total sea mínima.
4. Se repiten los pasos 2 y 3 hasta que la energía llega a un mínimo establecido (gradiente -1 es lo óptimo) o los puntos ya permanecen en sus posiciones de manera estable.

Los valores de los pesos de cada término, $\alpha_i, \beta_i, \gamma_i$ deben ser inicializados, pudiendo incrementar ligeramente el peso de los bordes por encima de la curvatura o el espaciado del contorno y condicionando, por tanto, la forma final de la región segmentada.

En el marco de esta tesis, se han realizado algunas pruebas con contornos activos, en el entorno discreto y se han verificado las limitaciones que se indican en la literatura. El resultado es fuertemente dependiente de la inicialización del contorno (lo que se conoce como parametrización), del número de puntos que constituyen este contorno inicial, de la inicialización de los pesos de cada término, del vecindario de búsqueda establecido, etc. Son bastante limitados en aquellas figuras con cambios bruscos en la curvatura o muchas protuberancias, siendo incapaces de adaptarse. Se muestra en la siguiente figura la secuencia de evolución de un contorno activo en el espacio discreto adaptándose a una forma sencilla.

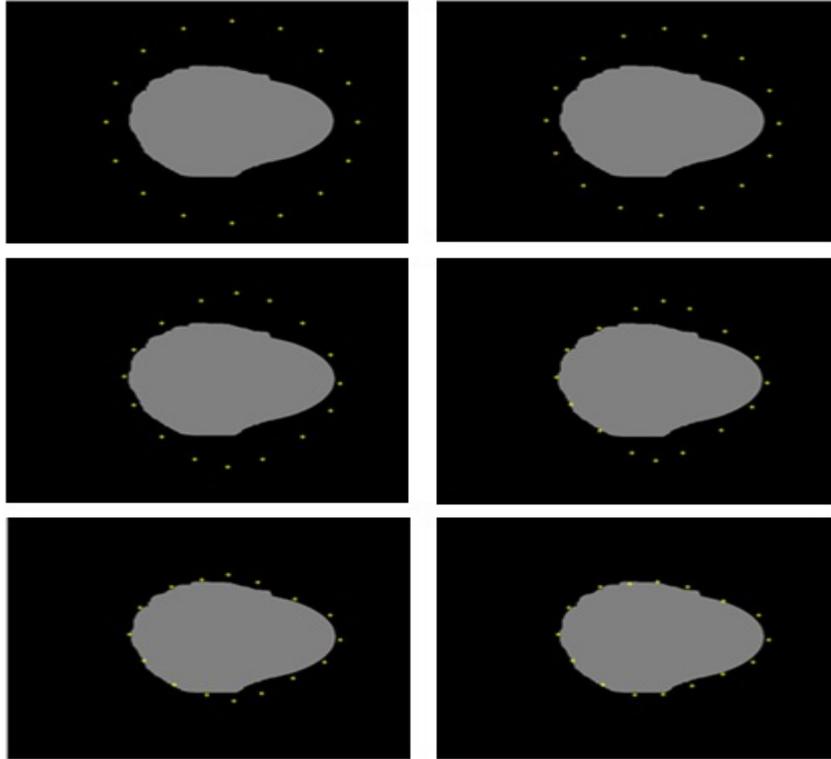


Figura 4.9. Secuencia de evolución de la snake hasta pegarse al borde del objeto

Una vez se ha comprendido el funcionamiento de un contorno activo en el espacio discreto, que es lo que inicialmente representaba el concepto de *snake*, se puede extender el método al espacio continuo.

De esta manera, los contornos activos ubicados en una imagen son expresados como procesos de minimización de energía. La característica objetivo es un mínimo de un funcional de energía. Este funcional de energía debe contener más que sólo información de bordes. Contiene información acerca de cómo el contorno puede ajustarse y curvarse, siendo su comportamiento similar al de una *snake*. En definitiva, el funcional de energía es función de la energía del contorno interno, de su energía limitante y de la energía de la imagen. El funcional de la energía es la integral de las funciones de la *snake* para el contorno completo $s \in [0,1]$ y se expresa de esta manera en la siguiente ecuación:

$$E_{snake} = \int E_{int}(v(s)) + E_{image}v(s) + E_{con}(v(s))ds \quad (4.8)$$

Esta es la energía que se debe minimizar donde,

E_{int} , es la energía interna, controla el comportamiento natural del contorno y por tanto, la colocación de los puntos.

E_{image} , es la energía de la imagen, atrae el contorno hacia los puntos seleccionados con información de bajo nivel sobre el cambio de región, pudiendo ser los bordes a través del cálculo del gradiente.

E_{con} , *constraint energy*, permite añadir información de alto nivel para controlar la evolución del contorno (opcional, puede aparecer o no).

La aproximación clásica inicial, propuesta por Kass, sólo considera la energía interna y la energía de la imagen, con lo que se proseguirá la descripción con este enfoque. En consecuencia, el contorno activo debe minimizar dicha ecuación. De hecho, el problema a resolver consiste en encontrar el conjunto de puntos $v(s)$ para el de la derivada de la energía es 0, esto es, $\frac{dE_{snake}}{dv} = 0$. El conjunto de puntos que resuelven esta ecuación pueden ser tanto un mínimo como un máximo, que se discrimina finalmente mediante la derivada de segundo orden. Para solucionar dicha ecuación, analicemos cuáles son los parámetros que controlan el comportamiento del contorno.

Los funcionales de energía son expresados:

$$E_{snake}(s) = E_{int}(v(s)) + E_{image}(v(s)) \quad (4.9)$$

Donde

$$E_{int} = \alpha(s) \left| \frac{dv(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2v(s)}{ds^2} \right|^2 \quad (4.10)$$

$$E_{image} = -\|\nabla I\| \quad (4.11)$$

$\alpha(s)$ – controla la contribución de la energía elástica debido al espaciado de los puntos.

si $\alpha(s)=0$ se elimina la influencia del espaciado, de manera que puede haber puntos desigualmente distribuidos, o incluso unos encima de otros.

$\beta(s)$ – controla la contribución de la energía de la imagen, y determina cómo es la curvatura del contorno final.

si $\beta(s) = 0$, puede haber esquinas con mucha curvatura, si aumentamos β , el contorno es más suavizado, y va a depender fuertemente de las características de la imagen, en concreto de los bordes.

Si se pasa del entorno discreto a un espacio continuo para resolver un problema de segmentación mediante un modelo de contorno activo, el problema matemático que debe resolver el investigador se conoce como técnicas variacionales aplicadas a segmentación de imagen.

Las técnicas variacionales son problemas matemáticos que consisten en buscar máximos y mínimos de funcionales continuos definidos sobre algún espacio funcional. Constituyen una generalización del cálculo elemental de máximos y mínimos de funciones reales de una variable. Las imágenes son un conjunto discreto de valores (1 canal) o vector de valores (RGB por ejemplo), pero en el espacio continuo se puede interpretar la imagen como una función suave, de manera que es admisible procesarla mediante técnicas variacionales, que además han resultado y resultan muy prometedoras y ventajosas para la segmentación de regiones. Estos funcionales de energía (típicamente en forma de integral) que se definen sobre las imágenes son tales que un mínimo de dicho funcional proporcionará una solución al problema de segmentación. Las ventajas son varias ya que permiten incorporar a los modelos muchas características geométricas como gradientes, normales, curvaturas, valores estadísticos. La resolución matemática al problema se puede reducir a la solución de Ecuaciones en Derivadas Parciales no lineales, que es un problema abordado ampliamente desde el punto de vista matemático y que se puede trasladar al espacio discreto para encontrar una solución.

Sin embargo, este método de segmentación por contornos activos presenta algunos inconvenientes. Entre ellos, que son dependientes de varios parámetros, dependen de la inicialización del contorno, no tienen la suficiente flexibilidad para ajustarse a formas muy complejas, y no son capaces de localizar varias regiones en la misma imagen, ni tampoco regiones interiores. En base a las limitaciones detectadas, se fueron abordando periódicamente otros enfoques.

Por ejemplo, se consideraron contornos abiertos en vez de cerrados (Waite and Welsh, 1990); se modificó el vecindario de la *snake* para poder converger y evitar posibles movimientos oscilatorios (Lai and Kin, 1994). También, con el objeto de permitir que la *snake* no sólo se contraiga sino que también se expanda, se diseñó para que evolucionara hacia fuerzas no atrayentes (Cohen, 1991) (Cohen and Cohen, 1993). Otros enfoques han sustituido el término de energía de la imagen que inicialmente se pega a los bordes, por términos que evalúan modelos estadísticos o textura (Ivins and Porril, 1995) y que pueden ser considerados precursores de los Active Shape Models (ASM) (Cootes et al., 1995) o Active Appearance Models (AAM) (Cootes et al., 1998), de los que se hablará más adelante.

Asimismo, para solucionar el problema de la parametrización adecuada, esto es, la elevada influencia del ajuste de los parámetros, se probó a controlar solamente un parámetro y expresar el resto de los términos en base a él. Se denominó a esto “regularización” (Lai and Kin, 1994).

Como se ha mencionado, un problema importante es la inicialización, que influye tremendamente en el resultado final. Para ello, se propuso embeber el contorno en evolución entre un contorno interno y otro externo (lo cual no deja de ser una primera aproximación a los level set) (Gunn and Nixon, 1997). Este primer enfoque fue resuelto con un algoritmo de programación dinámica (Amini et al., 1990), que también incluye forma

local y un solo término de parametrización, la “regularización” mencionada antes. El contorno interno se expande mientras el externo se contrae.

Otro problema importante de los contornos activos o *snakes* es que no es posible modificar la topología del contorno, es decir, la curva no puede dividirse en dos aunque rodee dos objetos que verifican las condiciones, no puede detectarlos simultáneamente y hay que actuar sobre la parametrización cada vez que se detecta una circunstancia excepcional.

Otra variación de los contornos activos, son las tijeras inteligentes o *scissors*. A veces, la evolución de los contornos activos es impredecible y puede requerir intervención del usuario. Una alternativa es que el sistema pueda optimizar el contorno en tiempo real al mismo tiempo que el usuario dibuja, son las *intelligent scissors* (Mortensen and Barrett, 1995). El funcionamiento es el siguiente: a medida que el usuario dibuja una línea, el sistema dibuja dicha línea lo más próxima posible a los bordes, interpretado como la zona con mayor contraste. Para ello, la imagen es inicialmente pre-procesada para ajustar los bordes, que es posible que sean los límites. Se emplean principalmente las magnitudes y orientaciones de gradiente en la identificación de los bordes.

Aplicación a segmentación hepática

En 2010 Ibrahim et al. recuperan técnicas de contornos activos relativamente sencillas sobre imágenes de MRI. El autor usa 8 conjuntos de datos para validación. Sobre estos datos se aplica un pre-procesamiento para eliminar el ruido, resaltar los bordes y aumentar el contraste. Selecciona una imagen (corte de la secuencia) inicial, donde el hígado aparece con forma elipsoidal o bastante circular, con bordes diferenciados y un contraste razonable con los órganos adyacentes. Con la transformada de Hough encuentra el centro del supuesto círculo. Lanza diámetros desde este punto cada 15° y obtiene pares de puntos, que constituyen el contorno inicial. Mediante la aplicación de la técnica de *intelligent scissors* traslada este contorno a la imagen del 2° corte (en la que en principio no hay mucha diferencia en forma ya que pueden estar separadas por 2-3 mm en el eje z de corte axial, dependiendo de la resolución), y vuelve a segmentar. Una vez se ha segmentado en una dirección de movimiento, por ejemplo en las imágenes ocasionadas por los cortes del eje z sobre el volumen, segmentan en otra dirección de movimiento y combinan los volúmenes segmentados, siendo el volumen intersección el que es considerado como el definitivo. El pre-procesamiento es muy importante para mejorar la segmentación en caso de que el contraste entre los elementos de la imagen sea bajo. El métodos presenta problemas de sub-segmentación en los resultados obtenidos (Ibrahim et al., 2010).

En los últimos años se han propuesto más métodos automáticos con contornos activos (Peng et al., 2011)(Peng et al., 2012) (Reska et al., 2012)(Luo et al., 2014) (Peng et al., 2014). Dichos métodos combinan modelos deformables con técnicas variacionales, para minimizar los funcionales de energía en busca de la solución global óptima. Esto se detalla en el apartado 4.4.

4.3.9.2. Isosurfaces y Level sets

Aproximación teórica

La primera aproximación a los métodos de conjunto de nivel (más conocidos por su término en inglés *level set*) fue propuesta en 1988 (Osher and Sethian, 1988) y fue trabajada posteriormente (Malladi et al., 1995) en el intento de solucionar las limitaciones identificadas en los contornos activos basados en las *snakes*, algunas de las cuales se han mencionado en el apartado anterior, como son la dependencia de varios parámetros regularizadores, de la inicialización, la escasa flexibilidad para adaptarse a determinadas topologías y la imposibilidad de fusionarse entre sí o dividirse en diferentes contornos acorde a los elementos que vayan encontrándose. Asimismo, es arriesgado cambiar la topología de la curva que evoluciona. De hecho, si la forma cambia de manera drástica, es necesario reparametrizar la curva.

Ante esta serie de limitaciones, surgen los *level set* como una representación alternativa, donde el cruce a cero (*zero-crossing*) de una función característica define la curva. Se trata de un método numérico para seguir la evolución de contornos y superficies. En vez de manipular directamente el contorno, dicho contorno está embebido como el *level set 0* (conjunto de valores zero) de una función de mayor dimensión, llamada *level set function*. La función *level set* evoluciona bajo el control de una ecuación diferencial. En cualquier momento, el contorno evolvente puede ser obtenido extrayendo el *zero level set*. La ventaja es que formas complejas pueden ser modeladas ya que cambios topológicos como fusión de regiones (*merging*) o división de regiones (*splitting*) son manejados implícitamente.

Los *level sets* evolucionan modificando su función envolvente implícita (*underlying embedding function*) en vez de la curva $f(s)$ de los contornos activos, como se muestra en la figura 4.10.

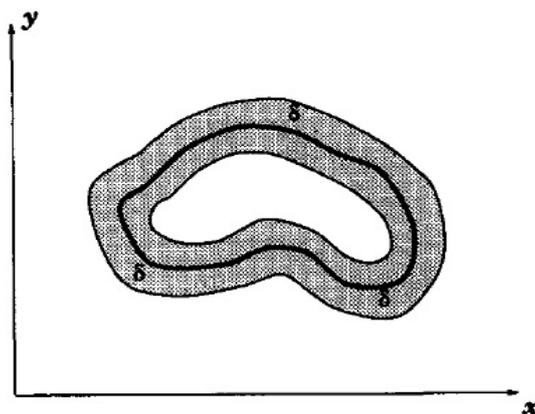


Figura 4.10. Banda estrecha de anchura δ alrededor del level set (Malladi, Sethian, 1995)

Para realizar la segmentación en vez de la propagación de la *snake*, la interfaz que se propaga es la curva C . Esta curva C puede evolucionar con una determinada velocidad acorde a esta ecuación diferencial:

$$\frac{d\vec{x}}{dt} = \vec{V}(\vec{x}), t \in [0, T], \vec{x} \in C \quad (4.12)$$

Esto indicará la ubicación del conjunto de puntos de la curva C en un instante de tiempo t . Esta es la formulación paramétrica del problema, considerando que los puntos del frente se están siguiendo uno a uno.

Otra alternativa es definir una función $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$, que describe el frente que evoluciona a través de su conjunto de nivel 0, esto es, $\phi(\vec{x}) = 0$. Esta es la formulación denominada euleriana de la evolución de la superficie, ya que no sigue el movimiento de cada punto uno a uno, sino que se encuentra implícito en la función ϕ . Si se desea que ϕ capture la posición del frente en cada instante de tiempo t , y en el caso de dos dimensiones (generalizable a n):

$$\phi(\vec{x}(t), t) = \phi(x(t), y(t), t) = 0 \quad (4.13)$$

Y derivando con respecto a la variable tiempo t :

$$\phi_t(x(t), y(t), t) + \frac{\partial \phi}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial \phi}{\partial y} \frac{\partial y}{\partial t} = 0 \quad (4.14)$$

Lo que se convierte en

$$\phi_t(x(t), y(t), t) + \nabla \phi \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t} \right) = 0 \quad (4.15)$$

Y a su vez en la que constituye la ecuación general de nivel zero o level set.

$$\phi_t + \nabla \phi \cdot \vec{V} = 0 \quad (4.16)$$

Los puntos de la imagen donde esta ecuación es válida caen en la curva C , es decir, en el conjunto de nivel cero de ϕ .

Aunque la técnica de *level set* permite que el contorno cambie su topología de manera radical y se adapte a cualquier tipo de curva y curvatura, puede quedarse estancado en algún mínimo local y por tanto, la función no convergerá en el mínimo global, proporcionando así la solución óptima, que en este caso representa la segmentación deseada.

Los elementos de la imagen que gobiernan la ecuación diferencial son el valor de intensidad media, así como los gradientes y bordes.

Las ecuaciones que representan los *level sets* se resuelven apoyándose en ecuaciones en derivadas parciales (EDP o PDE, siglas en inglés) para modelar lo que se denominan isosurfaces deformables. En un volumen pueden existir muchas isosurfaces anidadas. La

normal de una isosurface está dada por el vector gradiente normalizado en un punto dado. La curvatura de la isosurface puede ser calculada a partir de la envolvente de primer y segundo orden. La información de la forma de toda la superficie está dada por la matriz de normales en cada punto. Se puede calcular la curvatura media entre dos curvaturas principales. A la curvatura total se le llama también *deviation from flatness*, y es la raíz cuadrada de la suma de los cuadrados de las dos principales curvaturas. Una superficie puede expandirse de diversas maneras dependiendo de la forma, la curvatura de la superficie, y de fuerzas externas que condicionan esta evolución.

Las técnicas de *level sets* tienen un gran número de ventajas teóricas y prácticas sobre otros modelos de superficie convencionales, especialmente en el contexto de deformación y segmentación. Son flexibles, pueden representar superficies complicadas, identificar y respetar los agujeros, fusionar regiones en una sola o dividir una en múltiples objetos. Pueden incorporar muchos grados de libertad. Las formas están limitadas sólo por la resolución de las muestras. Las soluciones por EDP se resuelven mediante aproximación discreta (derivada como diferencia de píxeles contiguos, concepto habitual de gradiente en imagen).

Las técnicas de *level set* tuvieron una gran repercusión y se ha trabajado en variaciones que fueran limando sus deficiencias. La principal limitación es el elevado tiempo de cómputo ya que las direcciones de análisis de la energía de las isosurfaces o *level set* son muchas simultáneas, asimilando las posibilidades a las capas de una cebolla. Algunas de estas variaciones fueron las "*up-wind schemes*", donde la variación de la isosurface sólo se realiza en un sentido de movimiento, acotando las direcciones a investigar. Otras son los métodos "*narrow-bands*" o de banda estrecha, que surgen de la reflexión de que si sólo se está interesado en un *level set*, una única isosurface que segmente un determinado elemento, se reduzca el número de *level set* simultáneos. Así se reduce el tiempo de cómputo de manera notable. Se propone, por tanto, seguir la evolución del *level set* sólo en un vecindario de una isosurface ϕ . El resto de la imagen no contenida en esa banda se considera que no forma parte, de momento, del *level set* y por ello no merece la pena ser procesado. Se denomina un "*tube*" a la banda estrecha en la que el *level set* va evolucionando. Cuando el *level set* toca los límites de la banda estrecha, ésta debe recalcularse. La desventaja de este método es el coste también ya que la banda debe ser de cierta anchura, y el proceso de cálculo cada vez que el *level set* toca la banda es costoso.

Las técnicas de *level set*, en cualquier caso, necesitan grandes recursos, con lo que, para optimizar el tiempo de cómputo, se propuso actualizar cada vez en una pequeña intermediación de la frontera actual del zero-crossing, que es lo que se denomina métodos *Fast Marching* (Sethian, 1999a)(Sethian, 1999b). Estos métodos parten de la base que el *level set* se desplaza sólo en una dirección, cada píxel sólo se visita una vez, lo que simplifica la ecuación cuya solución óptima se puede encontrar más ágilmente, y se minimiza el tiempo de procesamiento necesario. Este conjunto de métodos ha sido ampliamente trabajado por la comunidad científica.

Aplicación a segmentación hepática

Existen varios trabajos que proponen las técnicas de *level set* aplicadas a segmentación de hígado en imagen abdominal. Chen et al. propusieron un método en el que, para mejorar las limitaciones de los *level set* en cuanto a que son dependientes de la inicialización y de que convergen mejor si son inicializados en las inmediaciones de la región a segmentar, se realizan 3 segmentaciones a partir de 3 inicializaciones diferentes del *level set*, y posteriormente se fusionan los resultados de la segmentación individual, logrando una segmentación mejor. Lo valida sobre imágenes de CT y MRI. Se trata de un método semiautomático, ya que precisa de la intervención del usuario para la inicialización de los contornos del hígado (Chen et al., 2009).

Un método de segmentación de hígado basado en *level set* y con mínima interacción con el usuario se propone sobre imágenes de CT en la referencia (Lee et al., 2007). Para inicializar el algoritmo el usuario especifica dos puntos semilla. Uno en la parte superior del hígado y uno en cada una de las puntas de los lóbulos del mismo. La imagen de velocidad de la evolución del *level set* se obtiene mediante el suavizado de la imagen de entrada con un filtro específico. A partir de cada una de estas semillas la región crece hasta topar con los gradientes elevados del borde del hígado. Estos resultados iniciales son refinados en una estrecha banda mediante Contornos Geodésicos Activos en el entorno de la inicialización (este método se describe en 4.3.9.6). Los resultados de cada imagen 2D se trasladan al siguiente corte o *slice* y son considerados la inicialización de un nuevo contorno geométrico activo, que hay que minimizar en el nuevo corte. Se repite el proceso hasta que las segmentaciones realizadas a partir de las semillas marcadas desde el corte de arriba y desde el corte de abajo se encuentran en la parte central del hígado. Se realiza un procesamiento posterior de suavizado y coherencia de resultados.

En otro trabajo identificado, a partir de una clasificación inicial basada en Maximum a Posteriori (MAP) sobre nivel de intensidad, se localiza la región con mayor probabilidad de ser hígado, y esta inicialización se refina posteriormente mediante técnicas de *level set*. Es objeto de este paper la localización de más estructuras corporales y emplea para ello otras técnicas además de las mencionadas, como clasificadores AdaBoost o procesamiento morfológico que se escapan del alcance de este capítulo (Furukawa et al., 2007).

La referencia (Wimmer et al., 2009) presenta un método de segmentación que aúna modelo de borde, modelo de región y modelo de forma. Representa la función conjunta con las tres componentes y lo minimiza mediante un *level set*. De la minimización resultante obtiene la región segmentada. Presenta un algoritmo que no es sólo para segmentar el hígado sino que puede aplicarse a otras situaciones. También intenta eliminar (minimizar) el número de parámetros a emplear. Como modelo de forma emplea estimación de densidad de Parzen (Parzen, 1962) dado que no supone que los modelos tienen una distribución normal. Esto no tiene por qué ser cierto si existe bastante variabilidad en los conjuntos de entrenamiento.

Las técnicas de *level set* son fuertemente dependientes de la inicialización y, para obtener buenos resultados, deben estar inicializadas cerca de la solución. Por ello es frecuente

realizar una segmentación burda por algún método y el ajuste fino a través de *level set*, cuando ya se está en las inmediaciones de la solución, a través de los *Fast Marching methods*, por ejemplo, que son las aproximaciones más rápidas, como ya se ha indicado. En otro de los trabajos localizados se propone realizar la segmentación burda a partir de clasificación *fuzzy*. El resultado de esta segmentación es la inicialización del *level set*, y algunos de sus parámetros reguladores se extraen también del proceso de clasificación *fuzzy*. Este método resulta eficiente si hay un contraste relativamente marcado entre el fondo y el objeto a segmentar, ya que las técnicas de *level set*, en general, presentan muchas fugas (sobre-segmentaciones) si no existe un contraste claro entre dos órganos adyacentes (Li et al., 2011).

Existen aproximaciones que mejoran el algoritmo de segmentación por *level set* incluyendo conocimiento previo a través de modelos basados en borde o en forma (Platero et al., 2011) y de esta manera evitar las fugas. Otra limitación de los *level set* es que hay un conjunto de parámetros que definir que regulan el comportamiento y evolución de las superficies, de manera que la incorrecta elección conlleva malos resultados. De hecho, dependiendo del corte de la secuencia en el que se aplique, es posible que los parámetros óptimos de la minimización sean diferentes, e incluso cambiantes dentro de cada corte en función de la curvatura, de la estrategia de crecimiento u otros condicionantes. Las técnicas de *level set* siguen siendo objeto de estudio y son muy empleadas.

4.3.9.3. Modelos Activos de Forma o Active Shape Models (ASM)

Aproximación teórica

La segmentación basada en modelos, en general, y en modelos de forma, en particular, surge para intentar solucionar varias limitaciones identificadas tras la aplicación de otras técnicas. La localización de elementos en imágenes con mucho ruido o donde éstos estén parcialmente ocultos es un problema complejo. Es necesario extrapolar la información oculta a partir de la disponible en la imagen. Esta circunstancia puede darse si, por ejemplo, se conoce a priori la forma global del objeto a segmentar, y de esta manera, el elemento deseado se puede identificar aunque no todos los puntos de su contorno se hayan posicionado correctamente. Esto puede permitir una buena segmentación incluso en condiciones difíciles. Así surgen los Modelos Activos de Forma o *Active Shape Models*.

Se propuso este nuevo enfoque como evolución de los contornos activos, en los que la flexibilidad del contorno está gestionada por propiedades locales de la imagen, como son la curvatura o los bordes. Se han descrito previamente éstos como métodos de segmentación flexible, que pueden evolucionar en el tiempo y cuya forma puede adaptarse para poder localizar las estructuras buscadas. Los contornos activos, en sus múltiples variaciones, actualizan las posiciones de sus puntos (en discreto) o de la curva (en continuo) en base a aquellos puntos en los que la energía total es menor, por ejemplo, por estar cerca de un borde o en el mismo borde. Sin embargo, no se puede saber a priori si el borde al que se están pegando es el adecuado o puede ser el de algún elemento contiguo de gradiente similar. En definitiva, se basa solamente en características locales de la imagen, y el

contorno, más allá de las propias restricciones de los otros términos de curvatura y espaciado, no está condicionado por nada.

Parece buena idea, por tanto, imponer una restricción a nivel global si se persigue segmentar un determinado elemento con unas características concretas de forma. Esta restricción global de forma acota el movimiento del contorno activo en un espacio de forma. Este fue el enfoque propuesto por Cootes en lo que denominó contornos activos de forma (ASM) (Cootes et al., 1995) y que dio origen a múltiples variaciones a posteriori, no sólo considerando la forma como restricción global, sino incluyendo también la textura u otros. Son los modelos flexibles de forma o modelos activos de forma. Siguen basándose en los principios de movimiento de los contornos activos pero, además, dichos movimientos están limitados por el modelo de forma generado previamente. Existe una variación de los ASM también propuesta por Cootes que añade la información de textura como condicionante del movimiento. Son los Modelos Activos de Apariencia o *Active Appearance Models* (AAM), (Cootes et al., 1998) y se describen posteriormente.

(Cootes et al., 1995) propuso generar el modelo de forma de un objeto, de manera que se contemplen todas las posibles variaciones en la forma de dicho objeto. Cootes considera que todos los objetos del mismo tipo, un modelo, comparten una región única de un sistema de referencia diferente al sistema de referencia de la imagen. En dicho nuevo espacio de referencia, cada elemento es un punto, y los ejes de dicho espacio, cualesquiera que sean sus dimensiones, maximizan las diferencias de esta clase con otra, de manera que un nuevo elemento perteneciente a ese modelo cae en el espacio acotado para ello.

(Cootes et al., 1995) propuso que la forma final de los contornos activos para poder ser considerados como solución al problema deben presentar una forma determinada y con una variación de la misma en cierto rango. La variabilidad de la forma se recoge a partir del procesado automático de un conjunto de ejemplos del objeto a segmentar, en los cuales se ha caracterizado la forma a partir de unos puntos de referencia identificados en todas las imágenes de entrenamiento. Las estadísticas sobre el ámbito de posición de dichos puntos describen las posibles variaciones de forma. Se suele usar Análisis de Componentes Principales para reducir la dimensionalidad (Jolliffe, 2005).

Los autovalores de más peso describen bastante bien la imagen de manera que se emplean dichos valores y los márgenes especificados para desechar o aceptar regiones susceptibles de corresponder con la forma del objeto a segmentar. Los márgenes en los límites de la forma admitida para que dicha región sea considerada como candidata a la región objeto de segmentación suelen estar marcados normalmente por $2\lambda_i$ o $3\lambda_i$, siendo λ_i los autovalores principales resultantes de calcularlos a partir del modelo de forma definido por la serie de ejemplos proporcionados. La mayor o menor flexibilidad en la forma viene determinada por $2\lambda_i$ o $3\lambda_i$, siendo 3 el factor en el que se exige más parecido a la forma media del conjunto del modelo (con los 3 primeros se describe en torno a 95-97% de los casos).

Los pasos para construir un ASM son los siguientes:

1. Modelo de Distribución de Punto (*Point Distribution Model*).

Lo primero es generar el modelo de la forma del objeto de estudio. Para ello es necesario hacer acopio de una serie de ejemplos de dicho elemento. El Modelo de Distribución de Punto (PDM) representa cómo están distribuidos los puntos de interés del objeto a modelar. La definición y obtención de estos puntos de interés, o *landmarks*, es clave. Es necesario seleccionar una serie de puntos, que aparezcan inamovibles en todos los ejemplos de la base datos. Estos puntos relevantes suelen ser las esquinas u otros puntos con gran cambio de curvatura, fáciles de identificar inequívocamente. Así mismo, para dar robustez al modelo, es recomendable la obtención de unos puntos secundarios, que pueden ser puntos adicionales equidistantemente espaciados entre 2 puntos principales. De esta manera se dispondrá de un número suficiente de puntos de interés que pueda representar la forma de manera fiable.

Una vez se dispone de todos los puntos de interés, es necesario posicionarlos en el mismo sistema de referencia, lo que se denomina alinear el conjunto de entrenamiento en dicho espacio de referencia. Es decir, cada conjunto de puntos correspondiente a una representación de la forma tiene sus coordenadas expresadas en el sistema de referencia de la imagen, ya que cada ejemplo puede tener dimensiones variables, con una escala y orientación cualesquiera. Para poder generar un modelo de forma es necesario que todos los puntos estén ubicados en el mismo sistema de referencia y de esta manera poder comparar las variaciones en los objetos. Para ello, se elige un origen de la nueva base, y se transforma cada conjunto de puntos a ella, eliminando la rotación y unificando la escala. Finalmente se dispone de un conjunto de puntos que representan el mismo objeto en el sistema de referencia único establecido, en la misma escala. A partir de la visualización de este conjunto de casos de entrenamiento, se puede apreciar cuánta dispersión existe entre sus puntos de interés, lo cual nos da una idea de la calidad de la base de datos empleada. Si la variabilidad es grande, es posible que el conjunto de *landmarks* no sea el adecuado para generar un modelo, o que la forma es realmente tan variable que no es factible modelarla, o que no se han escogido bien los puntos de interés principales o existe demasiada imprecisión en su cálculo.

Sobre este conjunto de puntos de interés de todos los ejemplos de formas, se puede calcular la forma media, el contorno medio y también la matriz de covarianza, la que representa la diferencia entre cada uno de los ejemplos de la base de datos con respecto al modelo. Es importante que la dispersión no sea elevada para garantizar que la forma puede ser descrita de una manera bastante compacta y con precisión. Este es el Modelo de Distribución de Punto, y representa el dominio de forma permitido (*Allowable Shape Domain* o ASD), es decir, el dominio en el que una nueva forma debe caer si representa a ese objeto.

2. Nuevo sistema de referencia para el modelo de forma

Cada ejemplo del conjunto de entrenamiento puede ser imaginado como un punto en el nuevo espacio de representación del modelo. Se asume que la nube de puntos es aproximadamente elipsoidal, y por tanto se pretende obtener los ejes principales y el centro de este nuevo espacio elipsoidal que representa dicha forma.

Dado un conjunto de N formas alineadas, se puede calcular la forma media, \bar{x} , lo que constituye el centro de dicho dominio de forma permitido elipsoidal, de esta manera:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.17)$$

Los ejes principales de este espacio elipsoidal pueden ser obtenidos mediante la aplicación de un Análisis de Componentes Principales (PCA) sobre la matriz de covarianza. Cada eje proporciona un modo de variación, en cuanto a que representa cómo se mueve el conjunto de puntos para representar dicha forma si el valor de dicho modo varía. De la aplicación de este PCA se obtiene una lista de autovalores o *eigenvalues*, que son los ejes principales, y sus respectivos autovectores o *eigenvectors*, que son los vectores que representan los nuevos ejes coordenados del nuevo sistema de referencia, expresados como relación con el sistema de referencia inicial, en el que están alineadas todas las formas. De esta manera, se puede establecer la correspondencia y el paso entre un espacio y otro de referencia, a través de la siguiente expresión:

$$x = \bar{x} + Pb \quad (4.18)$$

Con $P = (p_1, p_2, \dots, p_t)$ la matriz de los t primeros autovectores y $b = (b_1, b_2, \dots, b_t)$ un vector de pesos. Modificando este vector de pesos en límites aceptables, se pueden obtener nuevas formas, siempre dentro del espacio modelo. Límites típicos pueden ser

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \quad (4.19)$$

3. Búsqueda en la imagen

Una vez definido el modelo de forma a partir del conjunto de entrenamiento disponible, se pretende que el contorno activo realice la búsqueda de dicho elemento en una imagen cualquiera. No se trata de buscar elementos rígidos, que presenten siempre la misma forma, lo que se conoce de manera clásica como patrón, con poca flexibilidad con respecto al ejemplo, sino que se debe encontrar el conjunto de parámetros $\{b_1, b_2, \dots, b_k\}$ tales que el contorno siga representando un elemento de dicha forma, en cualquiera escala y rotación, pero absorba cierta variabilidad que pueda haber, acorde a los límites establecidos.

Se realiza la siguiente secuencia:

1. Se inicializa el contorno. Para ello, se puede usar el contorno medio calculado antes, para tener unos puntos iniciales a partir de los cuales buscar la forma. Es importante que esta inicialización no esté muy alejada de la solución final, dado que los puntos de dicho contorno no pueden “contorsionarse” demasiado buscando las nuevas mejores ubicaciones ya que entre sí están ligados por las condiciones globales del modelo de forma. Esta constituye una de sus limitaciones.
2. Se calcula un movimiento posible para cada punto del contorno. En la generación del modelo, en cada muestra de ejemplo, cada punto tiene su perfil de línea o *line profile*, es decir, la información de “textura” asociada a su posición. Esta información de textura es

sencilla ya que consiste en el valor de gris y derivada de la recta que pasando por el punto pertenece a la normal (en ambos sentidos). Esta información permite estimar si una posición u otra se asemeja más al valor de perfil de línea esperado, y en consecuencia, se propone un desplazamiento para dicho punto.

3. Se evalúan las propuestas de desplazamiento de todos los puntos del contorno, en base a criterios de energía, como corresponde a los contornos activos. Sin embargo, si el contorno con los nuevos valores de los puntos no verifica las condiciones de forma, todos los valores de los autovalores no están en límites, se establece los máximos valores de dichos autovalores que exceden, de manera que se calcula el máximo desplazamiento posible de los puntos.
4. Se repiten los pasos 2 y 3 hasta converger según el criterio establecido o los puntos apenas se desplazan ya de sus posiciones.

Hay que tener en cuenta que se está trabajando con 3 espacios de referencia: sistema de referencia de la imagen de entrada, sistema de referencia alineado del modelo, sistema de referencia de los ejes principales (o modos). El paso 2 se realiza sobre la imagen de entrada, el paso 3 se realiza sobre el sistema de referencia de los ejes principales, que indican una posición en el sistema de referencia alineado y que debe ser trasladado al sistema de la imagen de entrada. Se muestra en la siguiente figura un ejemplo de un ASM que detecta el contorno de unos pantalones.



Figura 4.11. Pantalones localizados mediante ASM. En azul el contorno inicial y en rojo la ubicación final de los puntos

Aplicación a segmentación hepática

Los ASM han sido ampliamente utilizados, siendo apreciados en ámbitos donde la poca definición de bordes, la oclusión o la variabilidad hacía imposible emplear métodos sin restricciones globales de forma. Han constituido también la base para multitud de desarrollos posteriores, como son los *Active Appearance Models*, que añaden información de textura en la generación del modelo.

Ambas dos aproximaciones han sido abordadas en diferentes ocasiones para el procesamiento de la imagen médica. En (Cootes et al., 1994) se revisa el uso de los ASM para localización de estructuras médicas y en (Cootes and Taylor, 2001) se recoge la trayectoria del trabajo realizado sobre ASM y AAM y su prometedora aplicación a imagen médica.

Sin embargo, no siempre pueden ser empleados ya que existen limitaciones. El principal problema surge si el elemento a segmentar es difícil de modelar, esto es, tiene una variabilidad de forma tan grande, por lo que sea, que no es posible definir un modelo que lo describa. Otra limitación es que, a la hora de que el modelo activo de forma se desplace hacia su convergencia, como está sometido a la restricción global de forma, hay que inicializarlo en puntos próximos a la solución. En caso contrario, es posible que no logre evolucionar adecuadamente y por tanto, no localice la solución final. Otro problema, no menos importante, surge del hecho de que, a pesar de ser una forma en principio modelable, la obtención de los puntos de referencia es muy difícil.

En el caso de aplicación objeto de esta tesis, si bien inicialmente se valoró como línea de trabajo, fue desechado. El hígado es un órgano con una forma definida y reconocible (aunque existe variabilidad entre individuos) pero la identificación de los puntos de interés es difícil. Sus formas redondeadas hacen que la identificación de puntos de referencia sea difícil y probablemente imprecisa sin interacción por parte del usuario. La automatización en la extracción de los mismos no va a ser robusta, sino que los estudios realizados con este enfoque se refieren a métodos de segmentación semiautomáticos, con gran tiempo de dedicación del usuario a la identificación manual de dichos puntos. Algunas de las experiencias identificadas son las referencias (Kainmüller et al., 2007)(Heimann et al., 2007)(Heimann and Meinzer, 2009)

Los métodos semiautomáticos o de interacción con el usuario están fuera del objetivo de esta tesis. Se deben buscar otros enfoques para solucionar el problema.

4.3.9.4. Modelos Activos de Apariencia o Active Appearance Models (AAM)

Aproximación teórica

Los ASM combinados con Análisis de Componentes Principales (PCA) en distribución de niveles de gris dan lugar a los Modelos Activos de Apariencia o *Active Appearance Models* (AAM) propuestos también por Cootes, como se puede encontrar en la referencia (Cootes et al., 1998). Constituyen un nuevo enfoque con respecto a ASM donde incluye la información de textura además de la de forma. Se usa un vecindario de cada punto para el análisis de textura. La metodología de los AAMs es similar a la de los ASMs en cuanto a usar unas imágenes de entrenamiento para identificar los autovalores que caracterizan la textura y la forma en un espacio transformado y que van a servir para identificar regiones que cumplan estos requisitos en el proceso de segmentación.

Aplicación a segmentación hepática

En la referencia (Cootes and Taylor, 2001) se recoge la trayectoria del trabajo realizado sobre AAM y su prometedora aplicación a imagen médica.

Un trabajo que realiza segmentación hepática mediante AAM se encuentra en la referencia (Beichel et al., 2001). En este trabajo se combinan los ASM y los AAM y lo extiende a tres dimensiones para segmentación de hígado sobre imágenes de CT. Para reflejar la bondad de los resultados sólo muestra imágenes, y no un análisis de resultados con respecto a métricas.

En el caso de segmentación hepática, Chen et al. proponen una combinación de *graph cuts* y de AAM para segmentación automática de hígado en 3D sobre imágenes de CT (X. Chen et al., 2012). Para la detección de los puntos de interés del hígado, que ha fijado en 35, se ayuda de la técnica de Live Wire que ajusta los puntos mal identificados a aquellos ubicados en un borde, aunque para ello necesita de intervención manual, con un enfoque similar a las "intelligent scissors" (Mortensen and Barrett, 1995). El algoritmo propuesto lo aplica sobre imágenes de CT y lo compara con los trabajos previos que se encuentran en las referencias (Kainmüller et al., 2007) y (Heimann et al., 2007) que fueron los que consiguieron mejores resultados en el MICCAI Challenge 2007, para lo cual usa las métricas allí propuestas. Los valores que obtiene para las métricas con este nuevo método son similares a los obtenidos por los otros dos, pero el tiempo de cómputo necesario para la ejecución ha sido reducido a la mitad.

Al igual que para el caso de los ASM aplicados a segmentación hepática, la gran dificultad estriba en la obtención automática de los puntos de interés, debido a las formas redondeadas y variables del hígado. De hecho, precisa de intervención del usuario para corregir las ubicaciones defectuosas, como se ha indicado. Es por ello, que no son viables para una segmentación automática.

4.3.9.5. Contornos Activos Sin Bordos o *Active Contours Without Edges* (ACWE)

Siguiendo la prometedora línea de trabajo de los contornos activos, *level set* y sus variaciones, aparecen en los años siguientes dos métodos que han alcanzado gran aceptación en la comunidad científica. Estos son los *Active Contours Without Edges* (ACWE) (Chan and Vese, 2001) y los *Geodesic Active Contours* (GAC) (Caselles et al., 1997), que se describen en éste y en el siguiente apartado, y son la base del modelo de contorno activo que se emplea en esta tesis, tal y como se explica posteriormente en el capítulo 5. Se van a referir sus siglas y nombre en inglés, por los que son conocidos.

Aproximación teórica

Los *Active Contours Without Edges* son un modelo de contorno activo propuesto por Chan y Vese en la referencia (Chan and Vese, 2001). Se basa en que el contorno activo no busca

gradientes o bordes en la imagen para poder minimizar el valor de su funcional, sino que se parte del supuesto de que existen dos regiones uniformes diferentes susceptibles de ser identificadas.

El trabajo de Chan y Vese se apoya en uno de los primeros intentos, junto con las *snakes* de Kass, de segmentar objetos en la imagen a través de la minimización de energía. El primer trabajo relacionado en el que se basan los ACWE es el del denominado funcional de Mumford-Shah (Mumford and Shah, 1989). Dicho funcional perseguía un doble objetivo: por una parte, eliminar el ruido y por otra, segmentar los objetos presentes en una imagen. Esto se consigue tras la minimización de este funcional:

$$F_{MS}(u, K) = \int_{\Omega} (u - I)^2 dx + \lambda \int_{\Omega - K} \|\nabla u\|^2 dx + v|K| \quad (4.20)$$

La función u es una suave aproximación de la función *piecewise* de la imagen I . Con el término “función *piecewise*” se refiere a una función híbrida, es decir, aquella que está formada por un conjunto de subfunciones, cada una de ellas definida en un intervalo del dominio de la función. Se divide la integral en dos partes ya que no existe ∇u en los saltos que define K . En el resto de los puntos de la imagen la función es suave. Los parámetros λ y v son constantes reales positivas.

Para aplicar este funcional de energía a la segmentación de imagen se consideran contornos cerrados, y de la minimización de la función (4.20) se obtiene un conjunto de elementos que consiste en una partición de la imagen I en regiones disjuntas Ω_i , de manera que $\cup_i \Omega_i = \Omega$, donde cada una de las particiones ha sido aproximada por una función suave $u_i: \Omega_i \rightarrow \mathbb{R}$.

La resolución de este problema no es sencilla ya que se combinan integrales y un conjunto de saltos definidos por K . Es necesario reformular matemáticamente el problema para encontrar un conjunto de minimizantes, que den una solución, que no tiene por qué ser única. A partir de aquí se empezó a trabajar en la solución de esta ecuación y un caso concreto de la misma son los ACWE, formulados en 2001.

En los ACWE, se define que el punto en que las regiones no son uniformes es el límite de las mismas y por tanto queda delimitado por el contorno de la región. Es decir, la función u sólo puede presentar 2 valores, c_1 y c_2 , siendo c_1 el valor medio de intensidad de la región K , y c_2 el valor medio de intensidad de la otra región, la que está fuera de K .

Es decir,

$$u(x) = \begin{cases} c_1, & x \in dentro(K) \\ c_2, & x \in fuera(K) \end{cases} \quad (4.21)$$

Entonces sólo existe un borde, y coincide con la snake de los modelos anteriores. El problema de segmentación de ACWE se convierte en

$$F_{CV}(K, u) = \lambda_1 \int_{inside(K)} (I - c_1)^2 dx + \lambda_2 \int_{outside(K)} (I - c_2)^2 dx + \rho \text{length}(K) + v \text{area}(inside(K)) \quad (4.22)$$

Se incorpora el área y la longitud de la curva como parámetros regularizadores, a menor área y menor longitud, menor energía que representa la función. Habitualmente $\lambda_1 = \lambda_2 = \rho = 1$ y $v = 0$.

Donde $c_1(K)$ y $c_2(K)$ son los valores medios de la intensidad dentro de la región K y fuera de ella, respectivamente, siendo estos valores cambiantes a lo largo de la evolución del contorno. El valor de la energía de ACWE no depende de la función detectora de borde como es el caso de los contornos geodésicos activos, que se explica en el apartado siguiente, sino en la diferencia de los valores medios de las 2 regiones en las que se asume que se divide la imagen.

Aplicación a segmentación hepática

Existen algunos trabajos en la literatura que proponen segmentación de hígado mediante ACWE. Por ejemplo, en la referencia (Iakovidis et al., 2006) se propone el uso de ACWE con una variación para controlar mejor las no uniformidades y ser un método de segmentación más robusto frente al ruido. Se valida dicho algoritmo sobre imágenes médicas diversas obtenidas mediante técnicas de ultrasonido, altamente ruidosas.

El modelo de Chan y Vese se aplica para la segmentación de hígado sobre MRI en la referencia (Cheng et al., 2008) donde emplea dicho método ACWE para realizar una segmentación burda inicial y generar un mapa de probabilidad de pertenencia a las diferentes clases para, posteriormente, refinar la segmentación mediante técnica de *level set* mejorada con conocimiento a priori de forma. Obtiene resultados razonables.

4.3.9.6. Contornos Geodésicos Activos o *Geodesic Active Contours* (GAC)

Aproximación teórica

Caselles propuso inicialmente los denominados Contornos Geométricos Activos o *Geometric Active Contours* (Caselles et al., 1993) como una mejora de los contornos activos partiendo de la ecuación de *level set*, (ver ec. 4.16) y añadiendo parámetros geométricos para controlar la evolución de la curva, de manera que ésta se desplazara hacia los bordes de la imagen pero a través de aquellos puntos mediante los cuales la longitud de la curva fuera menor. Un tiempo después, el mismo Caselles mejoró su propio método, y propuso en 1997 los conocidos Contornos Geodésicos Activos o *Geodesic Active Contours* (GAC) (Caselles

et al., 1997). Este nuevo modelo, en vez de partir de la ecuación de *level set* sobre la que aplica información geométrica, parte de la propuesta de Kass y sus *snakes* (ver ec. 4.8), y sobre ella presenta un algoritmo más eficiente para la evolución de la curva. Un GAC minimiza la longitud de una curva, donde se asigna menor longitud a las curvas situadas en puntos de la imagen donde el módulo del gradiente de la intensidad sea más grande.

Partiendo del modelo de la snake en el espacio discreto:

$$E_{snake}(s) = E_{int}(c(s)) + E_{image}(c(s)) \quad (4.23)$$

donde

$$E_{int} = \alpha(s) \left| \frac{dc(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2c(s)}{ds^2} \right|^2 \quad (4.24)$$

$$E_i^{image} = g \|\nabla I\| \quad (4.25)$$

con g una función detector de bordes suave, y $c(s)$ un conjunto de puntos o contorno en la imagen I .

y migrando al espacio continuo, el funcional del modelo de snakes queda de esta manera:

$$F_S(c) = \alpha \int_0^1 \|c'(q)\|^2 dq + \beta \int_0^1 \|c''(q)\|^2 dq + \lambda \int_0^1 g(\|\nabla I(c(q))\|) dq \quad (4.26)$$

Caselles et al. plantearon que el hecho de que los términos de regularidad estén controlados por α y β no implica que el contorno final sea regular. En la referencia (Caselles, 1993) se probó que para cualquier $\alpha, \beta \geq 0$, tal que $\alpha + \beta > 0$, el funcional $F_S(c)$ es capaz de extraer curvas con esquinas. Es decir, que la regularidad en la parametrización no implica regularidad en la curva. En dicha publicación se postula que el parámetro β sea posiblemente redundante y se considera el caso particular de la snake con $\beta = 0$. De esta manera, el funcional de la energía del contorno activo queda:

$$F_{S'}(c) = \alpha \int_a^b \|c'(q)\|^2 dq + \lambda \int_a^b g^2(\|\nabla I(c(q))\|) dq \quad (4.27)$$

Caselles apreció que este funcional $F_{S'}$, no es intrínseco, es decir, dos parametrizaciones diferentes de una curva no llevan a la misma solución final, lo cual no es el caso ideal para la localización de la mejor solución. Acorde a unos razonamientos en base a funciones hamiltonianas transformó el problema de minimizar $F_{S'}$, en la minimización del siguiente funcional:

$$J_{GAC}(c) = \int_a^b g(\|\nabla I(c(q))\|) \|c'(q)\| dq = \int_0^{L(c)} g(\|\nabla I(c(q))\|) ds \quad (4.28)$$

con ds el parámetro de la longitud de arco. Es decir el funcional mide la longitud de la curva, pero con un parámetro que indica un peso de cómo esa información se mide en la imagen. De esta manera la fuerza que controla la evolución de la curva finalmente se convierte en

$$F = g(I)K - \nabla g \cdot \vec{n} \quad (4.29)$$

Con K la curvatura del frente que evoluciona. En esta ecuación los únicos elementos geométricos que aparecen son la curvatura y el vector normal n al contorno que son independientes de la parametrización, y por tanto, el funcional que evoluciona es intrínseco, deseable para que la solución sea independiente de los parámetros iniciales.

Aplicación a segmentación hepática

Los *Geodesic Active Contours* de Caselles han alcanzado mucha popularidad en la comunidad científica y existen numerosos trabajos que los utilizan.

En la referencia (Foruzan and Yen-Wei, 2013), que ya ha sido citada con anterioridad en la descripción de los métodos de agrupamiento propone el uso de GAC para refinar una primera segmentación burda de hígado obtenida mediante métodos de agrupamiento no supervisado, mediante *k-means*. De esta manera, tomando la primera segmentación como inicialización del GAC, relativamente próxima a la solución final, se consiguen mejores resultados.

En la referencia (Peng et al., 2014) se evoluciona una variación de un GAC para segmentación de hígado sobre imágenes de CT, complementándolo con información de apariencia de la región. Realiza un buen análisis comparativo de este método con el tradicional GAC, aplicando la metodología establecida en el MICCAI Challenge 2007, cuyos resultados están recogidos en la referencia (Heimann et al., 2009). Los valores obtenidos por este método para cada una de las métricas sobre el conjunto de sets de prueba son muy buenos, rondando la tasa de acierto de la segmentación final el 82%. Este trabajo mejora una aproximación previa del mismo autor, referencia (Peng et al., 2011), donde obtenía un 80% final en tasa de acierto de la segmentación.

Recientemente, Oh et al. han publicado un trabajo en el que se propone un método denominado p-ESRAC, que consiste en un proceso híbrido basado en GAC, y por tanto, detección de borde, y la complementa con información de región, y por tanto de intensidad de una manera robusta frente al ruido, empleando una serie de secuencias diferentes que se van utilizando consecutivamente en los diferentes pasos del algoritmo de segmentación. Realiza una comparativa exhaustiva con GAC, obteniéndose mejores resultados con el nuevo método propuesto (Oh et al., 2014).

4.4. Técnicas variacionales para cálculo de la solución óptima de segmentación

Si bien las técnicas variacionales no pueden ser consideradas como un método de segmentación en sí mismas, son el eslabón matemático necesario para encontrar los mínimos de los funcionales de energía que representan los contornos activos (en sus múltiples variaciones). Localizar la región segmentada equivale a identificar el conjunto de puntos que minimizan dicha ecuación de energía de manera óptima. Esto implica localizar el contorno que constituye el mínimo global de dicha función, y que por tanto representa la energía mínima de la misma. Es necesario aplicar un método de minimización para encontrar la solución.

Los métodos de minimización son un amplio campo de investigación donde se concentran muchos esfuerzos. De manera burda, un candidato a mínimo o máximo se identifica mediante la resolución de la ecuación de la primera derivada, encontrando el valor que la hace 0, y el hecho de que sea máximo o mínimo se despeja a través de la segunda derivada. La localización de un mínimo global no es tan simple cuando se trata de resolver una función compleja que pueda presentar múltiples mínimos locales.

Entre los métodos de minimización se pueden encontrar desde el gradiente descendiente de primer orden hasta métodos más sofisticados como los algoritmos genéticos o evolutivos. Si se pasa del entorno discreto para resolver un modelo de contorno activo a un espacio continuo, se aplican las técnicas variacionales en segmentación de imagen. Las técnicas variacionales son problemas matemáticos que consisten en buscar máximos y mínimos de funcionales continuos definidos sobre algún espacio funcional. Constituyen una generalización del cálculo elemental de máximos y mínimos de funciones reales de una variable. En el espacio continuo se puede interpretar la imagen como una función suave, de manera que es admisible procesarla mediante estas técnicas variacionales, que además han resultado y resultan muy prometedoras y ventajosas para segmentación de regiones. Algunos ejemplos pueden encontrarse en las referencias (Peng et al., 2014), (Chen and Tagare, 2002).

Estas técnicas no son sólo empleadas en segmentación sino que también se usan en aplicaciones de eliminación de ruido (*denoising*) como se puede encontrar en las referencias (Caselles et al. 2011) (Chambolle and Caselles, 2010) (Chan et al., 2006).

Estos funcionales de energía que se definen sobre las imágenes son tales que un mínimo de dicho funcional proporcionará una solución al problema de segmentación. Las ventajas son varias ya que permiten incorporar a los modelos muchas características geométricas como gradientes, normales, curvaturas, valores estadísticos, etc. La resolución matemática al problema se puede reducir a Ecuaciones en Derivadas Parciales, que es un problema abordado ampliamente desde el punto de vista matemático y que puede ser traducido al espacio discreto.

Sin embargo, estas técnicas, si bien potentes, también presentan sus limitaciones, principalmente en lo relativo a la localización del mínimo global, esto es, que, durante la

evolución de la función, ésta no se quede estancada en un mínimo local, que no conduce a la segmentación óptima.

El problema matemático surge del hecho de que sólo las funciones estrictamente convexas pueden converger a un mínimo global. Las que sólo son convexas, pueden estancarse en mínimos locales. Uno de los métodos para solventar esta situación es el propuesto en la referencia (Chambolle, 2004). Chambolle desarrolló una aproximación matemática en la denominada formulación dual, en la que propone transformar una función de energía convexa en una función estrictamente convexa para poder alcanzar el mínimo global. Para ello, añade una penalización, otra función v , de manera que el cálculo del mínimo se realiza fijando v y minimizando con respecto a u ; y fijando u y minimizando con respecto a v separadamente. Ambas dos ecuaciones iteran hasta llegar a la convergencia. Esto se explicará en más detalle en el capítulo 5. Esta línea ha sido trabajada por matemáticos relevantes como Bresson en las referencias (Bresson and Chan, 2007)(Bresson et al., 2007) donde se pueden encontrar los detalles matemáticos de la implementación.

4.5. Segmentación hepática en 3D

Cualquiera de los métodos expuestos en el apartado 4.3 puede ser extendido a 3D con un esfuerzo razonable. La aproximación teórica es la misma pero en la implementación se añade la componente tridimensional en todos los cálculos asociados.

El uso de la tercera dimensión puede ser valioso para la estimación de la continuidad de los elementos, y no sólo teniendo en cuenta los ejes x e y . El gran inconveniente de estos métodos es que el procesamiento del volumen es, en general, mucho más costoso computacionalmente hablando que sobre una imagen 2D, y por tanto, son métodos lentos, que, dependiendo del caso de aplicación, deben ser rechazados. El desarrollo de ordenadores más potentes, con varios procesadores que permiten realizar tareas en paralelo, así como las GPU redundan en una reducción considerable del tiempo de cómputo necesario. Este análisis escapa del alcance de esta tesis.

Se ha identificado en la literatura algunos trabajos que realizan la segmentación del hígado en 3D. En la referencia (Rusko and Bekes, 2007) se propuso un método de crecimiento de regiones en 3 dimensiones para realizar posteriormente un post-procesado que permitiera seleccionar el hígado en imágenes de CT.

Okada et al. realiza una segmentación automática del hígado sobre imágenes CT basado en dos técnicas: atlas probabilístico en 3D para la localización de la región del hígado y modelos estadísticos de forma multinivel para refinar el volumen de hígado previamente localizado (Okada et al., 2008).

En la referencia (Gloger et al., 2010) se propone una segmentación automática sobre imágenes MRI del hígado en 3 fases basada en un crecimiento de regiones modificado y un umbralizado posterior. Se reduce la dimensión del vector de entrada con todas las características empleadas mediante análisis discriminante lineal multiclase y se generan

mapas de probabilidad que se usan posteriormente para la segmentación. Para el refinamiento de los resultados emplea conocimiento a priori.

En la referencia (Gambino et al., 2010) se propone un crecimiento de regiones en los 3 ejes simultáneamente sobre imágenes de CT basado en descriptores de textura obtenidos a partir de la matriz de coocurrencia (Haralick, 1979).

Entre los métodos de segmentación 3D sobre imágenes MRI identificados, se menciona la referencia (Reska et al., 2012) que propone una isosurface activa para la localización del hígado en MRI. Los autores comparan su método con respecto a una implementación de *snakes* en 2D y un algoritmo de crecimiento de regiones. Los resultados son comparables a los del estado del arte, pero mediante la isosurface activa la reducción del tiempo de cómputo es notable.

Cabe mencionar el software Mazda como una herramienta válida para la segmentación 2D y 3D de imágenes médicas tanto obtenidas mediante CT como por MRI (Strzelecki et al., 2013). La segmentación de los diferentes órganos se realiza a partir de descriptores de textura. El software permite la selección de todas las componentes del proceso, desde los descriptores de textura a emplear, el método para reducir la dimensionalidad, o el método de clasificación para obtener la región final segmentada. En el caso del hígado se han encontrado limitaciones y problemas para realizar una correcta segmentación sobre MRI. El principal inconveniente de este software es que se requiere un gran conocimiento experto por parte del usuario en técnicas de procesamiento de imagen y en clasificadores para poder obtener unos resultados razonables. Este conocimiento específico rara vez se encuentra en la práctica clínica, donde los métodos automáticos son los adecuados.

En la referencia (Huynh et al., 2014) se propone una segmentación del hígado en tres fases. Inicialmente se aplica una etapa de pre-procesamiento a la secuencia T1 con contraste para reducir el ruido y generar una imagen con contorno remarcado. Esta imagen mejorada se emplea como la función de velocidad de un algoritmo *fast - marching* en 3D para generar una primera superficie que de manera burda se parece a la forma del hígado. A partir de esta forma, que constituye la inicialización de un *Geodesic Active Contour* 3D se refinan los resultados de segmentación del hígado.

Goçeri et al. proponen un método de *level set* automático. Este método define inicialmente y de manera automática un contorno inicial, y calcula los valores de los pesos de cada término en el funcional de energía en cada iteración al mismo tiempo que evoluciona. Es decir, aborda al mismo tiempo y de manera automática dos limitaciones de los *level set*, que son la necesidad de una buena inicialización y la actualización de los pesos de cada término en el funcional de energía (Göçeri et al., 2014).

En esta tesis, tal y como se describirá en la aproximación teórica del capítulo 5, se va a realizar la extensión a 3D de un modelo de contorno activo en 2D. El contorno activo pasa a denominarse superficie activa. La base de funcionamiento es la misma. Dichas superficies definidas se desplazan y evolucionan en base a un criterio de minimización de energía, que

es calculada en base a características del volumen (gradientes, bordes) y que influye en la evaluación de las posibles posiciones de desplazamiento hasta converger a una solución. Pueden estar construidas con medidas estadísticas de parecido o con modelos de minimización de energía que balancean las fuerzas externas, de seguimiento de bordes, y las internas, que intentan preservar cierto suavizado.

La principal ventaja de usar superficies activas es que se pueden controlar a priori las características geométricas de los bordes, evitando agujeros, generando superficies con bordes suavizados, estructuras suavemente convexas o con pocas irregularidades. En teoría, la continuidad de los elementos en los 3 ejes es evaluada y por tanto el resultado de una segmentación en 3 dimensiones, teniendo el volumen como entrada va a ser mejor que un método que segmente una a una cada una de las imágenes 2D que representan los cortes del volumen que constituye un estudio real de cuerpo. Sin embargo, esta ventaja es al mismo tiempo su principal inconveniente dado que, cuando los elementos presentan bordes abruptos, estructuras pequeñas o de cambios de curvatura acentuados, se pierde definición, como es el caso del sistema vascular o los nervios.

En esta tesis se ha propuesto extender a 3D un contorno activo definido en 2D. Además su funcional de energía se va a minimizar mediante la propuesta aproximación dual al problema variacional de Chambolle, que también será extendida a una tercera dimensión, de manera que constituya un método de segmentación 3D.

4.6.Segmentación hepática sobre estudios multisequencia en MRI

Un tema importante abordado en esta tesis es la utilización de más de una secuencia de MRI como entrada para mejorar los resultados de la segmentación de hígado. Los escáneres de resonancia magnética pueden ser configurados atendiendo a diferentes protocolos de adquisición de manera que el campo electromagnético generado es distinto, lo que provoca una respuesta distinta de los tejidos humanos. Este procedimiento se usa habitualmente en la práctica clínica ya que existen ciertos patrones de comportamiento de las patologías y tejidos sanos identificados por los radiólogos que ayudan en el diagnóstico. Por ejemplo, es posible que una patología se manifieste con tono “claro” en una secuencia T1 y con tono “oscuro” en una secuencia T2. Este comportamiento ayuda a discriminar su naturaleza y establecer el diagnóstico.

Con esta premisa, parece razonable asumir que el uso combinado de varias secuencias va a contribuir favorablemente a la caracterización y localización de los diferentes elementos presentes en la imagen abdominal.

Existen dos inconvenientes principales en el manejo de múltiples secuencias. En primer lugar, las diferentes secuencias deben estar convenientemente registradas. El registro consiste en que cada píxel de la imagen en todas las secuencias debe referirse al mismo punto físico, para, de esta manera, tener la información múltiple alineada. El registro perfecto es difícil de conseguir en determinados casos, como es éste, en el que el hígado se ve afectado por movimientos respiratorios mientras la persona está siendo escaneada. El

problema del registro es un campo de investigación en sí mismo. No se va a abordar en la presente tesis.

La segunda cuestión importante que se debe resolver en una solución con múltiples secuencias es cómo combinar las mismas. Esto no es trivial ya que los resultados de la segmentación dependerán de ello.

Alguno de los trabajos preliminares identificados en la literatura han sido aplicados a clasificación de diferentes tejidos cerebrales, donde se ha considerado un valor añadido de la información múltiple sobre el mismo punto físico, basándose en operativas previas propuestas por la NASA para aplicaciones de detección remota (Vannier et al., 1991).

Cabe indicar que el número de referencias al respecto en el estado del arte para segmentación de hígado son limitadas. Esto es lógico partiendo de la base de que la técnica de adquisición más extendida en el procesamiento automático y segmentación de hígado es CT, que sólo tiene una secuencia, se adquiere en pocos segundos, se puede realizar en apnea y por tanto, no presenta grandes artefactos en la imagen. El número de trabajos que abordan la segmentación hepática sobre MRI es menor en general, y por tanto, el número de referencias localizadas que además empleen la opción multicanal es muy reducido. Se describen a continuación las localizadas y la manera en que gestionan las diferentes entradas.

En el estado del arte se pueden identificar dos tipos de enfoque para el manejo de las múltiples secuencias. Uno de ellos consiste en tratar cada secuencia por separado. Se diseña un algoritmo de segmentación común que es aplicado sobre cada canal, y los resultados obtenidos sobre cada uno de ellos se fusionan y combinan al final, para proporcionar la solución definitiva. Este es el caso de la referencia (Platero et al., 2008a) que propone un método para segmentar el hígado mediante el uso de dos canales, T1 y T2. Cada imagen de cada secuencia es procesada mediante el siguiente algoritmo: primero, se selecciona el máximo y mínimo del histograma entre los cuales se supone que el hígado está representado; a continuación se realiza una segmentación burda mediante detector de borde de Canny y operaciones morfológicas que es refinada a continuación mediante contornos activos. Se trata de un método semiautomático y se precisa de interacción por parte del usuario. Se aplica sobre los canales T1 y T2. Finalmente, las regiones resultantes obtenidas de la aplicación del algoritmo en ambas secuencias son fusionadas para proporcionar la región segmentada final.

El otro enfoque empleado para el manejo de la información multicanal, que es el más extendido, consiste en considerar la adición de nuevos canales de información como un problema de una dimensión mayor. Es decir, cada píxel viene representado por tantos niveles de gris como secuencias existen. El algoritmo de segmentación se aplica sobre este vector de características en la entrada, lo que aumenta la dimensionalidad del problema matemático a resolver. Computacionalmente hablando es más costoso. En esta línea se han identificado los siguientes trabajos.

Lin et al. combinan el uso de tres secuencias T2 de MRI a las que se le ha modificado el TR (Tiempo de Repetición) y el TE (Tiempo de Eco) para segmentación de cerebro y de hígado. El método propuesto usa FCM (*Fuzzy c-means*) penalizado. Fuzzy C-means es un método no supervisado para clasificar en base a centroides, empleando el mismo enfoque que propone el método *k-means* con niveles de intensidad (u otros) pero con grados de pertenencia basados en Fuzzy. Es decir, calcula la matriz de grados de pertenencia, calcula los centroides, recalcula la matriz y así iterativamente. Usa Fuzzy con penalización y una variación de Redes Neuronales Artificiales. De igual manera, la entrada es un vector de 3 valores. Se obtiene una única salida como resultado de la segmentación. La presentación de los resultados no es muy rigurosa en este trabajo, ya que no se evalúan con métricas o con una metodología descrita. Se muestran algunas imágenes de cortes específicos seleccionados pero cuyos resultados de la segmentación están sin cuantificar (Lin et al., 1996).

En la referencia (Iftekharuddin et al., 2006) se emplean 3 secuencias T1, T2 y FLAIR de MRI sobre imágenes de cerebro obtenidas mediante resonancia magnética. Sobre ellas, se detectan tumores mediante técnicas SOM (*Self Organizing Maps*) y los clasifica mediante técnicas avanzadas de SVM (*Support Vector Machine*). Los resultados de clasificación de los tumores aumentan notablemente si se emplea como entrada un vector de características con 3 componentes que se corresponden cada una con el valor de nivel de gris de ese pixel en cada secuencia. No se hace ningún procesamiento extra sobre la información multiseccuencial.

En otro trabajo identificado, Tankyevych emplea 3 secuencias de MRI para segmentar el hígado, músculo y grasa en las imágenes abdominales. El método propuesto consiste en aplicar Fuzzy c-means sobre un vector de 3 características, de manera que la incorporación de las secuencias supone un aumento de la dimensión del problema. El algoritmo de segmentación se basa en clasificación no supervisada y los resultados no son muy finos. Como conclusión de este trabajo, se extrae que el uso de múltiples entradas contribuye a la mejora en los resultados obtenidos (Tankyevych, 2006).

Este enfoque consistente en aumentar la dimensión del problema parece válido, ya que los resultados de segmentación se ven mejorados en todos los casos. El mayor inconveniente radica en que si el método de segmentación es complejo y el número de secuencias disponible aumenta, el tiempo necesario para ejecutarlo puede ser muy elevado, y es posible que los requerimientos de memoria no sean soportados por los ordenadores habituales en la práctica clínica.

Otra manera de abordar la información multiseccuencial para la segmentación automática de hígado sobre imágenes MRI se puede encontrar en la referencia (Oh et al., 2014), un método reciente, ya incluido en el apartado de *Geodesic Active Contours*. En él, los autores proponen un algoritmo donde en cada fase del mismo se emplea una determinada secuencia de las disponibles, de manera que la salida de cada paso del algoritmo constituye la entrada del siguiente estadio de procesamiento donde se usa una secuencia diferente. Por ejemplo, la imagen de fase retardada se usa para extraer el punto semilla que será usada en la imagen

de fase arterial. No combinan la información multiseuencia en sí, pero han diseñado un algoritmo donde se utilizan en cadena, hasta conseguir el resultado final, que además es muy bueno, acorde a las métricas que muestra. Este método se ha incluido en la discusión y comparación de los resultados obtenidos mediante el método propuesto en esta tesis.

4.7. Limitaciones de los métodos existentes para la segmentación de hígado en MRI

Una vez descrito el estado del arte e identificadas las limitaciones de los métodos de segmentación en general, y de segmentación hepática sobre imágenes adquiridas mediante MRI, en particular, se concluye que es necesario el desarrollo de nuevos métodos de segmentación capaces de realizar la localización del hígado de manera rápida, automática y salvando las dificultades donde los métodos más clásicos no llegan.

Como se ha expuesto anteriormente, existe un número creciente de trabajos que proponen métodos para segmentación del hígado sobre imagen de resonancia magnética, pero la tomografía axial computerizada (TAC o CT) sigue siendo la técnica más usada. La **utilización de resonancia magnética multiseccional** permite una descripción matemática mucho más rica y precisa de las estructuras anatómicas contenidas en la imagen al estar éstas descritas por sus propiedades en las diferentes secuencias (cada punto de la imagen está asociado a un valor numérico por cada una de las secuencias adquiridas formando un vector). Esto va a permitir desarrollar modelos de órganos mucho más precisos. Sin embargo, el número de trabajos sobre MRI, aunque está aumentando en los últimos años, es menor que el de los que trabajan sobre CT. La razón fundamental es porque, dentro de la dificultad que entraña el procesamiento de imagen médica en general, las imágenes de CT presentan mayor contraste que las de MRI. Las imágenes de MRI presentan mayor número de artefactos, el efecto de volumen parcial, con bordes difusos, poco contraste entre elementos, más ruido, zonas de iluminación no homogéneas, y en definitiva, mayor complejidad.

Las mayores limitaciones existentes a la hora de segmentar el hígado en un volumen bien de CT (menor riesgo) o de MRI es la dificultad en la delimitación del contorno del hígado (debido al efecto de volumen parcial sobre imágenes de resonancia magnética), sobre todo en los cortes centrales donde el hígado se pega a la pared abdominal. También puede ser problemática la distinción con respecto a otros órganos adyacentes, como son el intestino, los pulmones o los riñones, ya que los niveles de gris que muestra la imagen para la mayor parte de estos órganos son muy similares. El elevado nivel de ruido en las imágenes de MRI y las zonas que aparecen como borrosas o poco nítidas como resultado del movimiento de la persona (involuntario y respiratorio) mientras dura la captura (más lenta que sobre CT) tampoco contribuye a la obtención de buenos resultados en la segmentación de órganos y lesiones.

Muchos de los métodos desarrollados en el estado del arte se proponen como métodos semiautomáticos donde, de manera interactiva, el usuario señala puntos de interés o delimita zonas, para sortear estos problemas y minimizar los errores. En general obtienen

mejor resultado que los métodos automáticos pero implican una ralentización de la respuesta y un gran consumo de recursos, de manera que algunos emplean hasta 30 minutos en hacer una segmentación. Este tiempo es demasiado largo en la práctica clínica habitual.

Se ha mostrado en este capítulo que hay estudios que exploran la vía del uso simultáneo de varias secuencias obtenidas sobre la misma región física, de manera que la información complementaria puede ser valiosa para la mejora en la segmentación y caracterización de las lesiones presentes. Sin embargo, en el caso de imágenes obtenidas mediante resonancia magnética, donde cada secuencia puede suponer un tiempo de captura de 45 segundos, la adquisición en apnea es difícil, y por tanto, se acusan movimientos respiratorios entre las imágenes de las diferentes secuencias que se presentan desalineadas, es decir, cada píxel en cada secuencia no se corresponden exactamente con el mismo punto físico. Este es el problema del registro. El desarrollo de técnicas de registro es en sí mismo un campo muy extenso en procesamiento de imagen médica.

Por otra parte, se ha constatado que, en los trabajos identificados en la literatura y recogidos a lo largo del estado del arte en el presente capítulo, no todos los métodos han sido rigurosamente evaluados. Casi siempre se muestran imágenes con el hígado localizado, pero no todos los casos indican métricas y pruebas de validación acorde a las mismas y una segmentación de referencia establecida. Este hecho también quiere ser abordado en la presente tesis, mediante la definición de una metodología de medida y evaluación apropiada.

En esta tesis se trabaja sobre imágenes de MRI. Las técnicas de adquisición de imagen médica por resonancia magnética permiten disponer de varias secuencias diferentes obtenidas con diferentes protocolos de adquisición, en cuanto al campo electromagnético generado sobre el cuerpo humano. Esto proporciona una respuesta diferente de los distintos órganos y tejidos a cada estímulo recibido, y por tanto, información complementaria en las distintas secuencias.

Finalmente, el empleo de varias secuencias simultáneas en los métodos que trabajan sobre imágenes de MRI es muy reducido. Es aquí donde se ha visto la oportunidad de abordar un nuevo método sobre MRI que contemple la inclusión de varias secuencias en el mismo estudio. Estas múltiples secuencias aportan información complementaria de gran valor, que, presumiblemente, van a mejorar los resultados de la segmentación.

4.8. Conclusiones

El análisis detallado del estado del arte ha constatado que el desarrollo de nuevos métodos de segmentación es un área de trabajo interesante. Ha sido abordado de diferente manera por distintos investigadores, dando lugar a multitud de métodos, donde en los últimos años las técnicas variacionales para minimización de contornos o superficies activas está tomando un peso relevante en cualquiera de sus versiones. El tema objeto de esta tesis es de interés en la comunidad científica y con una amplia posibilidad de aplicación al tratamiento de la imagen médica, debido a la cantidad de órganos y tejidos que se pueden identificar

gracias a las tecnologías de adquisición de MRI, CT, PET, US, etc., que han experimentado un tremendo auge y una implantación masiva en las clínicas y hospitales para ayuda al diagnóstico como técnica no invasiva.

La aplicación de nuevos métodos de segmentación a la localización del hígado es un posible caso de aplicación, siendo un órgano de interés dada la elevada incidencia del cáncer de hígado y la relevancia que tienen un buen diagnóstico y la correcta determinación de la zona de la lesión para minimizar los riesgos de una intervención quirúrgica, y en definitiva, la calidad asistencial.

En esta coyuntura surge el tema propuesto en esta tesis, el desarrollo de un nuevo método automático para la segmentación de hígado mediante minimización de un contorno activo a través de la aproximación dual de Chambolle al problema variacional y con la propuesta de un nuevo descriptor compacto capaz de representar en una única secuencia la información disponible en varias secuencias, además, basado en modelo, lo que mejora la calidad resultante de la entrada y reduce de manera notable el tiempo de proceso necesario. El método se ha extendido a 3D. La validación ha sido exhaustiva mediante la implementación y cálculo de 9 métricas establecidas sobre una segmentación de referencia. Se ha comparado con los métodos del estado del arte que también realizan segmentación hepática sobre MRI. En el capítulo 5 se expone la aproximación teórica al método propuesto. En el capítulo 6 se describe la metodología de pruebas y los resultados obtenidos, para pasar a las conclusiones en el último capítulo.

Capítulo 5: APROXIMACIÓN
TEÓRICA

El objetivo de esta tesis es el desarrollo de nuevos algoritmos avanzados de procesamiento de imagen para la segmentación (delimitación) automática del hígado en imágenes de resonancia magnética.

En el capítulo 2 se ha mostrado la problemática inherente a la imagen médica debida a múltiples causas, y por tanto, la dificultad de ser procesada automáticamente y poder extraer información valiosa y precisa de ella. A lo largo del capítulo 3 se ha hecho un recorrido sobre los descriptores que son susceptibles de ser obtenidos de la imagen para componer el vector de características de entrada en un proceso de segmentación o clasificación. Por último, en el capítulo 4 se han presentado las técnicas de segmentación en 2D y 3D relevantes del estado del arte. Este estudio ha sido el punto de partida para los desarrollos realizados en esta tesis. A partir del conocimiento adquirido en el estado del arte y recopilado en los capítulos mencionados, se propone un nuevo método que permita mejores resultados en la segmentación automática de imagen médica.

Para llevar a cabo este proceso se trabaja sobre tres elementos clave:

- Descriptor compacto basado en modelo
- Modelo de contorno activo
- Método de minimización del contorno activo

Tal y como se ha explicado en capítulos anteriores, el procesamiento de una imagen cualquiera implica varias fases, tales como pre-procesamiento, extracción de características, obtención de descriptores o aplicación de métodos de segmentación o clasificación, que van a depender del objetivo final del estudio. En la medida en que se adecúen los métodos elegidos en cada una de dichas etapas al problema a resolver, se obtendrán mejores resultados en la segmentación y/o clasificación final.

El descriptor compacto que se propone es capaz de combinar la información de las múltiples secuencias de entrada y transformarlas en una sola entrada. Este descriptor se construye en base a un modelo estadístico multivariable de hígado previamente establecido. La secuencia o conjunto de imágenes de la nueva entrada generada mediante el descriptor compacto presenta un contraste mejorado con respecto a cada una de las secuencias iniciales por separado. Por ello, el descriptor compacto propuesto adopta una doble funcionalidad, de descriptor en sí y como método de pre-procesamiento para mejorar la calidad de la entrada.

El método de segmentación propuesto consiste en la minimización de un contorno activo. Es necesario definir el modelo de contorno activo y cómo minimizarlo. Ambos pasos van a permitir delimitar la región hepática en la imagen abdominal obtenida mediante tecnología MRI.

A partir de los dos modelos descritos en los apartados 4.3.9.5 y 4.3.9.6, Contornos Geodésicos Activos (*Geodesic Active Contours* o GAC) y (Contornos Activos Sin Borde (*Active Contours Without Edges* o ACWE), muy empleados en segmentación de imagen, tal

y como se ha visto en el estado del arte del capítulo 4, se han realizado múltiples variaciones en los últimos años. Estas variaciones han ido orientadas en dos líneas de investigación principalmente. Por una parte, se trabaja en la mejora de los modelos de contornos activos para que sean más robustos frente al ruido, solapes o zonas no uniformes en la imagen, entre otros. Por otra parte, se trabaja en técnicas matemáticas complejas, como son las técnicas variacionales, para resolver las limitaciones matemáticas que surgen en la resolución de los problemas representados por los contornos activos, como el hecho de que algunos funcionales de energía que dichos contornos representan no sean estrictamente convexos. Esto implica que no presentan un mínimo global, sino mínimos locales, donde pueden quedar estancados en el proceso de minimización, lo que ocasiona que el funcional de energía no converja adecuadamente en la solución óptima. Por consiguiente, se han ido proponiendo nuevos algoritmos, previa rigurosa justificación matemática, que sean capaces de hacer converger el funcional de energía en este proceso de minimización de una manera rápida y eficiente. Estas técnicas no sólo se han aplicado a segmentación sino que también son muy útiles en eliminación de ruido.

En esta tesis se ha elegido la aplicación de técnicas variacionales como método de minimización de un contorno activo definido. Entre las posibles técnicas variacionales, se ha escogido trabajar con la aproximación dual al problema variacional propuesto por Chambolle en 2D, que posteriormente ha sido trabajado por otros autores en problemas de segmentación y eliminación de ruido.

Además, en esta tesis, el modelo de contorno activo y el método de minimización asociado han sido extendidos de dos a tres dimensiones, es decir, el contorno activo se transforma en una superficie activa en un espacio de dimensión mayor. Ambos elementos, tanto el implementado en 2D como el 3D, son el núcleo de la metodología para la segmentación hepática sobre imagen obtenida mediante tecnología MRI con dos tipos de entrada distinta: monosecuencia y multisecuencia.

Finalmente, en este capítulo se describe el flujo completo del algoritmo, que se ejecuta de manera automática sin que el usuario tenga que interactuar con la aplicación en absoluto.

Se detallan en cada apartado todos los elementos mencionados en esta introducción.

5.1. Concepto de monosecuencia y multisecuencia

Monosecuencia (o secuencia) es el conjunto de imágenes (o cortes) obtenidos como resultado de la recepción de la respuesta proveniente de los tejidos del cuerpo a la emisión de señales electromagnéticas generadas por una máquina de resonancia magnética. Una secuencia contiene un número de cortes determinados (*slices*) que cubren el volumen completo del cuerpo que se ha decidido analizar, constituyendo cada corte una de las imágenes 2D. El conjunto de los cortes conforman la reconstrucción del volumen 3D.

Se entiende como multisecuencia la recogida de más de una secuencia en el proceso de captura. La máquina de resonancia magnética puede ser configurada con diferentes

protocolos de adquisición mediante la modificación de las características del campo electromagnético generado, lo cual provoca que la respuesta de los tejidos del cuerpo ante dicho campo sea distinta. Esta información se traduce en imágenes que presentan niveles de gris diferentes que caracterizan los mismos órganos y tejidos estimulados por estos campos distintos.

La obtención y análisis de varias secuencias en el mismo estudio se realiza de manera habitual en la práctica clínica para las tareas de diagnóstico, ya que es posible que determinada patología se muestre con tono “claro” en una secuencia y con tono “oscuro” en otra, y el diagnóstico asociado se pueda discriminar de esta manera. Es por ello que el uso combinado de varias secuencias, todas ellas referidas al mismo órgano, proporciona información complementaria que puede contribuir favorablemente en el proceso de segmentación.

Para la realización de esta tesis, se ha podido disponer de imágenes provenientes de 5 secuencias en estudios sobre pacientes adquiridos por la máquina Avanto de Siemens de 1.5 T de la que dispone el departamento de Diagnóstico por Imagen de la Clínica Vicente San Sebastián, ubicada en IMQ – Zorrozaurre (Bilbao). Estas secuencias elegidas se denominan en cada estudio médico de la siguiente manera:

- Secuencia 00: T2 BLADE
- Secuencia VIBE 05: en vacío axial
- Secuencia VIBE 06: arterial axial
- Secuencia VIBE 07: venoso axial
- Secuencia VIBE 08: tardío axial

El protocolo VIBE (*Volumetric Interpolated Breath-hold Examination*) se corresponde con la secuencia ultra rápida gradiente - eco que Siemens proporciona. Estas 5 secuencias constituirán la entrada de las pruebas denominadas “multisequencia” en cada uno de los métodos desarrollados y evaluados. Para la evaluación de los métodos sobre una sola secuencia se ha escogido la secuencia VIBE 05 en vacío axial, ya que, acorde a la información médica recibida, es una de las más utilizadas en la práctica clínica debido a la resolución y contraste que presenta. La información concreta de los parámetros de configuración principales de dichas secuencias se proporcionará en el capítulo 6, en la descripción del conjunto de datos disponibles para las pruebas.

Todas las secuencias representan la respuesta de la misma zona corporal ante distintos estímulos. Al emplear varias secuencias simultáneamente aparece el problema del registro, descrito en el capítulo 2, debido principalmente a los movimientos respiratorios, lo que puede ocasionar que el mismo píxel del mismo corte en las distintas secuencias no esté referido exactamente al mismo punto físico. Estos desplazamientos van a provocar errores en la segmentación. Se va a evaluar en el capítulo 6 de qué orden son estos errores y cómo afecta el problema de registro en este caso. En el capítulo 6, se detalla la metodología de las

pruebas realizadas y se muestran los resultados obtenidos para los métodos automáticos de segmentación desarrollados.

5.2. Descriptor compacto basado en modelo: generación de un mapa de distancias

El hecho de disponer de varias secuencias que proporcionan datos sobre los mismos puntos físicos del cuerpo es de gran valor para la segmentación, ya que se dispone de más información, en principio, complementaria. De hecho, el uso de más de una secuencia para realizar el diagnóstico es habitual en la práctica clínica, con lo que se pretende extraer información de manera automática a partir de las múltiples entradas con el objetivo final de que la segmentación resultante será más precisa.

Además, el uso de varios canales simultáneos puede también compensar algunos de los problemas y artefactos típicos de las imágenes de resonancia magnética, sobre todo, los efectos generados comparables a una iluminación no uniforme de la escena. Este problema se aborda a menudo con el uso de filtros anisotrópicos que intentan homogeneizar la imagen o con técnicas de mejora de calidad de la imagen (Perona and Malik, 1990) (Manjón et al., 2012) (Suzuki et al., 2013).

Sin embargo, en esta tesis se propone el uso combinado de múltiples secuencias para compensar este efecto de iluminación no uniforme de la escena y el bajo contraste existente entre algunos tejidos adyacentes. Esta es la hipótesis de partida para justificar el uso de la entrada multisequencia en el proceso de segmentación.

La cuestión que se plantea es cómo tratar la información de las múltiples secuencias. Desde el punto de vista de procesamiento de imagen, la entrada multisequencia, en este caso, de 5 secuencias diferentes, se traduce en la existencia de un vector de 5 características por cada píxel de la imagen.

Como se ha detectado en el análisis previo del estado del arte recogido en el capítulo 4, existen dos tendencias principales para el manejo de la entrada multisequencia. Una de las maneras de abordarlo es emplear directa y simultáneamente los 5 valores del vector, lo que aumenta la dimensión del problema a resolver y en definitiva, del funcional de energía a minimizar. Otra opción es reducir la dimensión del problema mediante la aplicación de técnicas de reducción de características, como puede ser Análisis de Componentes Principales (conocido como *Principal Component Analysis* o PCA), o Análisis Discriminante Lineal (conocido como *Lineal Discriminant Analysis* o LDA). Existen referencias de uso de estas técnicas en imagen médica (Heimann et al., 2007) (Qu et al., 2008).

Adicionalmente, la elección habitual de procesamiento es el análisis basado en píxel. En el enfoque adoptado en esta tesis se propone que las múltiples secuencias de entrada sean combinadas a través de un descriptor compacto estadístico que opera en un vecindario, que minimiza la variabilidad y mejora el contraste en la imagen resultante a partir de las

imágenes de MRI. Con este enfoque se asume una distribución gaussiana en la región representada sobre el hígado en la nueva imagen de entrada generada a partir de la combinación de las múltiples secuencias.

La propuesta en esta tesis es la generación, a partir de las imágenes de las 5 secuencias diferentes y del descriptor compacto propuesto, de un mapa de distancias en el que el valor de cada píxel representa la distancia que existe entre ese punto y un modelo gaussiano multivariable de hígado generado previamente. De esta manera, la entrada de 5 dimensiones se transforma en un mapa de distancias monodimensional sobre el que se realizará el proceso de segmentación. Este mapa de distancias o probabilístico de pertenencia a la clase hígado contiene la información complementaria y de valor proveniente de las 5 secuencias disponibles en cada estudio de resonancia magnética.

En este caso concreto, hay 5 secuencias disponibles en cada estudio abdominal por paciente. Estas secuencias fueron seleccionadas de acuerdo a la práctica clínica. El método propuesto es susceptible de ser extendido a otro número y tipo de canales. Cambiaría el modelo aprendido pero la metodología sería similar.

Las secuencias iniciales seleccionadas son T2 BLADE y 4 secuencias VIBE: axial en vacío, axial arterial, axial venoso y axial tardío. El protocolo BLADE es una secuencia T2 ponderada con eco turbo spin que permite una reducción substancial de los efectos del movimiento y presenta un mayor contraste comparada con las tradicionales T2 ponderadas o con peso. BLADE es el nombre propietario para las líneas paralelas superpuestas periódicamente rotadas con reconstrucción mejorada (PROPELLER) en los sistemas de MRI de Siemens (Pipe, 1999)(Hirokawa et al., 2008). El protocolo de las secuencias VIBE (*Abdominal Imaging with a Volumetric Interpolated Breath-hold Examination*) es una secuencia ultra rápida de eco gradiente, (Rofsky et al., 1999) y en este caso, también ha sido proporcionada por la máquina de resonancia de Siemens.

Con el objetivo de combinar las secuencias en un descriptor compacto, cada píxel del volumen 3D multiseccional captado es tratado como un vector, donde cada componente se corresponde con el valor de intensidad del píxel (i, j) en cada corte k y en cada secuencia disponible. Sin embargo, este enfoque no integra la variabilidad textural en las diferentes regiones de la imagen. Con el objetivo de proporcionar un descriptor capaz de combinar la información multicanal y la variación espacial se propone un descriptor $T(i, j, k)$ en el que el descriptor es una concatenación del valor medio de $I(i, j, k)$ en un vecindario del píxel (i, j, k) con el valor de desviación estándar de $I(i, j, k)$ en el vecindario.

$$T_{mean}(i, j, k) = \frac{\sum_{x=i-M}^{x=i+M} \sum_{y=j-N}^{y=j+N} \sum_{z=k-P}^{z=k+P} I(x, y, z)}{(2M-1) \cdot (2N-1) \cdot (2P-1)} \quad (5.1),$$

$$\mathbf{T}_{std}(i, j, k) = \sqrt{\frac{\sum_{x=i-M}^{x=i+M} \sum_{y=j-N}^{y=j+N} \sum_{z=k-P}^{z=k+P} (I(x, y, z) - \mathbf{T}_{mean}(i, j, k))^2}{(2M-1) \cdot (2N-1) \cdot (2P-1)}} \quad (5.2),$$

$$\mathbf{T}(i, j, k) = \{\mathbf{T}_{mean}(i, j, k), \mathbf{T}_{std}(i, j, k)\} \quad (5.3),$$

Donde M,N,P representa el tamaño del vecindario. Este descriptor contiene la variación espacial en cada secuencia y la información multisequencia de manera compacta.

Por otra parte, el conocimiento a priori se incluye mediante la integración del descriptor previo en un modelo estadístico multivariable que lo combina en una imagen de probabilidad. La probabilidad de pertenencia a la clase hígado está relacionada con una estimación de la distancia de cada píxel a dicha clase, asumiendo la clase como un modelo de distribución gaussiano. Los píxeles que presentan menores distancias pertenecen a hígado. En este punto, se ha considerado una distancia de Mahalanobis, a través de la siguiente ecuación 5.4.

$$d(i, j, k) = \sqrt{\sum_{s=1}^N \frac{(T_k(i, j, k) - \mu_s)^2}{\sigma_s^2}} \quad (5.4),$$

siendo $T_k(i, j, k)$ el valor de cada componente del descriptor T , μ_s el valor medio del modelo del hígado en cada secuencia específica, σ_s es la desviación estándar y N el número de secuencias disponible en la entrada.

De acuerdo a la ecuación (5.4) cada píxel de la imagen de distancias es el resultado de calcular la distancia de Mahalanobis entre el valor del mismo píxel en cada secuencia con respecto al valor medio del modelo de hígado en relación a la desviación estándar en esa misma secuencia en el modelo estadístico previamente generado. El enfoque propuesto se puede aplicar para una sola secuencia ($N = 1$) o a tantas como se requiera.

Se muestra a continuación, en la figura 5.1 un ejemplo del mapa de distancias generado para una entrada monosequencia sobre VIBE en vacío axial. En la figura 5.2 se muestran las 5 secuencias disponibles con información complementaria, y finalmente el mapa de distancias con respecto al modelo de hígado que combina la información y permite la reducción de la dimensión del problema mediante el descriptor compacto definido.

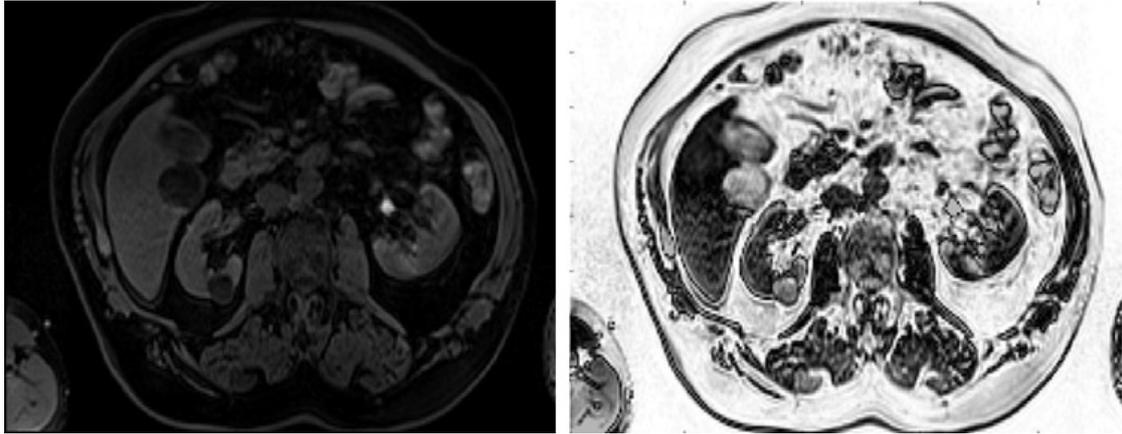


Figura 5.1. Mapa de distancias (dcha) de la entrada monosecuencia (izda) con respecto al modelo de distribución gaussiana del hígado generado

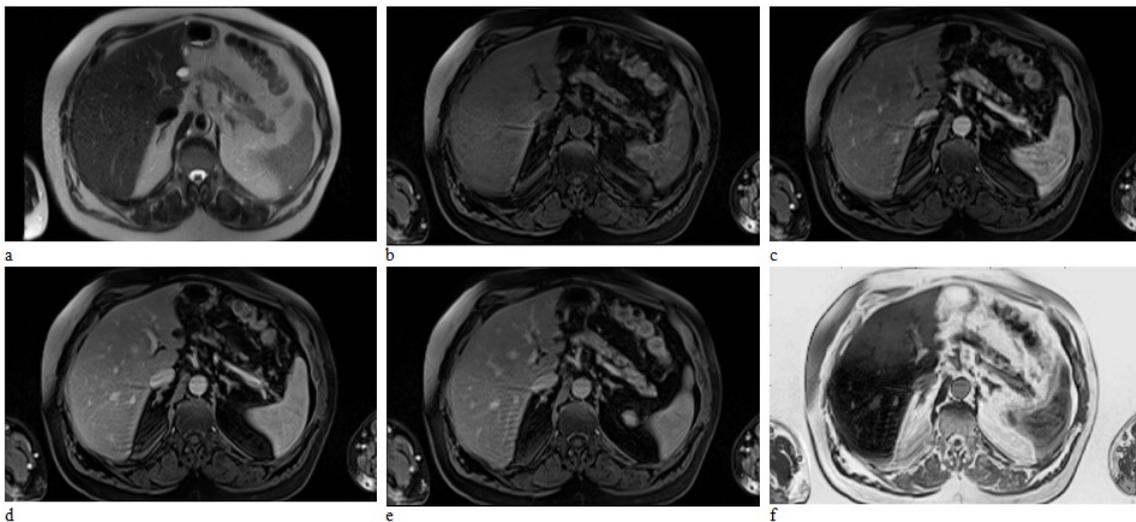


Figura 5.2. Imagen de distancias generada a partir de las imágenes de las secuencias separadas. a) Secuencia T2 BLADE; b) Secuencia VIBE: en vacío axial; c) Secuencia VIBE: axial arterial; d) Secuencia VIBE: axial venosa; e) Secuencia VIBE: axial tardío; f) Imagen de distancias generada a partir de la entrada multisección

Después de haber obtenido el mapa de distancias por el método propuesto, la visualización de los resultados proporciona una idea de cómo de buena es la nueva imagen, constituyendo una entrada con más contraste para el algoritmo de segmentación si se compara con la imagen de entrada inicial.

La imagen que representa el mapa de distancias con respecto al modelo de hígado generado previamente muestra la región del hígado más oscura. Esto es razonable ya que los píxeles que son hígado presentarán una distancia menor al modelo de hígado y una distancia menor se traduce en un nivel de gris bajo. Por el contrario, píxeles que representan otros tejidos tendrán un nivel de gris más elevado al representar mayor distancia al modelo. De esta manera, aumenta el contraste del hígado sobre los órganos y tejidos adyacentes en las

secuencias VIBE, en general, oscuras y de bajo contraste. Este aumento de contraste en la imagen que constituye la entrada al método de segmentación redundante en unos resultados más precisos, como se mostrará en el capítulo 6.

También se puede observar que entre el mapa de distancias obtenido a partir de una sola secuencia de la figura 5.1 y el mapa de distancias generado a partir de las múltiples secuencias de la figura 5.2 existen algunas diferencias. El mapa proveniente de la multisequencia está más suavizado, es más uniforme y homogéneo, lo que en principio permitiría realizar una segmentación más precisa.

Finalmente, asumiendo que el hígado se representa por un modelo gaussiano, la imagen de distancias es equivalente a un valor de probabilidad de pertenencia a la clase hígado a través de la siguiente expresión:

$$Prob(i, j, k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(d(i,j,k) - \mu)^2}{\sigma^2}} \quad (5.5)$$

Donde μ es la media, σ^2 es la varianza y σ es la desviación típica.

A modo de resumen, se puede decir que existe una gran variabilidad en las imágenes de MRI, debido a las máquinas de resonancia, la respuesta diferente de los tejidos de los distintos individuos o los protocolos de adquisición de las secuencias, entre otros, que conforman una escena no uniforme en cada corte. Para minimizar este efecto de iluminación irregular de la escena, tradicionalmente, se ha trabajado mucho en el desarrollo de filtros anisotrópicos potentes o técnicas de mejora de calidad en las imágenes. En el método presentado en esta tesis, se propone un descriptor basado en modelo que permite la generación de un mapa de probabilidad (a partir del mapa de distancias) de pertenencia al hígado.

Además, este descriptor puede aplicarse como una alternativa a los filtros anisotrópicos, si bien la asunción del modelo Gaussiano multivariable para la representación del hígado en la escena hay que explicarla. Se ha trabajado con la hipótesis de que el uso de varias secuencias y la información 3D reduce esta anisotropía en la escena. La imagen de mapa de distancias, resultado de la combinación de las múltiples secuencias, es más robusta al ruido y a los artefactos de la iluminación que la entrada inicial, y por tanto, la distribución de intensidad de píxel del hígado puede ser aproximada a una distribución gaussiana hasta un cierto grado. Con este descriptor propuesto y este enfoque se consigue una imagen con más contraste y más uniformidad que cada secuencia de entrada inicial. Esta imagen resultante de mapa de distancias constituirá la nueva entrada mejorada al método de segmentación.

Este descriptor compacto propuesto permite incorporar de manera consistente la información multisequencia en el marco de la segmentación 2D y 3D y tratarla como un problema monosecuencial. La dimensión del problema no aumenta incluso si se incorporan múltiples canales de información, lo cual redundante en un menor tiempo de cómputo del

algoritmo de segmentación. En el capítulo 6 se mostrará el aporte de esta representación compacta en la obtención de resultados más precisos.

Como se ha indicado anteriormente, el descriptor compacto propuesto adopta una doble funcionalidad, de descriptor en sí (combinando la información contenida en las múltiples secuencias) y como método de pre-procesamiento para mejorar la calidad de la entrada.

El modelo de hígado ha sido previamente obtenido con 3 estudios disponibles. Sobre la región de hígado anotada manualmente se han estimado los valores estadísticos de primer y segundo orden, media, desviación estándar y varianza para cada una de las secuencias. El modelo de hígado es por tanto dependiente de las secuencias adquiridas en el estudio. Si el número y la tipología de secuencias cambiaran se podría aplicar la misma metodología. Deberían estimarse los valores del modelo para el conjunto de secuencias disponibles en cada caso. Este modelo de hígado será utilizado para la generación del descriptor compacto.

5.3. Modelo de contorno activo

Los métodos de segmentación desarrollados e implementados en esta tesis consisten en la minimización de un contorno activo mediante técnicas variacionales en 2D y 3D. Se describe en este apartado el modelo de contorno activo.

El contorno activo inicial utilizado como base en los métodos propuestos en esta tesis se adapta de la siguiente referencia (Bresson, 2007). En este trabajo, se propone un nuevo modelo de contorno activo a partir de la combinación de dos técnicas ampliamente aceptadas: los Contornos Activos Sin Bordes (*Active Contours Without Edges* o ACWE) que se basa en la uniformidad de regiones para hacer evolucionar el contorno (Chan and Vese, 2001), y los Contornos Geodésicos Activos (*Geodesic Active Contours* o GAC), cuya evolución depende de los bordes de la imagen (Caselles et al., 1997). Se pretende aunar las ventajas de ambas técnicas en un nuevo modelo de contorno activo, en principio, más robusto acorde a los resultados presentados. Es decir, se combinan los enfoques complementarios de las dos técnicas de segmentación en un modelo mejorado capaz de detectar al mismo tiempo bordes a partir de los gradientes existentes en la imagen y regiones con niveles de intensidad homogénea.

En el caso concreto de segmentación de hígado, es de aplicación este modelo de contorno activo, dado que se pretende segmentar entre hígado o no hígado. Es decir, se puede asumir que existen dos regiones uniformes en la imagen, R_{in} y R_{out} con valores representativos medios c_1 y c_2 y delimitados por un contorno u . Los valores c_1 y c_2 no son constantes en todo el proceso sino que cambian a medida que el contorno evoluciona y se delimitan nuevas regiones R_{in} y R_{out} , siendo dependientes de la región delimitada en cada ciclo, $c_1(R_{in})$ y $c_2(R_{out})$.

El modelo de contorno activo implementado está diseñado para converger hacia los límites del hígado gracias al término gradiente en la energía, y a la maximización de la diferencia entre los valores estadísticos de distribución de intensidad en las regiones interior y exterior

(hígado / no hígado). La zona de “no hígado” no es uniforme en absoluto. Sin embargo, se ha realizado esta asunción dado que la mejora de contraste existente en el nuevo mapa de distancias mediante la obtención del descriptor compacto, la capacidad de la aproximación dual para la conveniente minimización del contorno activo en la consecución del mínimo global con independencia de la inicialización, junto con el hecho de que la zona interior del hígado sí se puede considerar relativamente homogénea, permiten que el contorno activo converja de manera óptima.

Con esta base, y tal y como expone Bresson en su trabajo (Bresson and Chan, 2007) el funcional de energía en el espacio continuo a minimizar se expresa de esta manera:

$$E_{ACWE}(\Omega_C, c_1, c_2, \lambda) = Per(\Omega_C) + \lambda \int_{\Omega_C} (c_1 - I(x))^2 dx + \lambda \int_{\Omega \setminus \Omega_C} (c_2 - I(x))^2 dx \quad (5.6)$$

Siendo I la imagen dada, Ω_C un subconjunto del dominio de la imagen delimitado por contorno cerrado, $Per(\Omega_C)$ el perímetro del conjunto Ω_C , λ es un parámetro positivo que regula el equilibrio entre el proceso de regularización y la fidelidad de la solución con respecto a la imagen original I , siendo c_1 y $c_2 \in R$.

Y el funcional de energía del modelo GAC es como sigue:

$$E_{GAC}(C) = \int_C g ds \quad (5.7)$$

Siendo g una función detectora de bordes, y s una función definida en el contorno C que separa el subconjunto de la imagen Ω_C que es el límite entre las dos regiones suaves, es decir $C = \partial\Omega_C = \partial(\Omega \setminus \Omega_C)$

El funcional de energía del contorno activo empleado en el método de segmentación propuesto en esta tesis viene representado por ambas energías conjuntamente representadas en las ecuaciones (5.6) y (5.7). Este modelo es el que se tomará como base para ser extendida a tres dimensiones y constituir una superficie activa.

5.4. Minimización del contorno activo: aproximación dual al problema variacional para minimización de funcionales de energía en 2D

En la sección 5.3 se ha expuesto el modelo de contorno activo o funcional de energía. En un proceso de segmentación es relevante que el método de minimización seleccionado sea adecuado para que el funcional de energía converja a una solución global y no se quede estancado en mínimos locales, que no proporcionan una segmentación óptima.

Los métodos de minimización son un campo amplio de trabajo. Explicado de una manera muy simple, el máximo o mínimo de una función se identifica a partir de resolver la ecuación de la primera derivada; y el hecho de averiguar si se trata de un máximo o un mínimo se concluye de la resolución de la ecuación de la segunda derivada. El problema es

que la ecuación a resolver puede ser realmente complicada, y en su resolución pueden aparecer varios mínimos, sobre los que hay que determinar si son locales o globales.

Aunque las imágenes son un conjunto discreto de valores, pueden ser interpretadas en un espacio continuo como una función relativamente suave. Como se ha indicado anteriormente, la energía que representa un contorno activo es lo que se denomina funcional de energía. Este funcional de energía es el que hay que minimizar para encontrar el mínimo global, el cual aportará la solución óptima, en este caso, la segmentación óptima de la imagen. La localización del mínimo global puede ser un proceso complejo, y no siempre se consigue. El conjunto de técnicas variacionales persiguen la obtención precisa del mínimo global y no local.

La resolución de los funcionales de energía mediante técnicas variacionales para la localización de los mínimos no es tan sencilla. Las indeterminaciones, los espacios n -dimensionales obligan a trabajar con EDP, Ecuaciones en Derivadas Parciales, que están resueltas, pero que no siempre proporcionan una solución dependiendo de la función definida. De hecho, matemáticamente hablando, una función posee un mínimo global si es estrictamente convexa. Si sólo es convexa tendrá mínimos locales además de un mínimo global, y esto no es deseable, ya que la función puede quedar estancada en un mínimo local en el proceso de minimización.

Con el objeto de localizar el mínimo en una función convexa y no estrictamente convexa, y transformar el funcional de energía en una función estrictamente convexa, y por ende, con un mínimo global y solución óptima, Chambolle propuso un nuevo enfoque matemático. Dicha propuesta se conoció inicialmente como *Mean Curvature Motion* para interfaces y su propuesta matemática se basa en la denominada aproximación dual (A Chambolle, 2004). Si la función u a minimizar no es estrictamente convexa y por tanto, tiene mínimos locales además del global, para calcular el mínimo global, propone transformar esta función estrictamente no convexa en una estrictamente convexa, añadiendo otra función v , de manera que la localización del mínimo se realiza fijando v y minimizando con respecto a u ; fijando u y minimizando con respecto a v , por separado, y se itera hasta la convergencia.

El funcional de energía empleado para caracterizar el contorno activo propuesto se representa por la relación ligada entre las dos energías dadas en ecuaciones (5.8) y (5.9). Este funcional es el siguiente:

$$E_1(u, c_1, c_2, \lambda) = TV_g(u) + \lambda \int r(x, c_1, c_2)u \, dx \quad (5.8)$$

Donde $TV_g(u)$ representa la variación total con peso de una función u con un peso g , que contiene información acerca de los bordes de la imagen I_0 a través de sus gradientes, y que se obtiene mediante esta expresión:

$$g(|\nabla I_0|) = \frac{1}{1 + \beta |\nabla I_0|^2} \quad (5.9)$$

Siendo β una constante positiva de la función g para detección de bordes, y $g = (g_x, g_y)$

La ecuación (5.8) proporciona la relación entre ACWE y GAC, siendo g expresado en la ecuación (5.9) la función detección de bordes y u una función característica, tal que 1_Ω , aparece la siguiente relación:

$$TV_g(u = 1_\Omega) = \int_{\Omega} g(x) |\nabla 1_{\Omega_c}| dx = \int_c g(s) ds = E_{GAC}(C) \quad (5.10)$$

De manera que la ecuación (5.8) puede ser reescrita de esta manera:

$$E_2(u = 1_{\Omega_c}, c_1, c_2, \lambda) = TV_g(1_{\Omega_c}) + \lambda \int r(x, c_1, c_2) 1_{\Omega_c} dx \quad (5.11)$$

En la ecuación (5.11) r se define como $r(x, c_1, c_2) = (c_1 - f(x))^2 - (c_2 - f(x))^2$, y en nuestra propuesta la entrada $f(x)$ es el mapa de distancias obtenido a partir del cálculo del descriptor compacto para el conjunto de N secuencias de entrada disponibles.

Esto significa que cualquier minimizante u de E_2 obtenido de un método de minimización cualquiera, el conjunto de puntos de la función u tal que u sea constante positiva, define un conjunto de puntos cuyo contorno es un mínimo global. En su trabajo muestra la prueba de la convergencia del funcional de energía en una solución total (A Chambolle, 2004).

En definitiva, la ecuación a resolver para el problema de minimización y por ende, de segmentación en este caso, es:

$$\min_{0 \leq u \leq 1} \left\{ E_2(u, c_1, c_2, \lambda) = TV_g(u) + \lambda \int r(x, c_1, c_2) u dx \right\} \quad (5.12)$$

La cual tiene el mismo conjunto de mínimos que la siguiente función convexa

$$\min_u \left\{ TV_g(u) + \lambda \int r(x, c_1, c_2) u + \alpha v(u) dx \right\} \quad (5.13)$$

Con $v(\xi) = \max \{0, 2 \left| \xi - \frac{1}{2} \right| - 1\}$ es una función de penalización siempre que α sea suficientemente grande en comparación con λ .

La ecuación (5.11) no es estrictamente convexa, lo que significa que posee mínimos locales que no son globales. Este problema de minimización presentado implica la resolución de una función no convexa dado que la minimización debe llevarse a cabo en una función que tiene sólo los valores c_1 y c_2 , que es una colección no convexa. Por tanto, el proceso de minimización puede llevar a un mínimo local y no a uno global. La formulación dual presentada garantiza que la ecuación (5.11) se pueda reescribir como una estrictamente convexa, de manera que sí presente un mínimo global, y por tanto, la solución global a nuestro problema. Esta función v es la penalización para transformar la función convexa en una función estrictamente convexa.

$$E^r(u, v, c_1, c_2, \lambda, \alpha, \theta) = TV_g(u) + \frac{1}{2\theta} \|u - v\|_{L^2}^2 + \int_{\Omega} \lambda r(x, c_1, c_2)v + \alpha v(v)dx \quad (5.14)$$

El problema de segmentación se ha convertido en el siguiente problema de minimización:

$$\min_{u,v} \left\{ E^r(u, v, c_1, c_2, \lambda, \alpha, \theta) = TV_g(u) + \frac{1}{2\theta} \|u - v\|_{L^2}^2 + \int_{\Omega} \lambda r(x, c_1, c_2)v + \alpha v(v)dx \right\} \quad (5.15)$$

se basa en la variación total con peso introducida de una función u con un peso g , que contiene información acerca de los bordes de la imagen I_0 a través de sus gradientes, y que se obtiene mediante la expresión (5.10). Dado que se ha demostrado que E^r es estrictamente convexa (Bresson et al., 2007), su mínimo puede ser obtenido minimizando u y v por separado, y luego iterando hasta que converjan. Con lo cual hay que considerar los siguientes problemas de minimización:

1) v fijado, se minimiza u como solución de:

$$\min_u \left\{ TV_g(u) + \frac{1}{2\theta} \|u - v\|_{L^2}^2 \right\} \quad (5.16)$$

2) u fijado, se busca un mínimo de v

$$\min_v \left\{ \frac{1}{2\theta} \|u - v\|_{L^2}^2 + \lambda \int_{\Omega} \lambda r(x, c_1, c_2)v + \alpha v(v)dx \right\} \quad (5.17)$$

La solución a (5.16) es

$$u = v - \theta \operatorname{div} p \quad (5.18)$$

Donde $p = (p_x, p_y)$ dado por

$$g(x)\nabla(\theta \operatorname{div} p - v) - |\nabla(\theta \operatorname{div} p - v)|p = 0 \quad (5.19)$$

La ecuación anterior puede solucionarse de una manera sencilla computacionalmente hablando por un método de punto fijo, de manera que el problema en el espacio continuo se traslada al espacio discreto. Esto es,

$$p^0 = 0$$

$$p^{n+1} = \frac{p^n + \delta t \nabla(\operatorname{div}(p^n) - \frac{v}{\theta})}{1 + \frac{\delta t}{g(x)} |\nabla(\operatorname{div}(p^n) - v/\theta)|} \quad (5.20)$$

La solución a (5.17) para localizar el mínimo de v es dada por:

$$v = \min\{\max\{u(x) - \theta\lambda r(x, c_1, c_2), 0\}, 1\} \quad (5.21)$$

Para minimizar la energía hay que iterar (5.20) y (5.21). Las constantes c_1 y c_2 se actualizan periódicamente cada n iteraciones definidas. En este proceso de minimización existen los siguientes parámetros configurables:

λ , regula cómo de suave se permite que sean los contornos, de manera que a menor λ , más suaves son los contornos (por tanto, se adapta peor a posibles irregularidades), y a mayor λ se permite más irregularidad en los bordes. Es necesaria la búsqueda de un compromiso acorde al objeto a segmentar,

β , indica la importancia del gradiente en la localización de los bordes,

N , indica el número de iteraciones a realizar,

θ , es una constante positiva escogida pequeña.

5.5. Extensión a 3D del modelo de contorno activo (superficie activa) y su proceso de minimización del funcional de energía

Las imágenes abdominales obtenidas mediante resonancia magnética, aunque habitualmente se analizan en 2 dimensiones y se procesan uno a uno todos los cortes de cada secuencia, en realidad constituyen un volumen 3D, que es la representación real del cuerpo. En los métodos de segmentación la información de la tercera dimensión es relevante y aporta conocimiento sobre la continuidad de las regiones segmentadas, ya que los volúmenes deben presentar continuidad en todas las direcciones. Por ello, se propone en esta tesis el uso de la tercera dimensión, lo cual contribuirá a que la segmentación sea más robusta. Esta extensión a una nueva dimensión es plausible que implique un aumento en el tiempo de cómputo necesario para minimizar la función de energía.

En una segmentación 3D, el modelo de contorno activo debe ser extendido a esta tercera dimensión, lo que se denomina una superficie activa. El funcional de energía que representa la superficie activa en el volumen de datos de entrada debe ser minimizado. El método de minimización elegido, en este caso, la aproximación dual al problema variacional debe ser también extendido para su aplicación en un espacio de una dimensión superior.

Este modelo de superficie activa y el método de minimización de su funcional de energía se aplica para segmentación hepática sobre imagen de resonancia magnética. La extensión a una nueva dimensión no es difícil desde un punto de vista teórico. Para ello, se han considerado las siguientes premisas:

- Las funciones características u , v se mantienen, pero ahora u , v , los contornos activos minimizantes del funcional de energía, son en tres dimensiones en vez de en dos, y por ello, se denominan superficies activas. Para estas superficies activas,

sigue vigente el concepto $c_1, c_2 \in R$, que son los valores de intensidad de la región interior de la superficie activa y el valor de intensidad de la región exterior, respectivamente, pero estimados en un volumen en vez de en una área.

- La entrada al procesamiento, anteriormente una imagen 2D, se convierte ahora en un volumen 3D.
- Se emplea para resolver la ecuación en derivadas parciales el mismo método del punto fijo, expuesto en la ecuación (5.20) pero dotándolo de una dimensión más, y siendo cada punto $p = (p_x, p_y, p_z)$
- El cómputo de la función de distribución de probabilidad para la estimación de c_1 y c_2 se actualiza dentro y fuera de la superficie activa, al igual que antes sucedía dentro y fuera del contorno activo en 2D, siendo el cálculo de c_1 y c_2 iterativo.
- La función de bordes g se extiende de 2D a 3D de una manera sencilla mediante el cálculo del gradiente en los 3 ejes, esto es, $g = (g_x, g_y, g_z)$

Aunque la minimización de la superficie activa se realice en este caso sobre el volumen de información de la secuencia, para poder evaluar los resultados y compararlos con los obtenidos con el método en 2D, se visualizará el volumen en cortes 2D superponiendo a cada corte el conjunto de puntos delimitado por la función u que minimiza. Sobre estas imágenes 2D se aplicarán las métricas de medida identificadas, cuyos resultados se muestran en detalle en el capítulo 6.

Los parámetros configurables en el proceso de minimización que se han mencionado en método en 2D también son de aplicación aquí:

λ , regula cómo de suave se permiten que sean los contornos, ya que a menor λ , más suaves son los contornos (se adapta peor a posibles irregularidades), y a mayor λ se permite más irregularidad en los bordes. Es necesaria la búsqueda de un compromiso dependiendo del elemento a segmentar en el volumen de entrada.

β , indica la importancia del gradiente en la localización de los bordes.

N , representa el número de iteraciones del algoritmo para alcanzar el mínimo global.

5.6. Algoritmo completo de segmentación

En los apartados previos se ha realizado la aproximación teórica al núcleo del método propuesto, constituido por el uso conjunto de:

- 1) nuevo descriptor compacto basado en modelo para la generación de un mapa de distancias que mejora la calidad en la imagen de entrada y maneja la entrada multisequencia;

- 2) la definición del modelo del contorno activo (o superficie activa) que se va a emplear en el proceso de segmentación;
- 3) el método de minimización propuesto del funcional de energía (modelo variacional) representado por el contorno activo (o superficie activa) definidos;

En este apartado se describe el algoritmo completo de segmentación, representado en la siguiente imagen con este flujo de procesamiento.

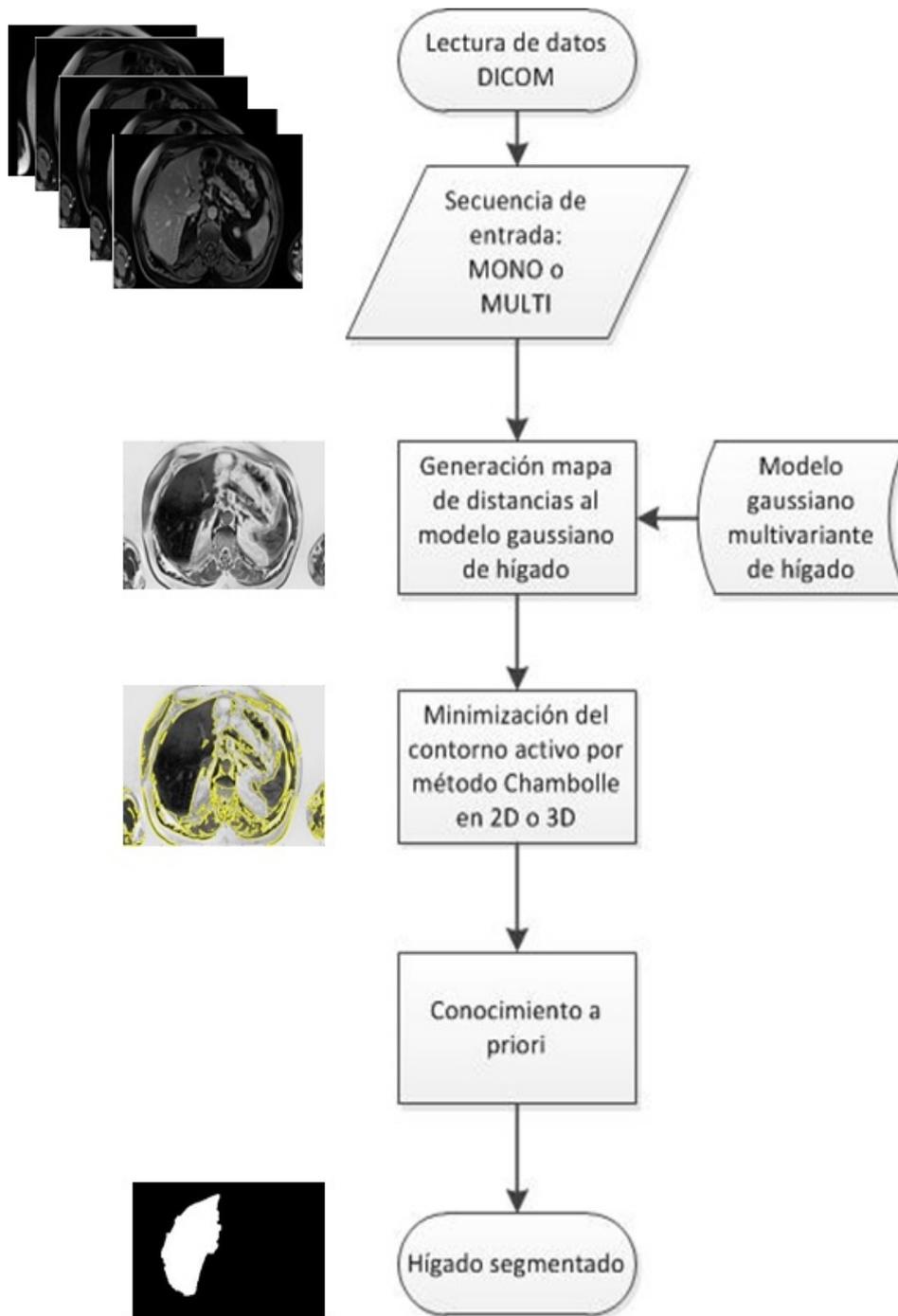


Figura 5.3. Diagrama de flujo del proceso completo de segmentación

Se describen a continuación cada uno de los pasos del proceso.

5.6.1. Entrada monosecuencia o multisecuencia

Como se ha explicado en el apartado 5.1 se dispone de dos tipos de entrada: monosecuencia o multisecuencia, con 5 secuencias diferentes simultáneas. Cada una de estas secuencias se corresponde con la respuesta que los tejidos del cuerpo producen cuando son sometidos a diferentes campos electromagnéticos. El uso combinado y adecuado de la información múltiple se considera más valioso que la entrada única, y es uno de los elementos que se validan en esta tesis. La manera de emplear dicha información multisecuencia puede ser diversa, acorde al estado del arte expuesto en el capítulo 4. En esta tesis se propone un nuevo descriptor compacto que reduce la dimensión de las 5 secuencias a 1. Este descriptor permite además la generación de un mapa de distancias, que constituirá la nueva entrada, y que proporciona información de más valor que cada entrada por separado.

Se validará el algoritmo sobre ambas entradas: monosecuencia (secuencia VIBE en vacío axial) y multisecuencia (secuencia T2 BLADE y las 4 secuencias VIBE), tal y como se ha detallado en el apartado 5.1.

5.6.2. Generación de un modelo gaussiano multivariante

El modelo estadístico multivariable de hígado mencionado se realiza en base a 3 estudios disponibles. Se ha escogido el tamaño de la región de hígado que se emplea para generar el modelo gaussiano multivariable, que es de 10 x 50 píxeles en el caso monosecuencia, y 10 x 50 x 5 píxeles en el caso multisecuencia. Se calcula este valor en los 10 cortes comprendidos entre el corte 30 y el 40, que tienen porciones de hígado suficientemente grandes para tomar un muestrario relevante. Se ha elegido un tamaño en el eje 'y' mayor que en eje 'x' dado que la no uniformidad en la iluminación de la imagen se produce en este eje, y es por tanto necesario considerar una superficie mayor para que el modelo generado sea capaz de absorber esta variabilidad. Para la generación de este modelo de hígado se han calculado los valores de media y desviación estándar del conjunto de píxeles indicado.

5.6.3. Descriptor compacto: Generación de un mapa de distancias basado en modelo a partir de la entrada

La imagen o imágenes de entrada se van a reemplazar por un mapa de distancias donde se representa la distancia de cada píxel a un modelo probabilístico multivariable de hígado que ha sido generado previamente. Este modelo pretende absorber la variabilidad existente en los niveles de gris entre individuos y diferentes capturas, y por tanto, se persigue que este mapa de distancias generado reduzca los problemas de iluminación no uniforme y aumente el contraste y la nitidez de los bordes entre los distintos tejidos.

Este mapa de distancias generado basado en modelo constituye la imagen de entrada al algoritmo de segmentación sobre la cual el contorno activo (o superficie activa en 3D) será minimizado para encontrar la solución óptima.

El método propuesto en esta tesis puede ser aplicado a otros protocolos de adquisición con diferentes máquinas de resonancia magnética y con mayor o menor número de secuencias. El comportamiento del hígado en las diferentes secuencias elegidas es almacenado en el modelo estadístico multivariable de hígado generado inicialmente. El descriptor compacto propuesto permite el manejo y combinación de la información de todas las secuencias disponibles, sin aumentar la dimensión de la entrada, y permitiendo la generación del mapa de distancias de hígado con respecto al modelo generado del mismo como única entrada mejorada al método de segmentación en si mismo.

5.6.4. Minimización del contorno activo definido mediante los métodos descritos en 2 y 3 dimensiones

Una vez se dispone del mapa de distancias, bien procedente de la entrada monosecuencia o multisequencia, se aplica el algoritmo de segmentación propiamente dicho. Como se ha descrito en las secciones anteriores, se va a minimizar un modelo de contorno activo que aúna los Contornos Activos Sin Bordes o *Active Contour Without Edges* de Chan y Vese, y los Contornos Geodésicos Activos o *Geodesic Active Contours* de Caselles, en el intento de mejorar la segmentación a través de la sinergia entre las dos técnicas, de manera que se busca la localización de regiones homogéneas en valor de intensidad (sin detección de bordes en sí), y la detección de bordes mediante gradiente y regulación del movimiento de la curva con parámetros geométricos. Para obtener la segmentación óptima, esto es, el mínimo global resultado de la minimización del funcional de energía que el contorno activo representa en la imagen, se va a aplicar y adecuar la aproximación dual al problema variacional propuesto por Chambolle, descrita anteriormente en la ecuación (5.15).

Tanto el modelo de contorno activo como esta aproximación dual al problema variacional han sido extendidos a 3D como una de las aportaciones de esta tesis.

Ambos modelos de contorno activo, tanto en 2D como en 3D (superficie activa) son aplicados sobre la imagen del mapa de distancias generado a partir de la secuencia o secuencias de entrada y basado en el modelo probabilístico de hígado previo.

5.6.5. Etapa de post-procesado

La dificultad de las imágenes de MRI, como se ha mostrado en los capítulos 2, 3 y 4 es que presentan poco contraste entre los diferentes elementos presentes, siendo los bordes poco nítidos. Esto implica que la segmentación precisa es difícil. En la imagen abdominal, junto con el hígado, se encuentran los riñones, la grasa abdominal, los pulmones, los intestinos... con características visuales similares a las del hígado. Por lo tanto, a pesar de que los métodos de segmentación complejos, como es el caso de esta combinación de contorno activo robusto con una minimización muy eficiente, obtengan buenos resultados, quedan en la imagen un montón de partículas residuales que no son hígado, cuyo contorno (puntos de fuerte gradiente al ser los bordes) también han sido delimitados por la función

característica. Es necesario, tras el proceso de minimización, la aplicación de una etapa de post-procesado.

La etapa de post-procesado implementada ad-hoc consiste en la aplicación de conocimiento a priori y filtrado de la imagen para este refinamiento de la segmentación inicial, de manera que, de forma automática, la región del hígado queda seleccionada y bien delimitada.

En el procesamiento de imágenes médicas y extracción de información posterior existen muchos estudios que han empleado el conocimiento a priori como parte importante en el desarrollo, como se ha mostrado en el estado del arte. De hecho, los métodos de Modelos Activos de Forma (*Active Shape Model* o ASM) y Modelos Activos de Apariencia (*Active Appearance Model* o AAM) descritos en los apartados 4.3.9.3 y 4.3.9.4 del estado del arte se basan en conocimiento a priori de forma o información de textura, recogidos en un modelo.

En la segmentación hepática sobre imagen de resonancia magnética se ha considerado que también se puede hacer uso del conocimiento existente. Cuando una persona se somete a una prueba de resonancia se introduce en el escáner correspondiente, lo cual le obliga a estar en una posición determinada y sin apenas posibilidad de moverse. Es por ello, que a pesar de los movimientos respiratorios y los pequeños desplazamientos que provocan los problemas de registro, y de la variabilidad existente entre los tamaños de los hígados en cada paciente, existe un rango de posiciones donde siempre se localiza el hígado. El hígado, si bien se encuentra en la zona derecha del cuerpo humano, en las imágenes reconstruidas por el escáner Avanto de Siemens se ubica siempre en la zona izquierda de la imagen. Aunque el área y forma de las secciones del hígado en 2D sea diferente a lo largo de todos los cortes de la secuencia, existe una región común en todas las imágenes de los cortes que conforma un eje que siempre atraviesa el hígado. Es decir, las regiones o partículas resultantes del proceso de segmentación inicial que no se encuentran en la zona izquierda son desechadas, ya que no se corresponderán con el hígado. Además el área del hígado es la mayor de todas las partículas segmentadas ubicadas en la zona izquierda de la imagen. Este hecho también sucede con ambos extremos del hígado. Con estas dos premisas, se realiza un filtrado del resto de partículas.

Por otra parte, las técnicas variacionales, como se ha explicado en el estado del arte, pueden localizar regiones interiores, y dado que el objetivo de esta tesis es la segmentación del hígado, posibles segmentaciones interiores, como algún vaso o quiste, deben ser asumidas como pertenecientes al volumen del hígado y por tanto, se realiza una operación de rellenado de agujeros.

En la figura 5.4 se muestran los efectos de la fase de post-procesado.

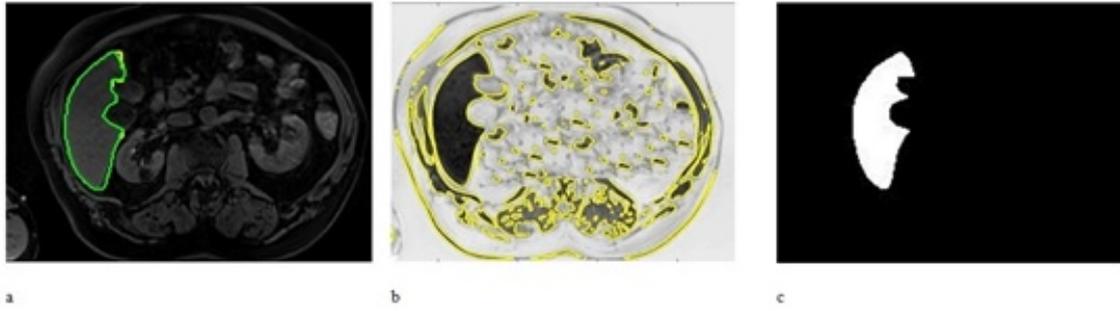


Figura 5.4. Ejemplo de la contribución de la etapa de post-procesado a la segmentación automática final visualizado en un corte 2D de la secuencia VIBE. a) El hígado es segmentado con el método propuesto de superficie activa 3D sobre la secuencia VIBE vacío axial (contorno verde es el resultado de la segmentación automática y amarillo la segmentación de referencia); b) salida del proceso de minimización sobre el mapa de distancias para entrada multisequencia mediante descriptor compacto con impurezas en la imagen; c) hígado segmentado después de la etapa de post-procesado, la región en la izquierda de la imagen y de mayor tamaño permanece.

Los resultados obtenidos de las pruebas de validación de este algoritmo de segmentación, para dos y tres dimensiones, y con entrada monosequencia y multisequencia se detallan en el capítulo 6.

5.6.6. Ajuste de parámetros

Los parámetros establecidos para el proceso de minimización del algoritmo se obtuvieron con 60 iteraciones, actualizando c_1 y c_2 cada 10 iteraciones, $\lambda = 0.01$, $\beta = 50/255^4$, $\delta t = 1/8$.

En las aproximaciones con contornos activos la inicialización es de vital importancia, y tradicionalmente, es recomendable que el contorno inicializado se encuentre próximo a la solución final para que converja adecuadamente. En el método presentado en esta tesis esto no es necesario. El principal objetivo de la minimización del funcional de energía que representa el contorno o superficie activa mediante técnicas variacionales es la obtención del mínimo global del funcional, y no mínimos locales. Con esta premisa, la inicialización no constituye ningún problema, ya que el algoritmo consigue llegar al mínimo global, y por tanto, a la segmentación óptima del objeto.

El método es independiente de las condiciones de inicialización. En este caso, se ha escogido $u = v = \frac{I_0}{\max(I_0)}$, $c_1 = c_2 = p_x = p_y = p_z = 0$, siendo I_0 la imagen inicial.

Capítulo 6: PRUEBAS Y
RESULTADOS

En el capítulo anterior se ha descrito la aproximación teórica de los desarrollos propuestos en esta tesis. En este capítulo se define una metodología de pruebas para la correcta valoración de los algoritmos desarrollados y su comparación con otras técnicas. Asimismo, se exponen los resultados obtenidos para cada uno de los métodos de segmentación propuestos. Finalmente se muestra una comparativa de los resultados obtenidos para cada uno de los métodos junto con los de otros trabajos del estado del arte, lo cual ayudará a la extracción de las conclusiones del capítulo siguiente.

6.1. Metodología de las pruebas

El objetivo del presente capítulo es recoger de manera organizada y rigurosa las pruebas de evaluación de los métodos propuestos en esta tesis para la segmentación de hígado en MRI. Los métodos valorados son:

- Segmentación en 2D con entrada monosecuencia
- Segmentación en 2D con entrada multisecuencia
- Segmentación en 3D con entrada monosecuencia
- Segmentación en 3D con entrada multisecuencia

Siendo el método de segmentación (en 2D o 3D) el descrito en el apartado 5.6.

Para valorar la bondad de dichos métodos, se cuantificará la segmentación del hígado sobre las imágenes y se comparará con la segmentación de referencia (denominada *ground truth* o *gold standard*) mediante las métricas de evaluación. Se ha definido en el capítulo 5 qué se entiende con el término monosecuencia y multisecuencia. Cada adquisición realizada por MRI para un paciente se denomina un estudio, el cual comprende el conjunto de secuencias disponible. Cada una de dichas secuencias cuenta con 63 imágenes, que se corresponden con el número de cortes 2D de dicho volumen o secuencia a lo largo del eje axial del cuerpo humano.

Una vez establecida la referencia de segmentación y las métricas de evaluación se procederá a validar los desarrollos. Para ello, se aplicará cada uno de los diferentes métodos de segmentación propuestos sobre el estudio de entrada y los resultados de dicha segmentación serán un conjunto de 63 imágenes correspondientes a cada uno de los cortes de la secuencia de máxima resolución. Estas imágenes, al igual que la segmentación de referencia, son binarias y, como resultado de este proceso, sus píxeles han sido etiquetados como 1 (hígado) o 0 (no hígado).

Con los volúmenes segmentados y de referencia, y tras la aplicación de las expresiones correspondientes, se obtienen los valores de las métricas para cada método de segmentación.

En este capítulo se muestran los resultados obtenidos. En los métodos propuestos se combinan 3 elementos importantes:

- 1) Descriptor compacto para generación del mapa de distancias
- 2) Modelo de contorno activo (2D y 3D)
- 3) Métodos variacionales como técnicas robustas de minimización (2D y 3D) para regularización y suavizado

Es importante valorar si la unión propuesta de estos elementos constituye un método de segmentación que proporcione resultados precisos. Se debe valorar también si la extensión a 3D aporta valor sobre el desarrollo 2D. Otro elemento a valorar es si el descriptor compacto propuesto para la generación del mapa de distancias, que constituye la nueva entrada generada a partir de las entrada multisequencia, es válido para el manejo de la información múltiple, y permite la inclusión de información de valor complementaria que contribuya a mejores resultados en la segmentación final.

6.1.1. Conjuntos de datos. Monosequencia y multisequencia

Las imágenes disponibles para el desarrollo y validación de los métodos han sido adquiridas mediante el escáner de resonancia magnética Avanto de Siemens de 1.5 T disponible en la CVSS Radiología Clínica (IMQ, Bilbao). Cada uno de dichos estudios contiene imágenes provenientes de 5 secuencias, las denominadas:

- Secuencia 00: T2 BLADE
- Secuencia VIBE 05: en vacío axial
- Secuencia VIBE 06: arterial axial
- Secuencia VIBE 07: venoso axial
- Secuencia VIBE 08: tardío axial

Para generar el modelo estadístico multivariable del hígado, se han empleado 3 estudios. Además, se dispone de otros 6 estudios adicionales para la realización de las pruebas, todos ellos de hígados sanos. Cada caso de prueba contiene la información de las diferentes secuencias. Cada una de estas secuencias está constituida de 63 imágenes (cortes o *slices*) de tamaño 160 x 256 píxeles, que presentan una separación relativa en el eje axial de 2.5 mm. La profundidad de píxel es 12 bits.

El grosor de corte (*slice thickness*) es de 2.5 mm, y la resolución espacial en el plano axial (*in plane resolution*) es de 1.95 mm para las imágenes de 160 x 256 píxeles. Los parámetros configurados para la adquisición de las secuencias VIBE AX IPAT son tiempo de repetición (*Repetition Time*) de 6.02 milisegundos, tiempo de eco (*Echo Time*) es de 2.65 milisegundos, número de ecos (*Echo Numbers*) es 1. Para la secuencia T2 BLADE el tiempo de repetición (*RepetitionTime*) es 8167.75 milisegundos y tiempo de eco (*Echo Time*) es 106 milisegundos. Todos los temas legales y éticos, incluido el consentimiento del paciente para el uso de las imágenes con fines de investigación, son gestionados por el socio clínico indicado. Las imágenes están completamente anonimizadas y sin datos personales.

Se ilustra el conjunto en la siguiente imagen:

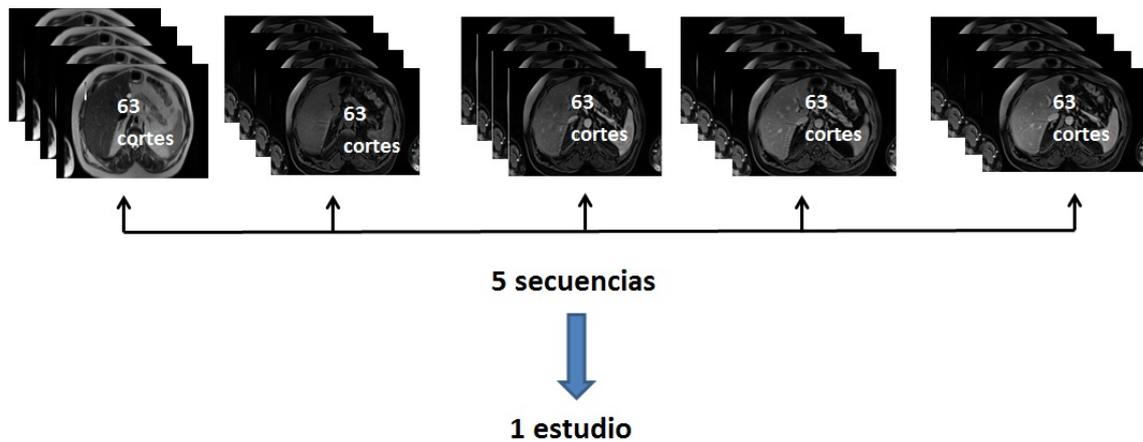


Figura 6.1. Cada estudio está compuesto de 5 secuencias. Cada secuencia se compone de 63 cortes.

Se dispone de 6 estudios (conjuntos de datos) para la validación. En total, el número de imágenes sobre las que se realizará la validación es de 1.890 imágenes (6 conjuntos de datos x 5 secuencias/conjunto x 63 imágenes /secuencia).

Las 5 secuencias juntas constituyen la entrada del análisis del comportamiento de la segmentación en imágenes multisequencia en cada uno de los métodos desarrollados y evaluados.

Para el análisis de los métodos que se ejecutan sobre una sola secuencia, se ha elegido el protocolo VIBE, acorde al criterio clínico de calidad en la imagen, y dentro de las 4 secuencias generadas con dicho protocolo, se ha seleccionado la VIBE en vacío axial para pruebas de los algoritmos sobre una única secuencia. El protocolo VIBE se corresponde con la secuencia ultra rápida gradiente - eco que Siemens proporciona.

6.1.2. Segmentación de referencia

La segmentación de referencia, también conocida por su término en inglés, *ground truth* o *gold estándar* se emplea para referirse a la “verdad” contra la cual se van a comparar los resultados de segmentación. En este caso, se cuenta con varias segmentaciones de referencia del hígado contra las que verificar los resultados de la segmentación. Las imágenes que constituyen las secuencias de ambos casos han sido segmentadas manualmente.

Se ha contado con la ayuda del doctor radiólogo Gonzalo Solís de la Clínica Vicente San Sebastián Radiología Clínica (IMQ) y su equipo para la identificación y segmentación del hígado en las imágenes de resonancia magnética. Esta tarea consiste en visualizar uno a uno todos los cortes de una secuencia, y seleccionar en la imagen el área hepática. Para ello, es

necesaria una herramienta software que pueda facilitar esta labor. Con este objetivo, se ha implementado un módulo software que permite cargar la imagen deseada, y mediante un conjunto de herramientas de zoom y polilínea permite seleccionar manualmente el hígado en cada una de las imágenes. Una vez el usuario está conforme con la selección, se almacena para su posterior uso en la fase de validación. De esta manera, la zona seleccionada como hígado (contenida dentro de la polilínea) será la que posea el valor '1', y lo que no es hígado, será etiquetado con valor '0'. Es decir, los hígados de referencia están constituidos por un conjunto de imágenes binarias, tantas como cortes haya en la secuencia inicial, con la que verificar los resultados de la segmentación, siendo los píxeles con valor de '1' aquellos en los que el hígado está presente. Para la evaluación de los resultados de segmentación se compararán los resultados de cada corte de la secuencia con el corte de referencia que ha sido segmentado manualmente.

Se muestra a continuación una imagen de la interfaz de usuario de la herramienta SW desarrollada ad-hoc para poder realizar de manera manual esta segmentación de referencia.

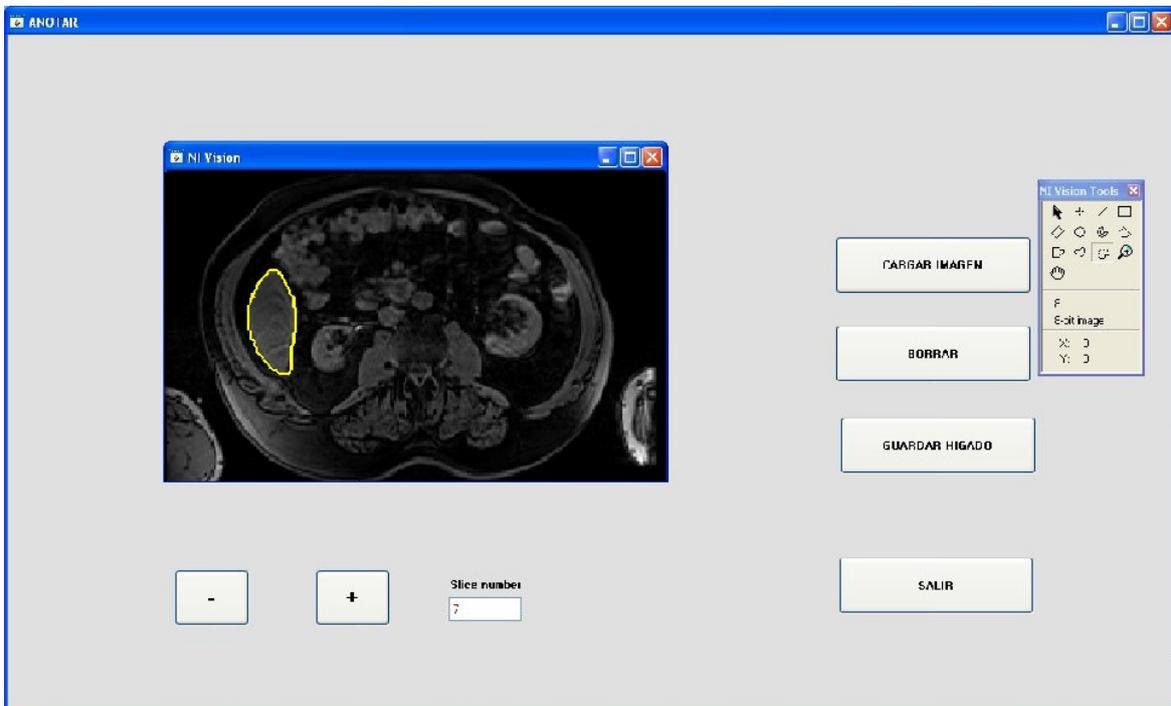


Figura 6.2. Ejemplo de segmentación manual de una imagen con el módulo SW desarrollado ad-hoc

La segmentación manual ha sido realizada 2 veces para poder evaluar la variabilidad entre las personas en la realización de la segmentación de referencia. Además, una de las dos personas ha realizado el proceso de segmentación manual sobre las imágenes dos veces. De esta manera se podrá comprobar la variabilidad en las segmentaciones de referencia entre ambas personas y para el mismo sujeto en distintos momentos de tiempo, lo que se conoce con su término en inglés *inter-rater variability* e *intra-rater variability*, respectivamente. Estos valores dan información de la complejidad del proceso de segmentación.

6.1.3. Métricas

La identificación y uso de unas métricas adecuadas para validar los métodos de segmentación es objeto de estudio en la comunidad científica. Después del análisis realizado, para la evaluación de los métodos propuestos en esta tesis se van a emplear un conjunto de métricas bien establecidas y que son empleadas para cuantificar la bondad de los resultados en los diversos trabajos referenciados. Estas métricas representan distancias, porcentaje de solape entre volúmenes, o medidas de precisión y similitud. No todos los trabajos calculan todas, de hecho, la mayor parte de las publicaciones muestran resultados parciales con una o dos métricas. Sin embargo, en este trabajo se ha decidido calcular las 9 métricas identificadas como más frecuentes, para que sea posible comparar los resultados obtenidos con varios de los trabajos más relevantes.

La mayor parte de las métricas se basan en solape volumétrico y distancias entre superficies. Existen variantes de dichos conceptos con modificaciones de las expresiones matemáticas para calcularlas, pero en general están clasificados en estos dos grandes grupos. Dependiendo del objeto que se quiera comparar se emplean unas u otras. Por ejemplo, si se desea valorar cómo de bien se realiza el cálculo de volumen de tumores, se emplearán las de solape volumétrico. En otras circunstancias, es más conveniente la evaluación de la calidad de la segmentación por otras vías. Se detalla el cálculo de las 9 métricas seleccionadas para la fase de validación. En el caso del MICCAI Grand Challenge 2007 se eligieron las cinco primeras (Heimann et al., 2009).

1. Error de Solapamiento Volumétrico o *Volumetric Overlap Error* (VOE): el error de solapamiento volumétrico entre 2 conjuntos de vóxeles se indica en porcentaje y se calcula:

$$VOE = 100 \left(1 - \frac{V_S \cap V_{GT}}{V_S \cup V_{GT}} \right) \quad (6.1)$$

con V_S el volumen segmentado y V_{GT} la referencia

Este coeficiente obtenido de dividir la intersección de volúmenes y la unión de los mismos se conoce como el coeficiente de Tanimoto o Jaccard. En una segmentación perfecta, este error es 0, y 100 si los volúmenes segmentados no coinciden nada. Esta es una de las métricas más populares para la evaluación de la precisión de la segmentación.

2. Diferencia relativa de volúmenes o *Relative Volume Difference* (RVD): la diferencia relativa de volúmenes entre 2 conjuntos de vóxeles se calcula con la siguiente expresión:

$$RVD = 100 \left(\frac{V_S - V_{GT}}{V_{GT}} \right) \quad (6.2)$$

con V_S el volumen segmentado y V_{GT} la referencia

Un valor de 0 significa que los dos volúmenes son idénticos, pero no significa que se superpongan o que sean el mismo. Por esta razón este valor nunca debe ser usado como única medida de la bondad de un método. En combinación con otros valores indica si hay

fugas o sub-segmentación, lo que se conoce como *oversegmentation* o *undersegmentation*. Por ello son valores con signo, aunque como medida del volumen relativo se emplea en valor absoluto.

3. Distancia media de superficie simétrica o *Average Symmetric Surface Distance (ASD)*: se proporciona en milímetros y se calcula en base a los vóxeles de superficie de 2 segmentaciones, la segmentada V_S y la de referencia V_{GT} . Los vóxeles de superficie se definen como aquellos que tienen al menos un vóxel no-objeto en su vecindario de 18 (9 arriba y 9 abajo). Para cada píxel de superficie de V_S , se calcula la distancia euclídea al vóxel de superficie de V_{GT} más cercano, usando la técnica del vecino más cercano aproximado y se almacena. Para proporcionar simetría (y ser considerado métrica) se aplica el mismo proceso para cada vóxel de V_{GT} con respecto a V_S . La media de todas las medidas almacenadas es el valor final, siendo 0 una segmentación perfecta.

De manera analítica: Sea $S(A)$ el conjunto de vóxeles de superficie de A . La distancia más corta de un vóxel cualquiera v a $S(A)$ se define como:

$$d(v, S(A)) = \min_{s_A \in S(A)} \|v - s_A\| \quad (6.3)$$

siendo $\| \cdot \|$ la distancia Euclídea.

La distancia media simétrica entre las superficies A y B se calcula:

$$ASD(A, B) = \frac{1}{(|S(A)| + |S(B)|)} \left(\sum_{s_A \in S(A)} d(s_A, S(B)) + \sum_{s_B \in S(B)} d(s_B, S(A)) \right) \quad (6.4)$$

Esta métrica junto con el error de solapamiento volumétrico es una de las más usadas.

4. Error cuadrático medio de la distancia simétrica de superficie o *Root Mean Square Symmetric Surface Distance (RMSD)*: el valor se proporciona en milímetros. Esta métrica está fuertemente correlada con la anterior. Conceptualmente es igual pero se calcula la raíz cuadrada de la distancia antes de ser almacenada. En una segmentación perfecta también se obtendría el valor de 0.

$$\begin{aligned} RMSD(A, B) & \quad (6.5) \\ &= \sqrt{\frac{1}{(|S(A)| + |S(B)|)} \left(\sum_{s_A \in S(A)} d^2(s_A, S(B)) + \sum_{s_B \in S(B)} d^2(s_B, S(A)) \right)} \end{aligned}$$

Esta métrica, como ya se ha indicado, está fuertemente correlada con la distancia media. La ventaja que presenta es que grandes desviaciones del contorno verdadero se penalizan más.

5. Máximo de la distancia simétrica de superficie o *Maximum Symmetric Surface Distance (MSD)*: esta medida se proporciona en milímetros y se determina de manera similar a las 2 métricas previas. Se la conoce también como distancia de Hausdorff. Las

diferencias entre los diferentes conjuntos de vóxeles se calculan también mediante distancia euclídea, y la distancia máxima simétrica es el valor máximo del conjunto de distancia. En una segmentación perfecta este valor es 0. Esta métrica es sensible a los bordes y devuelve el error máximo (caso de fugas en la segmentación). Se requiere en aplicaciones de planificación quirúrgica, donde es más importante el error máximo que la media de errores.

$$\text{MSD}(A, B) = \max \left\{ \max_{S_A \in S(A)} d(S_A, S(B)), \max_{S_B \in S(B)} d(S_B, S(A)) \right\} \quad (6.6)$$

Siendo $S(A)$ es el conjunto de vóxeles de la superficie A, y $S(B)$ es el conjunto de vóxeles de la superficie B, en este caso el volumen segmentado y el de referencia.

El segundo grupo de métricas comprende valores que miden la precisión del resultado de segmentación. Estas cuatro métricas adicionales son:

6. Tasa de positivos verdaderos o True Positive Rate (TPR): esta métrica se proporciona en porcentaje y mide la tasa de aciertos en la segmentación, es decir, el porcentaje de píxeles correctamente segmentados con respecto a la referencia. En una segmentación perfecta este valor es 100. Se calcula de la siguiente manera:

$$\text{TPR} = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (6.7)$$

$$TP = V_{GT} \cap V_S$$

$$TN = I - (V_{GT} \cup V_S)$$

$$FP = V_S - TP$$

$$FN = V_{GT} - TP$$

Siendo V_S el volumen segmentado, y V_{GT} la referencia de segmentación para todas las métricas, I es la imagen completa, P es el número de valores positivos de la imagen (etiquetados como '1' o elemento segmentado), N el número de valores negativos de la imagen (etiquetados como '0' o fondo). Un valor de 100 caracteriza una segmentación perfecta.

7. Tasa de falsos positivos o False Positive Rate (FPR): esta métrica se proporciona en porcentaje y representa la tasa de falsos positivos, es decir, el número de píxeles clasificados como el elemento segmentado y no lo son. En una segmentación perfecta este valor es 0. Se calcula de esta manera:

$$\text{FPR} = \frac{FP}{N} = \frac{FP}{FP + TN} \quad (6.8)$$

8. Precisión o Accuracy (ACC): esta métrica indica en porcentaje la precisión de la segmentación realizada. En una segmentación perfecta este valor es 100. Se calcula así.

$$ACC = \frac{TP + TN}{P + N} \quad (6.9)$$

9. **Coefficiente de Similitud de Dice** o *Dice Similarity Coefficient (DSC)*: se indica en porcentaje y es una medida de similitud definida entre dos volúmenes. Un valor de 100 indica una segmentación perfecta. Se obtiene mediante esta expresión:

$$DSC = \frac{2 |TP|}{2|TP| + |FP| + |FN|} \quad (6.10)$$

Una vez elegidas las métricas para la validación de resultados, con la segmentación de referencia y la segmentación obtenida mediante el método propuesta, se aplican las expresiones matemáticas correspondientes y se obtienen los valores de las 9 métricas por cada corte de la secuencia. Se va a proporcionar como valor de cada métrica el valor medio obtenido para todas las imágenes disponibles, esto es, las 63 imágenes de cada una de las 5 secuencias en cada uno de los 6 conjuntos de pruebas, es decir, de las 1.890 imágenes.

La siguiente tabla recoge las expresiones de todas las métricas calculadas:

<p><i>Volumetric Overlap Error (VOE)</i></p>	$VOE = 100 \left(1 - \frac{V_S \cap V_{GT}}{V_S \cup V_{GT}} \right)$
<p><i>Relative Volume Difference (RVD)</i></p>	$RVD = 100 \left(\frac{V_S - V_{GT}}{V_{GT}} \right)$
<p><i>Average Symmetric Surface Distance (ASD)</i></p>	$ASD(A, B) = \frac{1}{(S(A) + S(B))} \left(\sum_{s_A \in S(A)} d(s_A, S(B)) + \sum_{s_B \in S(B)} d(s_B, S(A)) \right)$ $d(v, S(A)) = \min_{s_A \in S(A)} \ v - s_A\ $

<p><i>Root Mean Square Symmetric Surface Distance (RMSD)</i></p>	$RMSD(A, B) = \sqrt{\frac{1}{(S(A) + S(B))}} \times \sqrt{\sum_{s_A \in S(A)} d^2(s_A, S(B)) + \sum_{s_B \in S(B)} d^2(s_B, S(A))}$
<p><i>Maximum Symmetric Surface Distance (MSD)</i></p>	$MSD(A, B) = \max\left\{\max_{s_A \in S(A)} d(s_A, S(B)), \max_{s_B \in S(B)} d(s_B, S(A))\right\}$
<p><i>True Positive Rate (TPR)</i></p>	$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$ $TP = V_{GT} \cap V_S$ $TN = I - (V_{GT} \cup V_S)$ $FP = V_S - TP$ $FN = V_{GT} - TP$
<p><i>False Positive Rate (FPR)</i></p>	$FPR = \frac{FP}{N} = \frac{FP}{FP + TN}$
<p><i>Accuracy (ACC)</i></p>	$ACC = \frac{TP + TN}{P + N}$
<p><i>Dice Similarity Coefficient (DSC)</i></p>	$DSC = \frac{2 TP }{2 TP + FP + FN }$

Tabla 6.2. Expresiones matemáticas para el cálculo de las métricas

6.1.4. Problemas que afectan los resultados: el registro de las secuencias y la calidad de la segmentación de referencia

Uno de los problemas existentes para la correcta segmentación de las imágenes médicas en el caso de disponer de varias secuencias es el registro de las mismas, como se ha explicado en el capítulo 2. Si dos imágenes no están convenientemente registradas, la información disponible en ambas no se corresponde exactamente con el mismo punto físico del espacio, lo cual es un problema para segmentar. Este fenómeno se produce ligeramente en las imágenes de estudio en esta tesis, ya que la máquina de resonancia magnética a partir de la cual se han obtenido las secuencias no registra las mismas convenientemente.

Este problema de registro afecta a los resultados en dos aspectos: por una parte, la calidad de la segmentación y por otra, la determinación de una segmentación manual de referencia.

Debido al problema de registro existente, la calidad de la segmentación en algunos cortes es mala, sobre todo en los cortes iniciales y finales, que es donde más se acusa el efecto, ya que en una secuencia puede haber hígado en un determinado corte y en el mismo corte de otra secuencia no.

En el desarrollo propuesto en esta tesis, no se ha trabajado en métodos de registro, se trata de un campo de trabajo amplio que excede el alcance establecido en esta tesis. Por ello, se ha aplicado un algoritmo para registro deformable basado en información mutua disponible en el paquete software de ITK (Mattes et al., 2003). La secuencia utilizada como referencia en la segmentación es la VIBE en vacío axial, acorde a la información médica recibida. Esta secuencia también es la empleada en el proceso de segmentación manual para la generación de la segmentación de referencia.

Como resultado de este proceso, se observa que las 5 secuencias no están completamente registradas entre sí, encontrándose la secuencia 00 por una parte, y por la otra, el conjunto de 4 secuencias VIBE. Inicialmente se ha decidido generar una segmentación de referencia de cada uno de los grupos de secuencias que no están registrados entre sí, es decir, de la secuencia 00 y de una secuencia VIBE. Las diferencias de registro se acusan más en los cortes iniciales y finales.

Aun así la similitud entre los 63 cortes es del 98,47 % con una desviación estándar de 0.72. Por tanto, existe un error medio de desplazamiento relativo entre las secuencias de 1,53%. Se concluye que no son errores considerables de desplazamiento.

Por criterio clínico se ha elegido la secuencia VIBE como aquella que proporciona más información en cada imagen. Para la evaluación de los métodos tanto sobre una sola secuencia, como en la segmentación multisequencia, se empleará la segmentación de referencia generada manualmente sobre la secuencia denominada VIBE 05 en vacío axial.

El registro multimodal o multicanal en sí mismo es una línea extensa de trabajo y excede el alcance de esta tesis. En cualquier caso, se apunta este problema de registro que, si fuera

abordado debidamente, podría mejorar los resultados de la segmentación, sobre todo en los cortes iniciales y finales del estudio, donde el desajuste es mayor.

También debe considerarse que los resultados de la segmentación pueden verse influenciados por la calidad de la segmentación de referencia, cuyo proceso de obtención se ha explicado en el apartado 6.1.2. Esta segmentación manual puede diferir dependiendo del radiólogo que la realice e incluso el mismo radiólogo puede introducir variaciones en distintos momentos de tiempo, debido a la baja precisión y nitidez de los bordes en algunos cortes de la secuencia. Se ha cuantificado el error que la misma persona ha introducido en la anotación de una secuencia en dos instantes de tiempo distintos. Para ello, se han obtenido las métricas sobre las dos segmentaciones manuales diferentes, lo que se denomina *intra-rater variability*, es decir, variabilidad de la propia persona que establece el ratio de acierto. Ambas se han realizado sobre cada una de las 63 imágenes de una secuencia VIBE, que es la que se ha adoptado como referencia de segmentación. Se han recogido los resultados en la tabla que se muestra a continuación.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
1.78 ± 1.11	0.78 ± 0.69	0.37 ± 0.24	1.09 ± 0.57	5.81 ± 3.56	98.65 ± 0.85	0.11 ± 0.09	99.85 ± 0.14	98.75 ± 0.58

Tabla 6.3. Métricas obtenidas para dos segmentaciones manuales realizadas por la misma persona (*intra-rater variability*)

De la anterior tabla, se puede apreciar que en este caso, los errores son pequeños, siendo el valor de MSD la métrica más afectada por la diferencia en la segmentación manual. Dado que ambas segmentaciones se consideran válidas, se ha escogido una de ellas como segmentación de referencia del proceso de pruebas y validación.

Como se ha explicado previamente, se dispone de otra segmentación manual realizada por una segunda persona. La variabilidad entre ambos validadores, lo que se conoce como *inter-rater variability*, se recoge en la siguiente tabla.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
10.40 ± 2.99	4.45 ± 3.70	2.41 ± 0.92	3.26 ± 1.20	11.82 ± 6.28	91.99 ± 3.42	0.55 ± 0.49	98.96 ± 0.56	93.39 ± 1.69

Tabla 6.4. Métricas obtenidas para dos segmentaciones manuales realizadas por distintas personas (*inter-rater variability*)

Como se puede observar, la variabilidad entre personas es mayor que la variabilidad entre dos segmentaciones realizadas por la misma persona, como es de esperar. Se ha escogido como segmentación de referencia una de las realizadas por uno de los validadores.

6.2. Generación del modelo estadístico multivariable de hígado

El modelo estadístico multivariable de hígado se realiza en base a 3 estudios disponibles, en la manera descrita en el apartado 5.6.2. De manera resumida, y en lo referente a la implementación, el tamaño de la muestra de hígado que se emplea para generar el modelo gaussiano multivariable de hígado es de 10 x 50 píxeles en el caso monosecuencia, y 10 x 50 x 5 píxeles en el caso multisecuencia. Se calcula este valor en los 10 cortes comprendidos entre el corte 30 y el 40, que tienen porciones de hígado suficientemente grandes para tomar un muestrario relevante. El hecho de que el tamaño en el eje 'y' sea mayor que en el eje 'x' es debido a la no uniformidad en la iluminación de la imagen que se produce en este eje 'y', y es por tanto necesario abarcar una superficie mayor para que el modelo generado sea capaz de absorber esta variabilidad. Para la generación de este modelo de hígado se han calculado los valores de media y desviación estándar del conjunto de píxeles indicado.

6.3. Segmentación en 2D por minimización de contorno activo mediante técnicas variacionales y basado en modelo

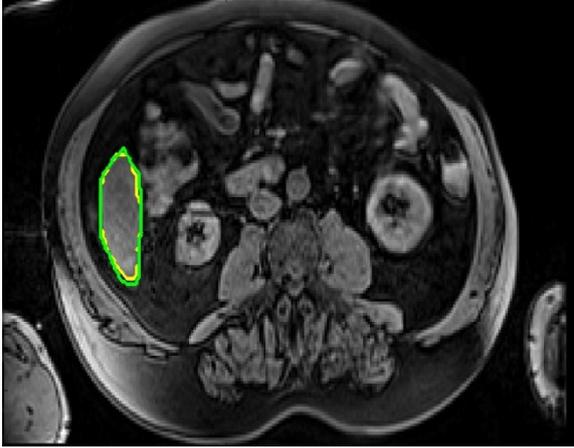
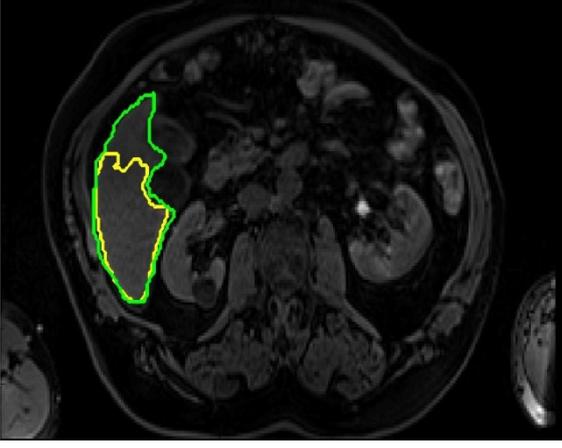
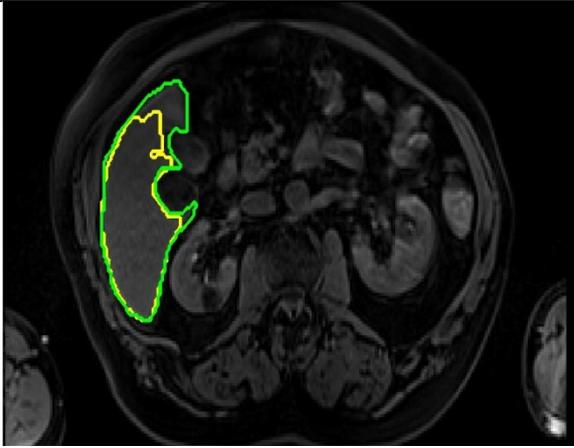
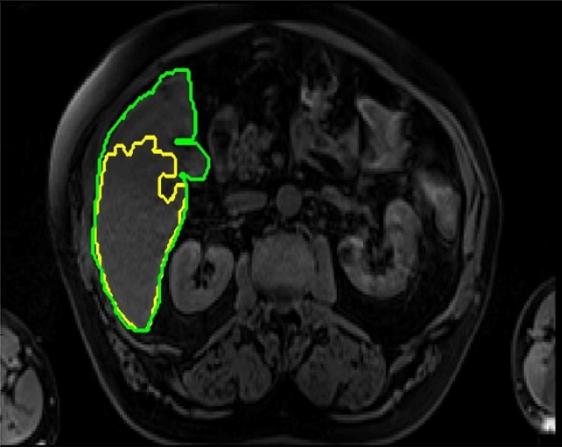
En esta sección se van a exponer los resultados obtenidos mediante el primero de los métodos propuestos para la segmentación de hígado, que consiste en la minimización del contorno activo definido mediante la aproximación dual al problema variacional realizado en 2D, cuyo fundamento teórico ha sido detallado en el apartado 5.4. Se va a validar el método sobre los dos tipos de información disponible en la entrada: una sola secuencia o varias secuencias simultáneas de la misma adquisición. Se muestran en las siguientes páginas los diferentes cortes del estudio con el resultado de la segmentación con respecto a la segmentación de referencia disponible así como las métricas obtenidas.

En el proceso de minimización, que va a regir la evolución del contorno activo en cada iteración en búsqueda de la solución óptima y tal como se ha detallado en el apartado 5.4 del capítulo anterior, existen algunos parámetros configurables, como λ , que gestiona la regularización de la segmentación, es decir, cómo de suaves pueden ser los contornos. A mayor λ , más suaves son las curvas, y se ajusta peor a superficies con formas muy irregulares. Para poder ajustarse bien a contornos abruptos, es necesario fijar valores de λ pequeños. β indica la importancia del gradiente en la localización de los bordes, t es un intervalo de paso, y N es el número de iteraciones que debe realizar el proceso de minimización si no ha convergido antes. Los parámetros establecidos para encontrar una segmentación precisa fueron conseguidos con valores de $\lambda = 0.01$, $\beta = 50/255^4$, $t = 1/8$ y $N = 60$. Los valores c_1 y c_2 se actualizan cada 10 iteraciones, como ya se ha indicado en el capítulo anterior y se recuerda aquí.

La implementación del proceso de minimización mediante el método indicado es independiente de las condiciones de inicialización. En cualquier caso, es necesario elegir unos valores iniciales y éstos han sido $u = v = \frac{I_0}{\max(I_0)}$, $c_1 = c_2 = p_x = p_y = p_z = 0$, siendo I_0 la imagen inicial.

6.2.1. Entrada Monosecuencia (raw VIBE data)

En la siguiente tabla se exponen los resultados obtenidos para la segmentación sobre algunos de los cortes de la secuencia VIBE 05, con los datos tal y como son obtenidos por la máquina de MRI. No son procesados. No se ha generado el mapa de distancias mediante la aplicación del descriptor compacto propuesto. Las imágenes tienen muy poco contraste y las técnicas de segmentación más frecuentes no proporcionan buenos resultados. Se van a validar las técnicas variacionales sobre esta imagen tan pobre. Sobre cada imagen de la secuencia VIBE se muestra el contorno del hígado segmentado por nuestro método (amarillo) y el contorno del hígado de referencia (verde).

	
Corte 5	Corte 20
	
Corte 24	Corte 28

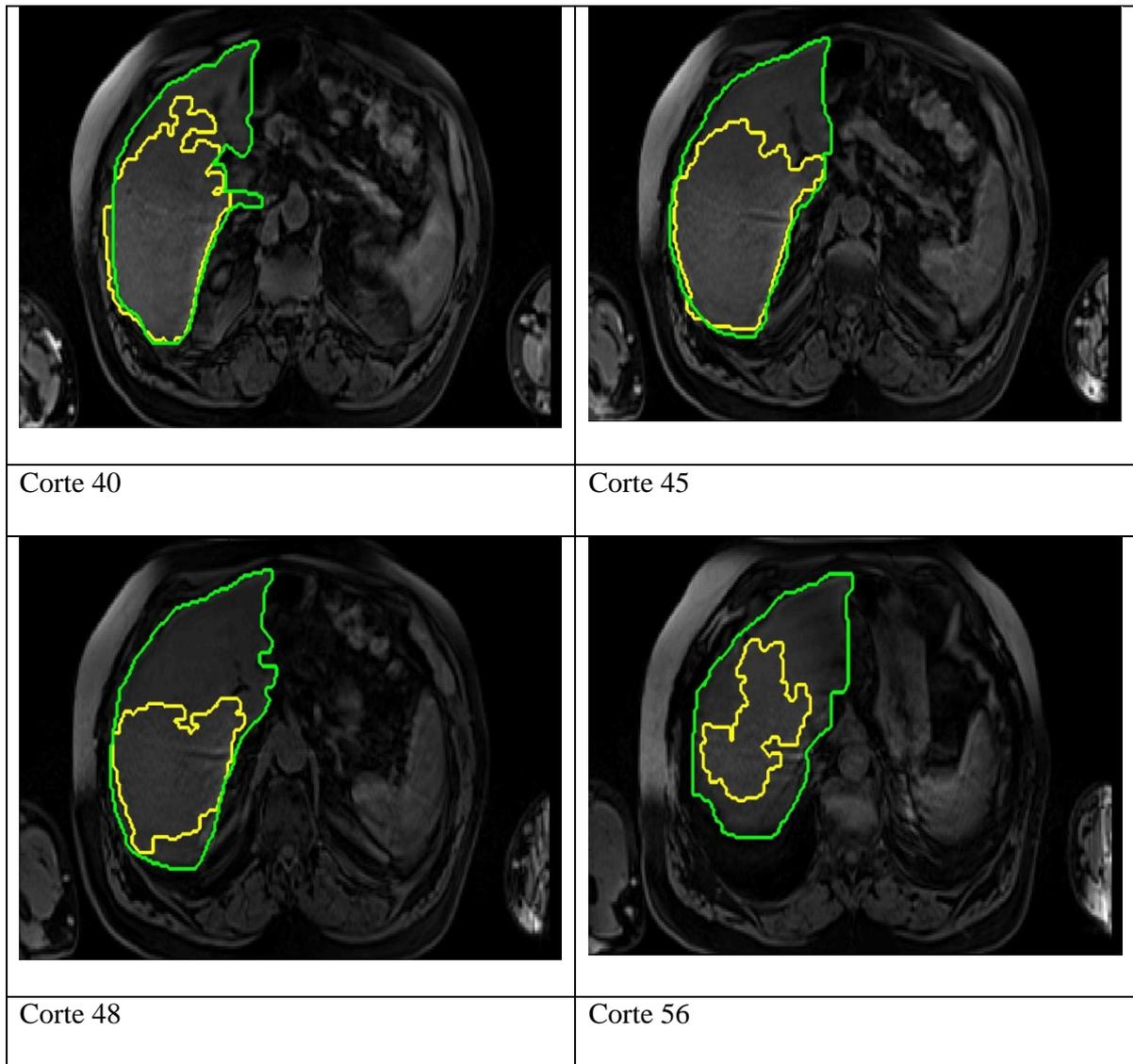


Figura 6.3. Imágenes de diferentes cortes de la secuencia VIBE 05 con hígado segmentado por método de minimización de contorno activo en 2D y entrada raw VIBE (amarillo) y contorno de hígado de referencia (verde)

En esta prueba, no se ha localizado el hígado en casi la mitad de los cortes. La razón es el bajo contraste existente entre los tejidos presentes en la imagen. El ojo humano puede intuir el contorno en algunas imágenes, pero con baja precisión. En determinados cortes la realización de la segmentación es muy compleja. La realización automática de esta tarea, incluso para un contorno activo avanzado como es el propuesto, no proporciona buenos resultados. Es necesario mejorar la calidad de la imagen, aumentar el contraste entre los elementos presentes en ella.

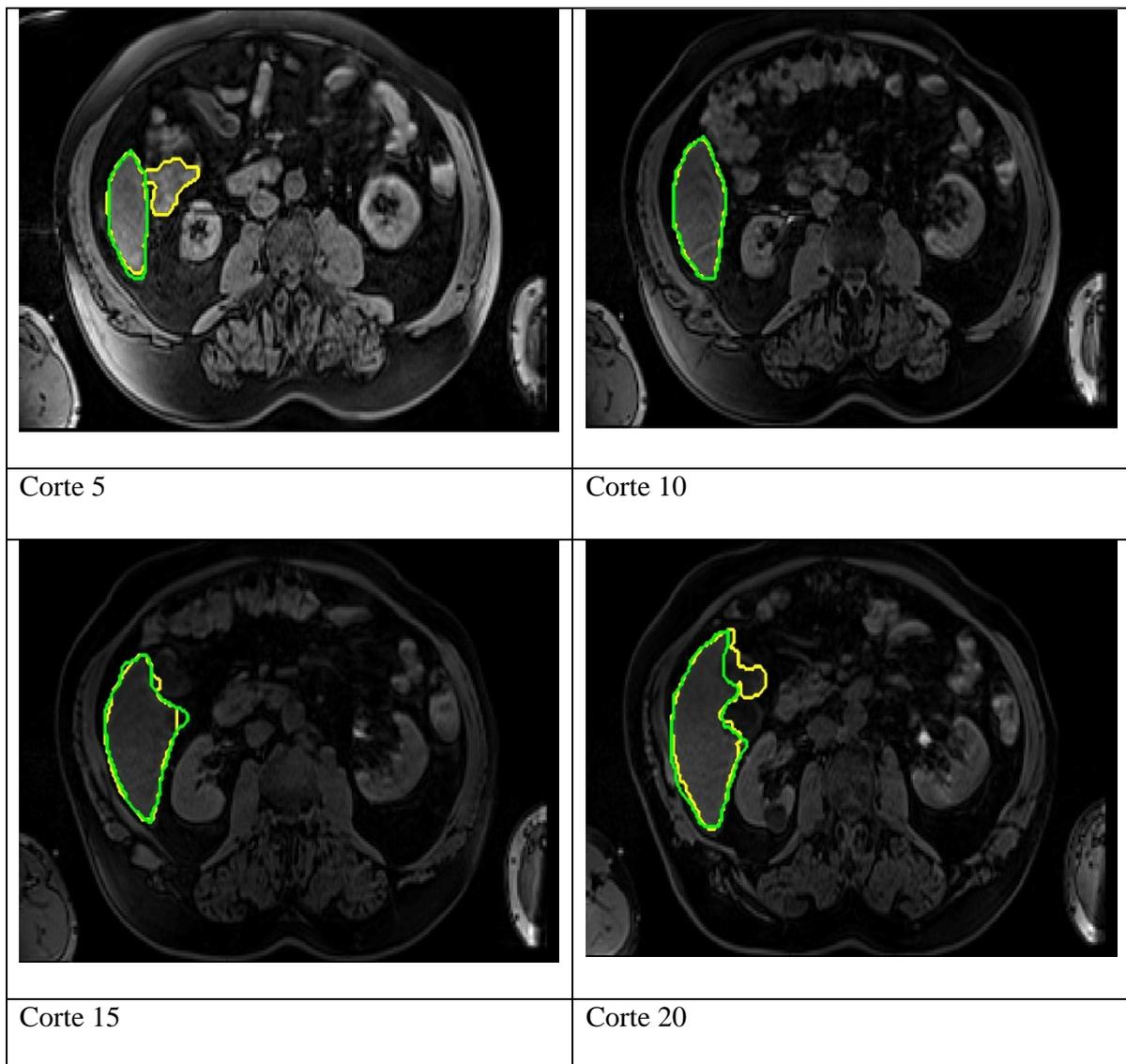
Las métricas obtenidas con esta entrada se muestran a continuación:

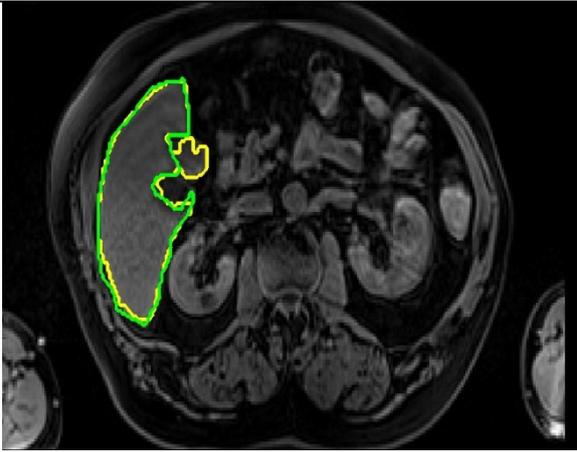
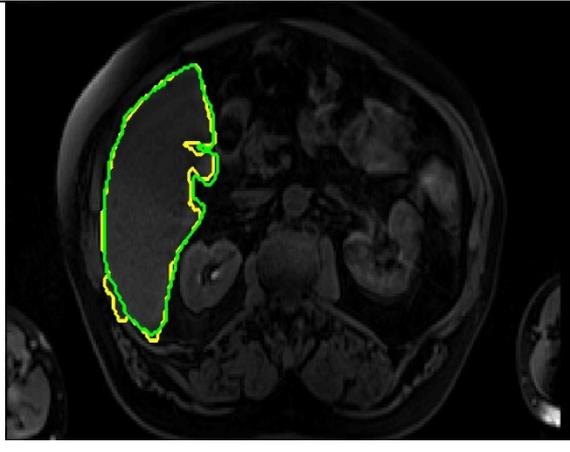
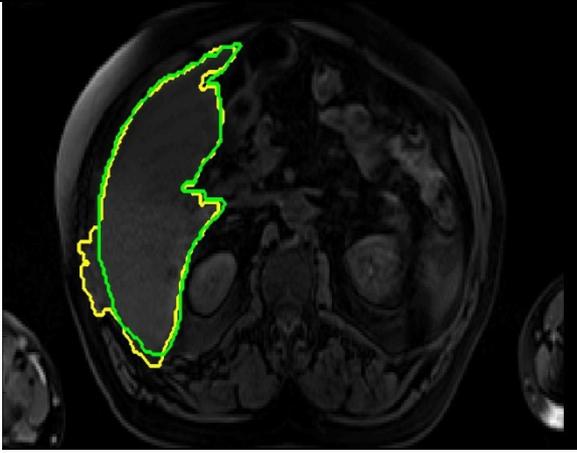
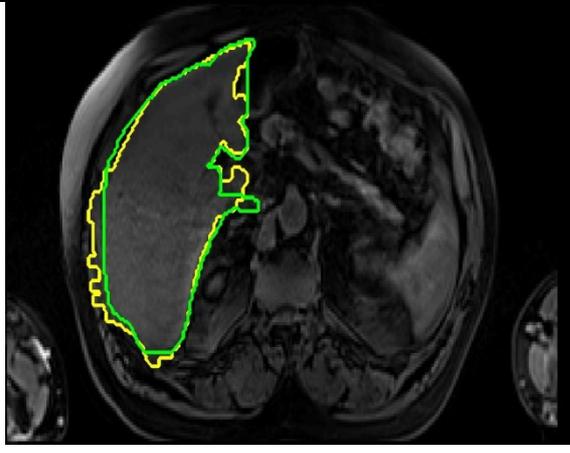
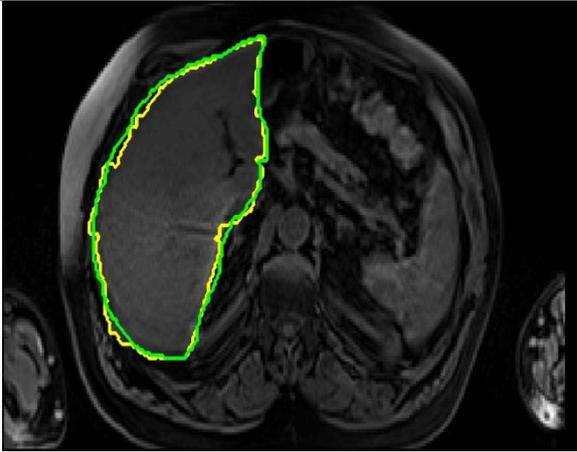
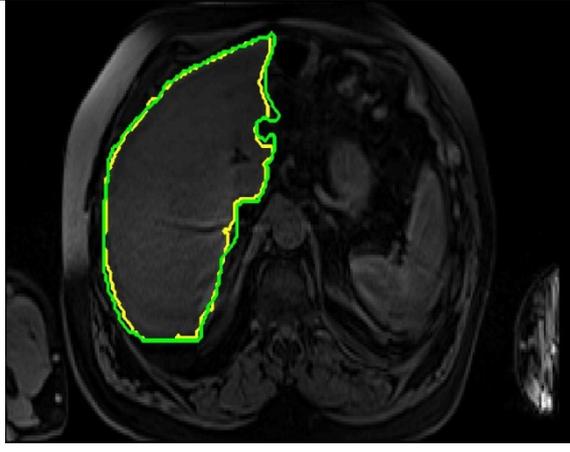
VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
39.38 ± 20.79	41.78 ± 28.66	14.05 ± 11.28	20.32 ± 13.64	60.09 ± 31.69	66.18 ± 22.11	0.22 ± 0.36	96.31 ± 3.06	73.21 ± 17.71

Tabla 6.5. Métricas obtenidas de los resultados de segmentación mediante el método de minimización de contorno activo en 2D y entrada imagen de grises monosecuencia VIBE 05 (sin aplicación del descriptor compacto para generación del mapa de distancias)

6.2.2. Entrada Monosecuencia (con mapa de distancias)

En la siguiente tabla se exponen los resultados obtenidos para la segmentación sobre algunos de los cortes de la secuencia VIBE 05. Sobre cada imagen de la secuencia VIBE se muestra el contorno del hígado segmentado por nuestro método (amarillo) y el contorno del hígado de referencia (verde).



	
Corte 25	Corte 30
	
Corte 35	Corte 40
	
Corte 45	Corte 50

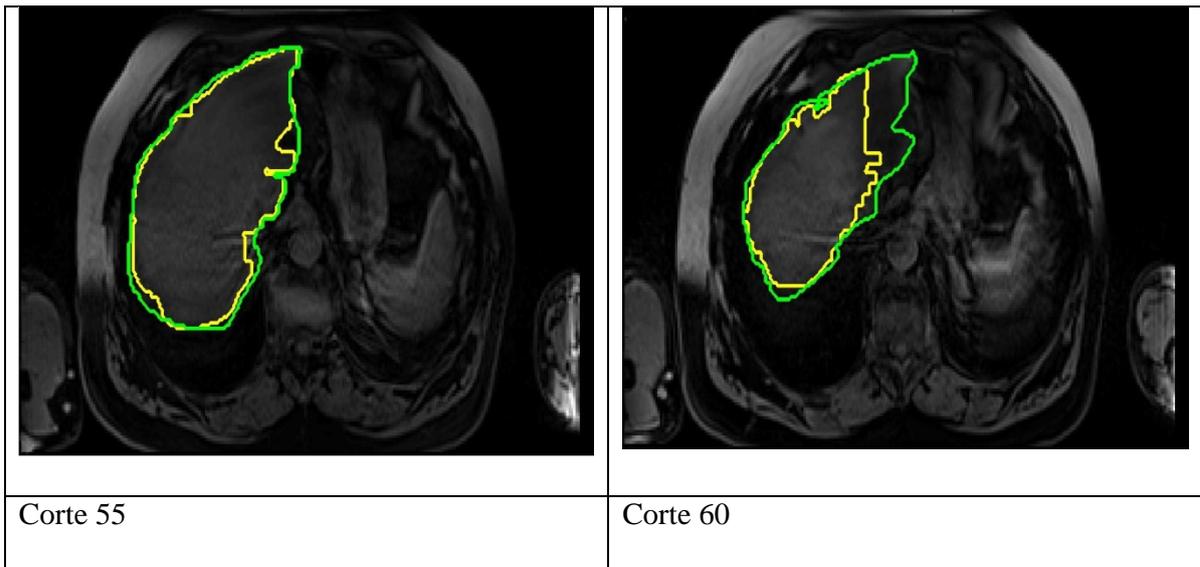


Figura 6.4. Imágenes de diferentes cortes de la secuencia VIBE 05 con hígado segmentado por método de minimización de contorno activo en 2D y entrada monosecuencia (amarillo) y contorno de hígado de referencia (verde)

En la tabla 6.5 se muestran los valores de las métricas obtenidos, con respecto a la segmentación de referencia generada manualmente sobre la secuencia VIBE 05, en la manera en que ha sido descrita en 6.1.2. Estos valores son la media de las métricas obtenidas para los 6 estudios disponibles. Como en este apartado se está analizando el caso monosecuencia, se ha validado sobre las 63 imágenes de cada secuencia VIBE 05 en cada uno de los 6 estudios para pruebas disponible, es decir 63 imágenes x 6 estudios, lo que hace un total de 378 imágenes. Se ha adoptado como valor de la métrica final, como se ha indicado, la media de los valores obtenidos para todos los cortes en cada secuencia disponible.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
12.41 ± 6.11	7.89 ± 6.31	3.04 ± 1.24	4.93 ± 1.95	19.33 ± 7.36	94.00 ± 5.10	0.51 ± 0.32	98.97 ± 0.65	93.07 ± 3.54

Tabla 6.6. Métricas obtenidas de los resultados de segmentación mediante el método de minimización de contorno activo en 2D y entrada monosecuencia

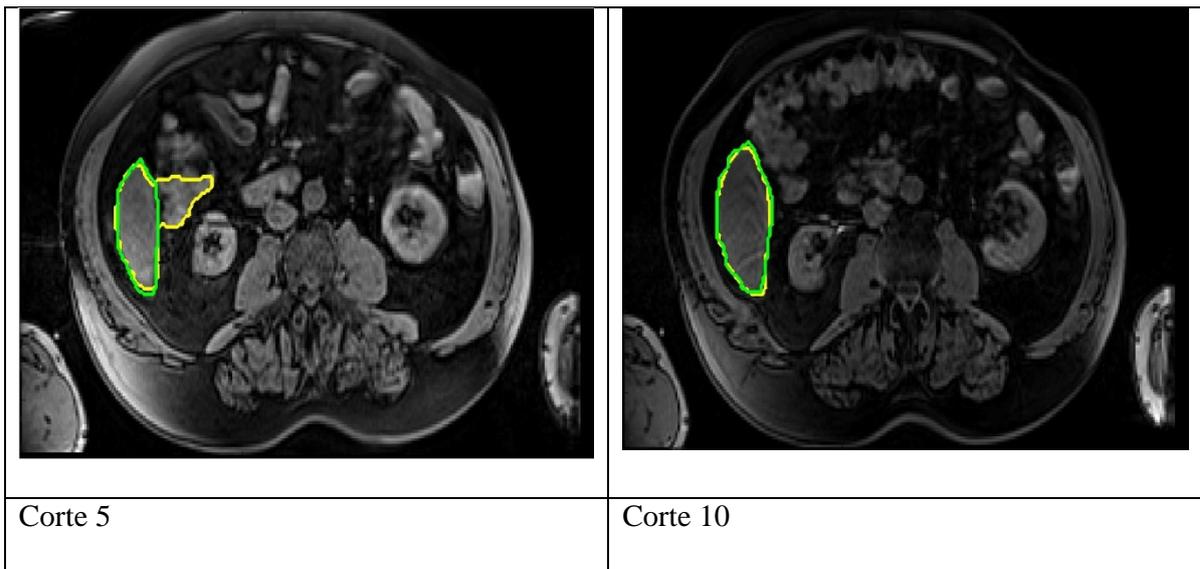
Una de las ventajas cualitativas que se aprecian de la observación de los resultados es que las técnicas variacionales propuestas, permiten que los contornos de las superficies que se adaptan y minimizan hasta encontrar los bordes de las regiones sean totalmente flexibles. Existe un parámetro configurable, en este caso λ , que permite controlar lo suave o abrupto que pueden ser los bordes de la curva, como ya se ha indicado. El valor adoptado de λ es 0.01, que proporciona un equilibrado suavizado con bordes no demasiado abruptos y con flexibilidad suficiente para adecuarse a todas las curvaturas existentes.

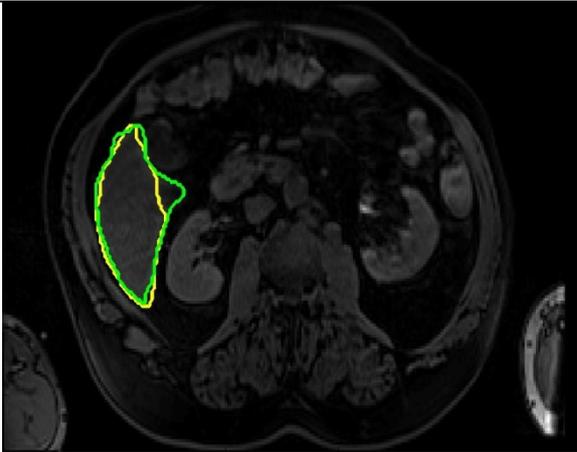
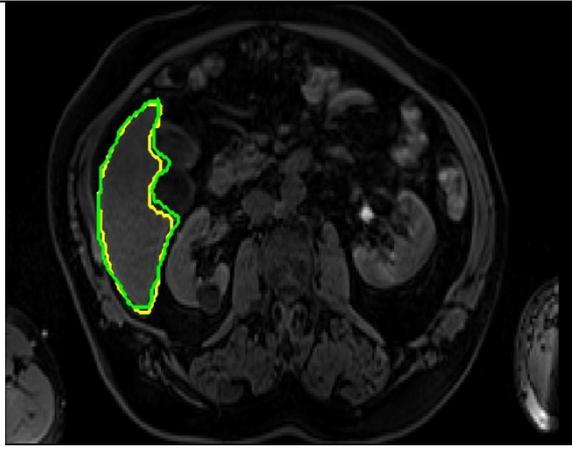
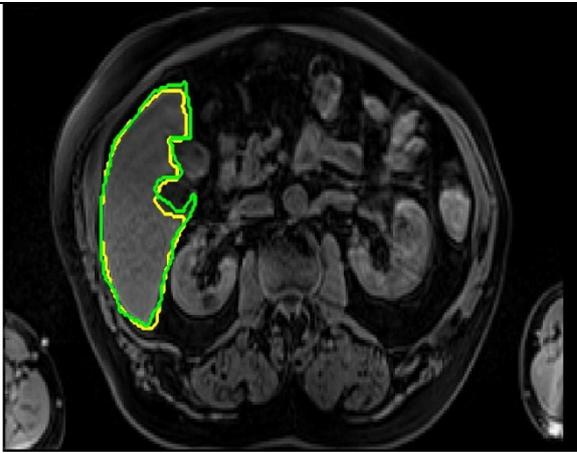
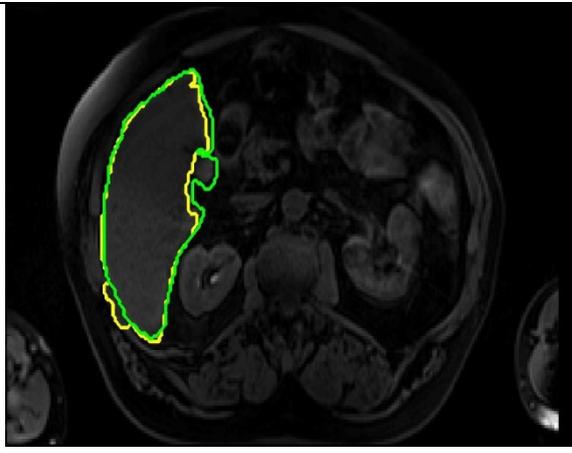
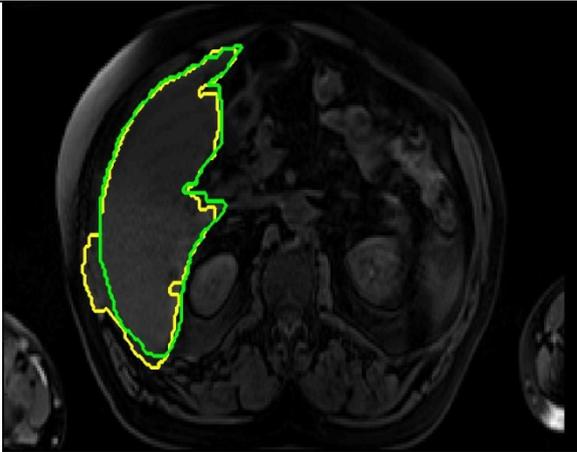
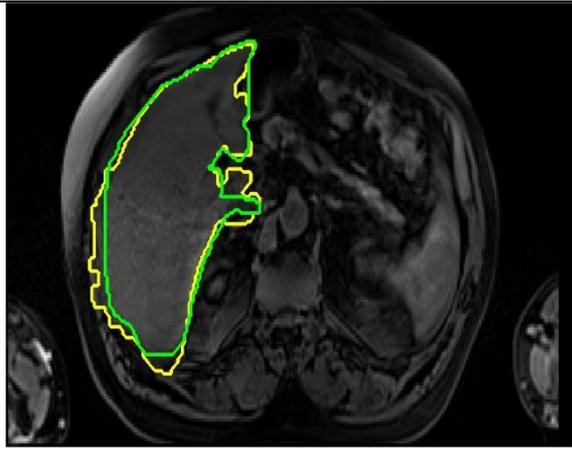
6.2.3. Entrada Multiseuencia

Análogamente al caso anterior, se ha aplicado la misma metodología para la segmentación de la imagen acorde a la clasificación de los píxeles como pertenecientes al hígado o no. Se dispone de 5 secuencias, lo que convierte cada estudio del paciente en información en 4D de dimensiones 160 x 256 x 63 x 5 píxeles. La aproximación propuesta permite tratar corte a corte la secuencia, disponiendo de información característica en las 5 secuencias, de dimensiones 160 X 256 x 5 y no sólo una imagen 2D de 160 x 256, como en el caso monosecuencia, y donde se ha transformado la información multiseuencia mediante un descriptor compacto que permite reducir la dimensionalidad.

De manera análoga a como se ha realizado previamente, se va a aplicar el método de segmentación propuesto en el capítulo 5 sobre la entrada de datos multiseuencia. Para poder emplear el potencial de la información complementaria de la multiseuencia, se propone aplicar el método de segmentación sobre la nueva imagen de entrada generada mediante el descriptor compacto, que representa un mapa de distancias construido a partir de las 5 secuencias disponibles, como se ha explicado en el apartado 5.2. De esta manera la entrada de 5 secuencias se ha transformado en un mapa de distancias de cada píxel de la imagen al modelo de hígado previamente generado. La imagen resultante presenta mayor contraste.

Los resultados obtenidos de la aplicación del método en 2D sobre la nueva entrada multiseuencia se muestran a continuación. Se muestran los mismos cortes de la misma secuencia VIBE mostrada antes, con el hígado segmentado por el método propuesto con contorno amarillo, y el hígado de referencia tras la segmentación manual con contorno verde.



	
Corte 15	Corte 20
	
Corte 25	Corte 30
	
Corte 35	Corte 40

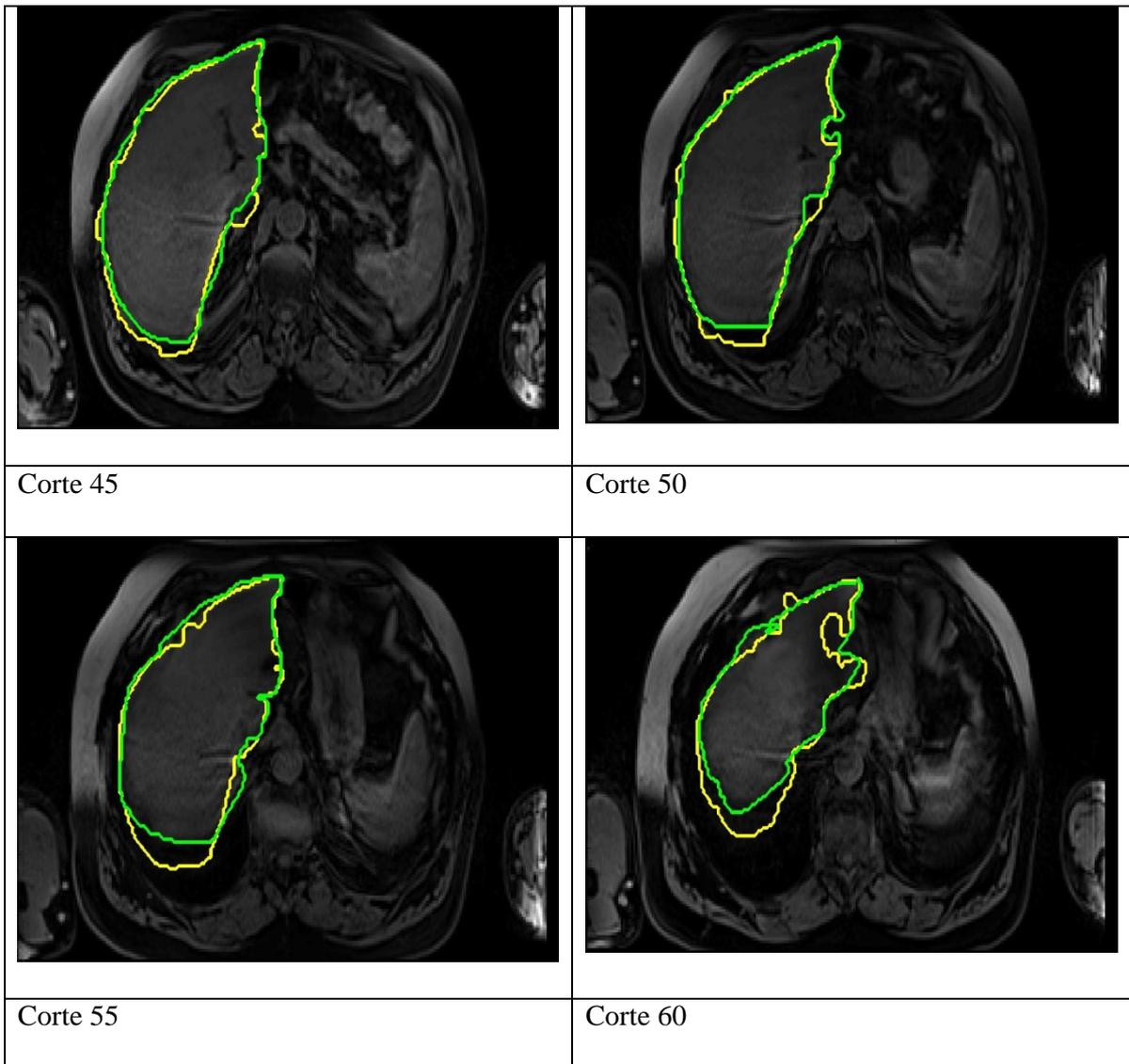


Figura 6.5. Imágenes de diferentes cortes de la secuencia VIBE 05 con hígado segmentado por método de minimización de contorno activo en 2D y entrada multisequencia (amarillo) y contorno de hígado de referencia (verde)

A simple vista se aprecia que la segmentación es cualitativamente mejor que la obtenida a partir de una única secuencia. Las limitaciones identificadas en el caso anterior han sido compensadas por la información complementaria existente en las secuencias disponibles.

Tras este análisis cualitativo, la comparación estricta de la bondad del método se realiza a partir de las métricas obtenidas según la metodología establecida.

En la tabla 6.6 se muestran los valores de las métricas obtenidas, con respecto a la segmentación de referencia generada manualmente sobre la secuencia VIBE 05. Estos valores son la media de las métricas obtenidas para los 6 estudios disponibles. Como en este apartado se está analizando el caso multisequencia, se ha validado sobre las 63 imágenes de cada una de las 5 secuencias en cada uno de los 6 estudios para pruebas disponible, es decir

63 imágenes x 6 estudios x 5 secuencias / estudio, lo que hace un total de 1.890 imágenes. Se ha adoptado como valor de la métrica final, como se ha indicado, la media de los valores obtenidos para todos los cortes.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
1.98 ±	2.01 ±	0.88 ±	2.27 ±	10.59 ±	99.62 ±	0.23 ±	99.81 ±	98.64 ±
1.82	1.92	0.73	1.70	7.77	0.11	0.22	0.19	0.92

Tabla 6.7. Métricas obtenidas de los resultados de segmentación mediante el método de minimización de contorno activo en 2D y entrada multisequencia

De la realización de las pruebas y del análisis de los resultados de la tabla tras la aplicación del método propuesto se pueden extraer varias conclusiones:

- El método de segmentación mediante un contorno activo en 2D minimizado por técnica variacional y aproximación dual de Chambolle sobre una entrada multisequencia de imágenes, con la aproximación de ser aplicado sobre la nueva imagen generada mediante el descriptor compacto propuesto y que representa un mapa de distancias con respecto al modelo estadístico de hígado multivariable proporciona mejores resultados que si se aplica el mismo método sobre una sola secuencia.
- La aproximación propuesta al problema multisequencia aporta valor, es decir, que la entrada al método de segmentación sea un mapa de distancias generado a partir de la distancias de los valores de gris de los píxeles de las secuencias con respecto al modelo estadístico multivariable obtenido del hígado, favorece que los resultados sean comparativamente mucho mejores que sobre una sola secuencia, y mediante aplicación de la misma técnica variacional, en comparación con los resultados presentados en el apartado anterior.
- El mapa de distancias que se genera a partir de las 5 secuencias proporciona una imagen de mejor calidad en cuanto a uniformidad de la iluminación y mayor contraste entre órganos y tejidos. De esta manera, la región perteneciente al hígado aparece más oscura y con más contraste con respecto a la zona circundante. Además este método ha permitido reducir la dimensionalidad del problema, mediante el descriptor compacto propuesto que representa la información de las 5 secuencias en una sola secuencia.
- Los mejores resultados de la segmentación se obtienen en los cortes centrales de las secuencias, siendo especialmente buenos entre los cortes 11 y 55, la zona central del hígado (salvo alguna excepción). Esto es razonable dado que los cortes iniciales y finales de cada secuencia se corresponden con las zonas más externas del hígado, donde más confuso es establecer los límites (incluso para el anotador manual) y donde las secuencias aparecen peor registradas.

- El tiempo necesario para la realización de la segmentación 2D tanto sobre entrada monosecuencia como multisecuencia es inferior a 1 minuto, por lo que se considera un método válido para ser integrado en la práctica clínica.

6.4.Método de segmentación 3D por minimización de contorno activo mediante técnicas variacionales y basado en modelo

En esta sección se van a exponer los resultados obtenidos de la segmentación de hígado mediante el contorno activo definido y minimizado mediante la aproximación dual al problema variacional propuesta por Chambolle extendido sobre una entrada en tres dimensiones, cuyo fundamento teórico ha sido detallado en el apartado 5.5. De igual manera al caso anterior se va a validar el método sobre los dos tipos de entrada disponible: una sola secuencia o varias secuencias adquiridas sobre el mismo paciente en la misma sesión.

Se pretende validar la bondad de la extensión a 3D del contorno activo en sí y de la solución dual al problema variacional para su minimización que previamente se ha implementado en 2D. Parece razonable suponer que el hecho de considerar todo el volumen que representa la secuencia permitirá disponer de más información para que las funciones a minimizar sean capaces de encontrar los puntos de separación entre las regiones existentes en el volumen, y por tanto, obtener una segmentación óptima.

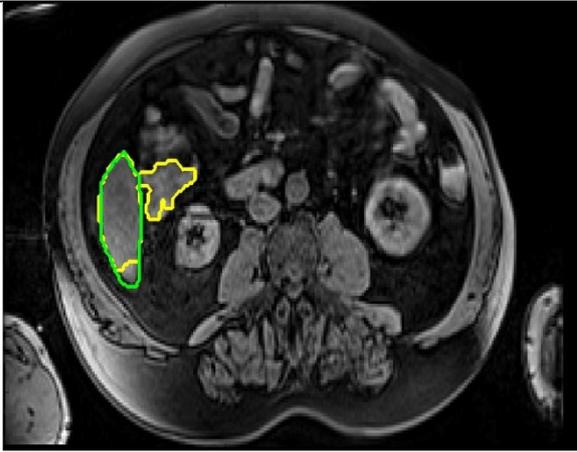
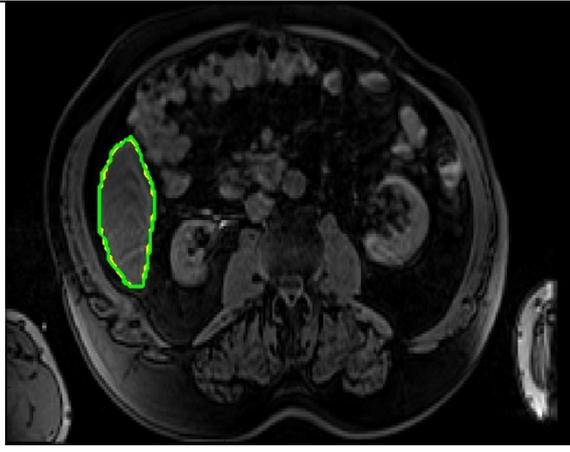
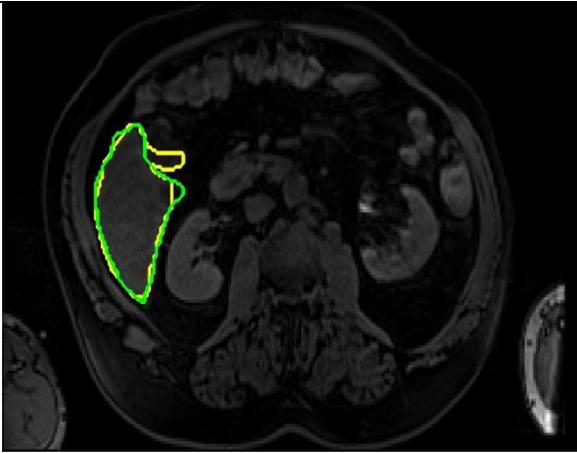
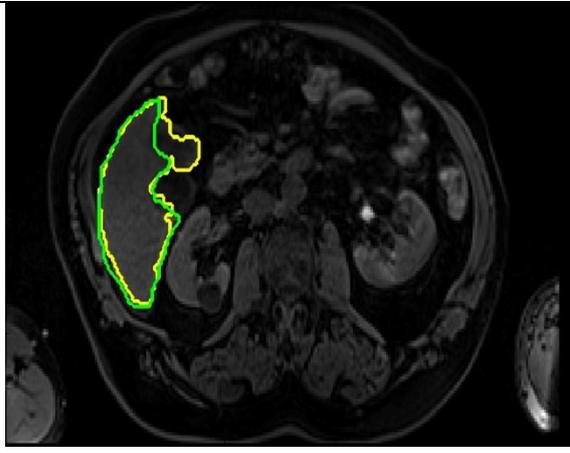
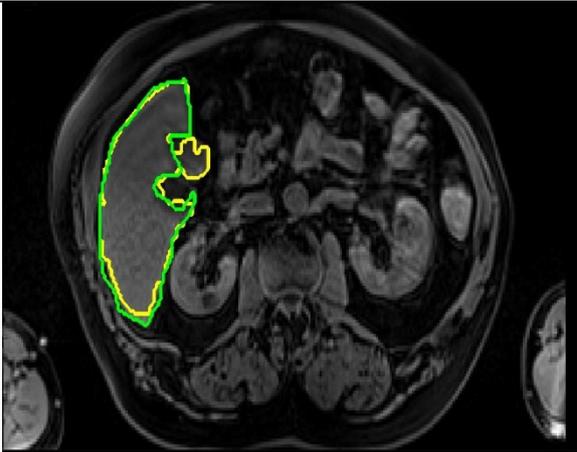
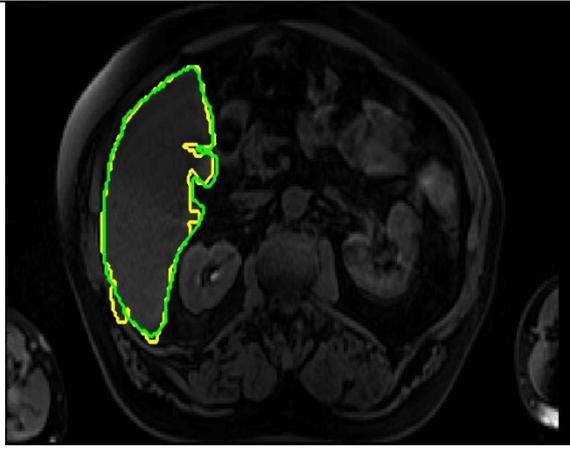
El flujo de procesamiento y la metodología de pruebas es idéntica a la expuesta para el caso precedente en 2D.

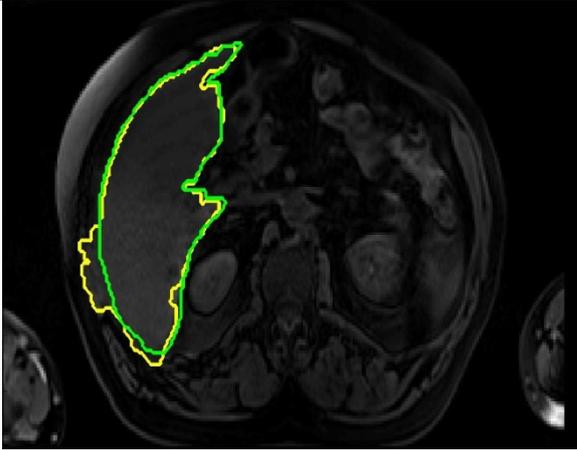
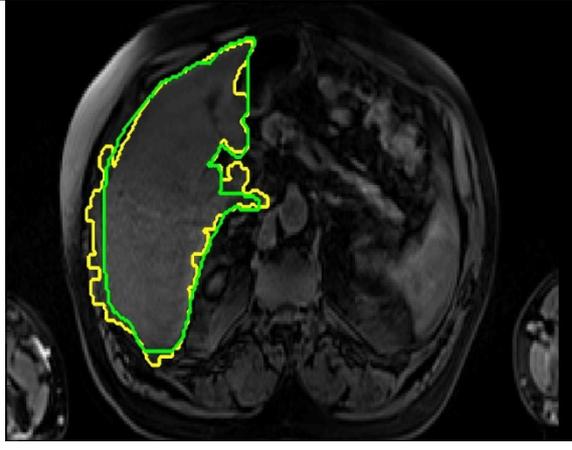
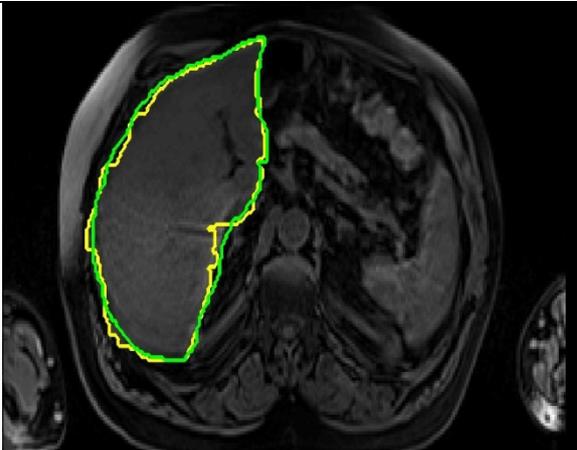
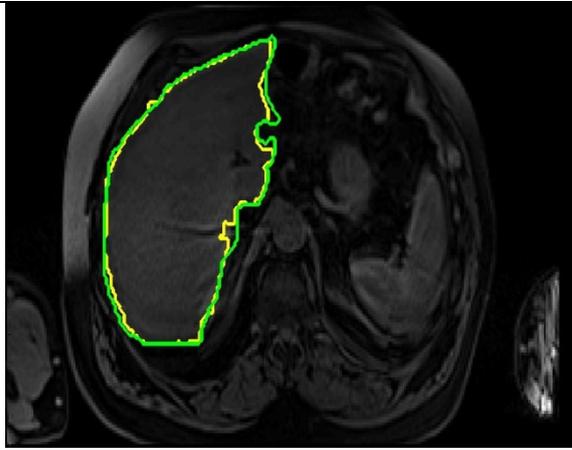
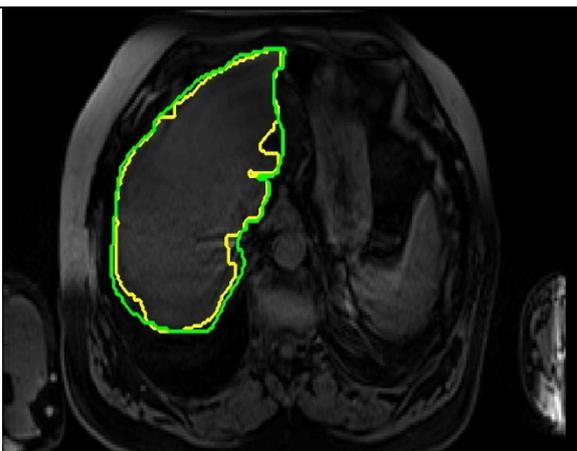
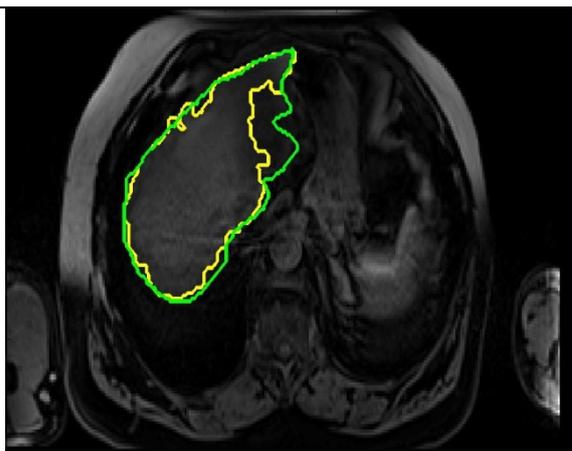
6.4.1. Entrada Monosecuencia

Se ha extendido el método de segmentación 2D a 3D, mediante la implementación de las premisas indicadas en la sección 5.5 de esta tesis. Una vez hecho esto, se valida el método en 3D sobre la secuencia VIBE 05, y sus 63 cortes. De esta manera las funciones minimizantes deben localizar las regiones en un volumen de 160 x 256 x 63 píxeles.

El proceso de segmentación se realiza en 3D, pero para validarlo, comparándolo con el método en 2D, y visualizarlo rigurosamente con las métricas adoptadas y análogamente a los procesos anteriores, se ha realizado un barrido 2D en el eje z generando cortes similares a los que se disponía previamente una vez ha finalizado el proceso de segmentación.

Se muestran a continuación algunos cortes con el hígado segmentado señalado con contorno amarillo junto con la segmentación de referencia con contorno verde.

	
Corte 5	Corte 10
	
Corte 15	Corte 20
	
Corte 25	Corte 30

	
Corte 35	Corte 40
	
Corte 45	Corte 50
	
Corte 55	Corte 59

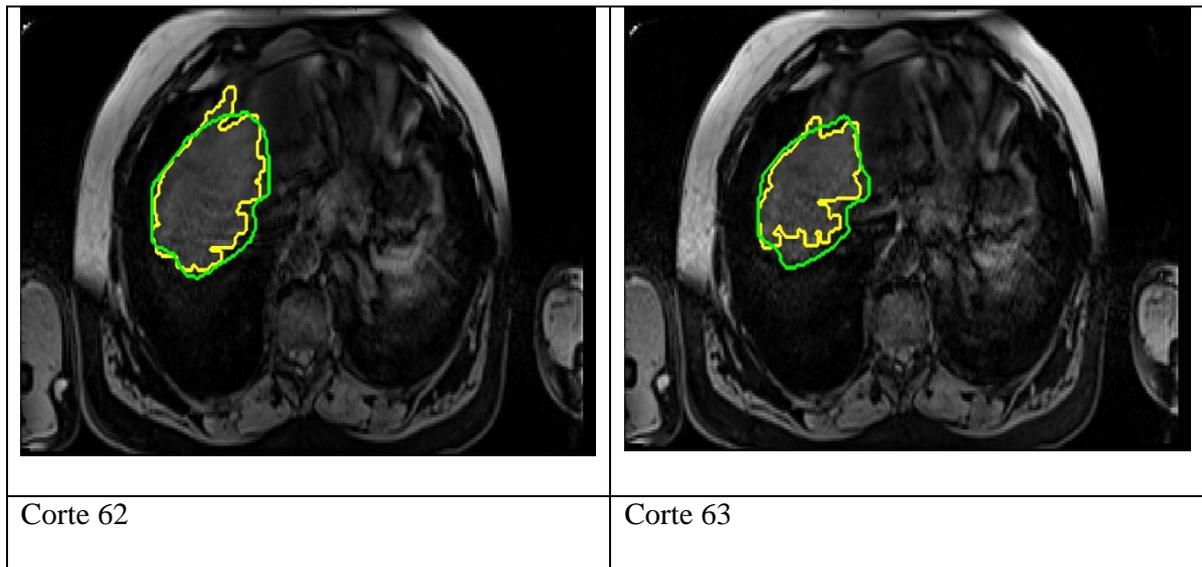


Figura 6.6. Imágenes de diferentes cortes de la secuencia VIBE 05 con hígado segmentado por el método de minimización de contorno activo en 3D y entrada monosecuencia (amarillo) y contorno de hígado de referencia (verde)

En la figura 6.6 se pueden apreciar algunos cortes seleccionados de la secuencia VIBE 05, donde el hígado segmentado automáticamente ha sido marcado con el contorno amarillo. Se puede apreciar que las curvas son suaves y se adaptan a todo tipo de estructuras, ciñéndose al borde perfectamente.

A continuación, en la tabla 6.7 se muestran los valores de las métricas obtenidos por este método extendido a 3D con respecto a la segmentación de referencia disponible generada manualmente sobre la secuencia VIBE 05. Estos valores son la media de las métricas obtenidas para los 6 estudios disponibles. Como en este apartado se está analizando su aplicación sobre entrada monosecuencia, se ha validado sobre las 63 imágenes de cada secuencia VIBE 05 en cada uno de los 6 estudios para pruebas disponible, es decir 63 imágenes x 6 estudios, lo que hace un total de 378 imágenes. Se ha adoptado como valor de la métrica final, como se ha indicado anteriormente, la media de los valores obtenidos para todos los cortes en cada secuencia disponible.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
11.86 ± 4.96	4.83 ± 3.00	3.07 ± 0.99	4.55 ± 1.82	17.91 ± 7.30	94.24 ± 1.90	0.48 ± 0.41	99.05 ± 0.47	93.48 ± 2.86

Tabla 6.8. Métricas finales de los resultados de segmentación por el método de minimización de contorno activo en 3D y entrada monosecuencia

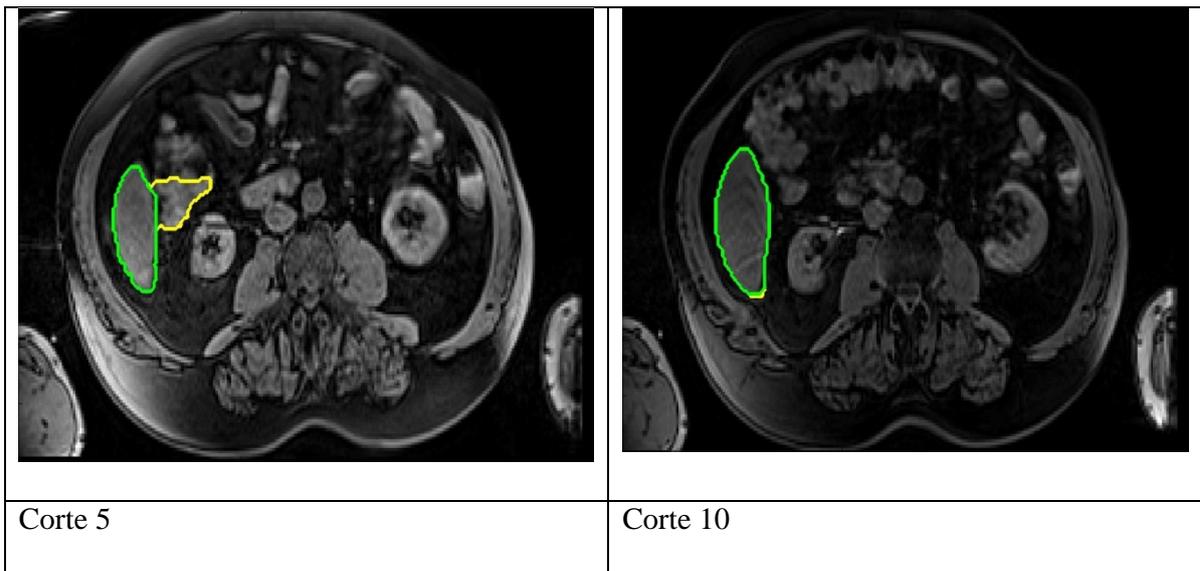
Los resultados obtenidos son muy buenos. El único inconveniente de la aplicación de este método de segmentación es la carga computacional que implica, ya que segmentar un volumen en 3D mediante minimización de un contorno activo definido es costoso ya que añade una dimensión más al problema. El tiempo empleado en realizar una segmentación de

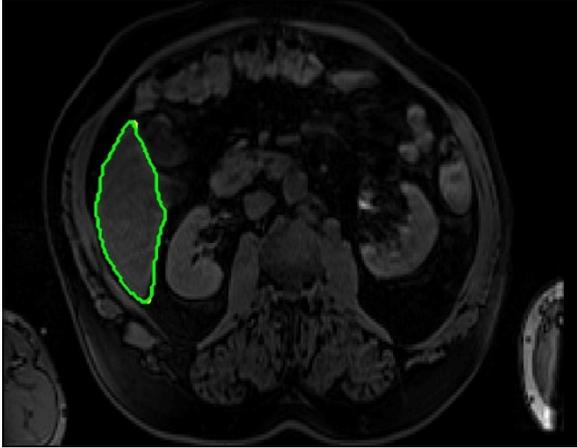
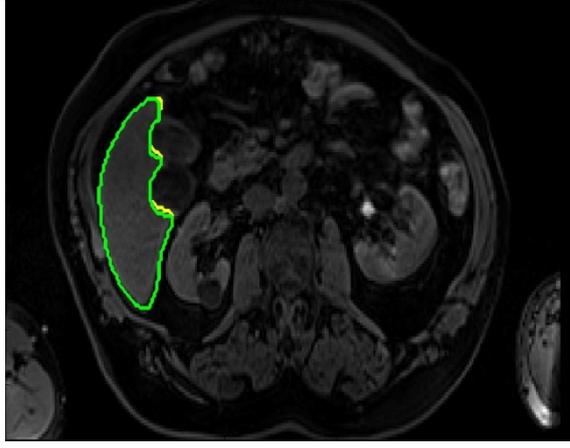
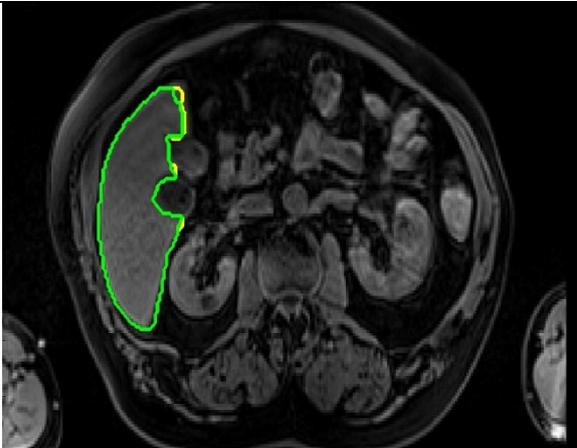
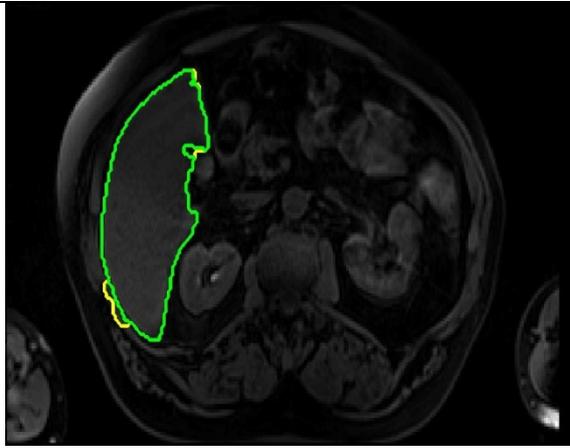
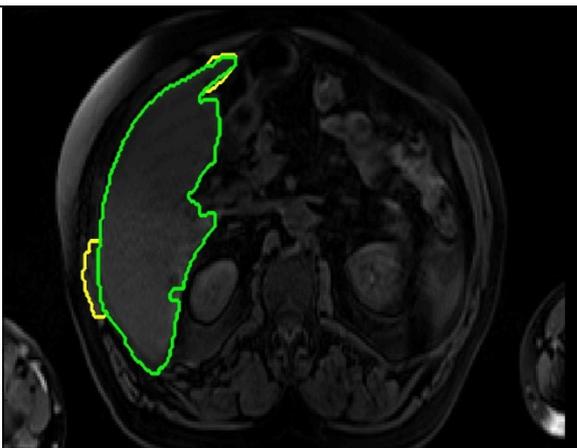
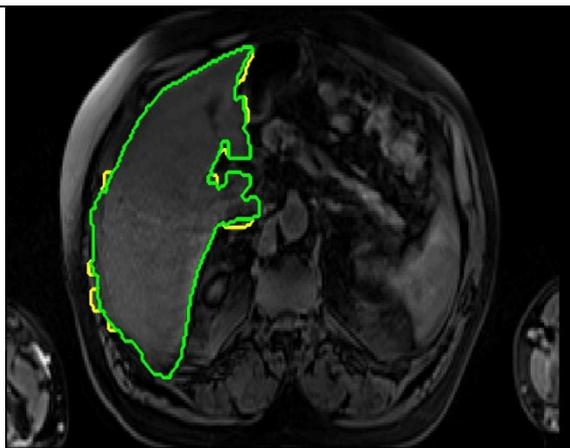
la secuencia completa es de aproximadamente 40-50 segundos en un Dell Optiplex 780, Intel core 2 quad CPU Q9650 a 3 GHz y de 64 bits. Se presenta como un método valioso para conseguir una buena segmentación del órgano.

6.4.2. Entrada Multisequencia

De forma similar a los casos anteriores, se va a aplicar el método en 3D sobre la entrada multisequencia disponible. El manejo de la entrada multisequencia se realiza mediante la obtención del descriptor compacto propuesto, de igual manera que se ha realizado para la segmentación en dos dimensiones. Con las 5 secuencias disponibles y mediante el descriptor compacto propuesto se va a generar el mapa de distancias al modelo estadístico multivariable obtenido del hígado. Este mapa de distancias de cada uno de los 63 cortes de las secuencias, considerados como un volumen, constituye la nueva entrada del método de segmentación en tres dimensiones objeto de estudio en el presente apartado.

Se muestran a continuación, de manera análoga a los apartados anteriores, los resultados obtenidos. Se presentan los diferentes cortes de la secuencia VIBE en vacío axial, con el hígado segmentado señalado con el contorno en amarillo y el hígado de referencia con el contorno en verde.



	
Corte 15	Corte 20
	
Corte 25	Corte 30
	
Corte 35	Corte 40

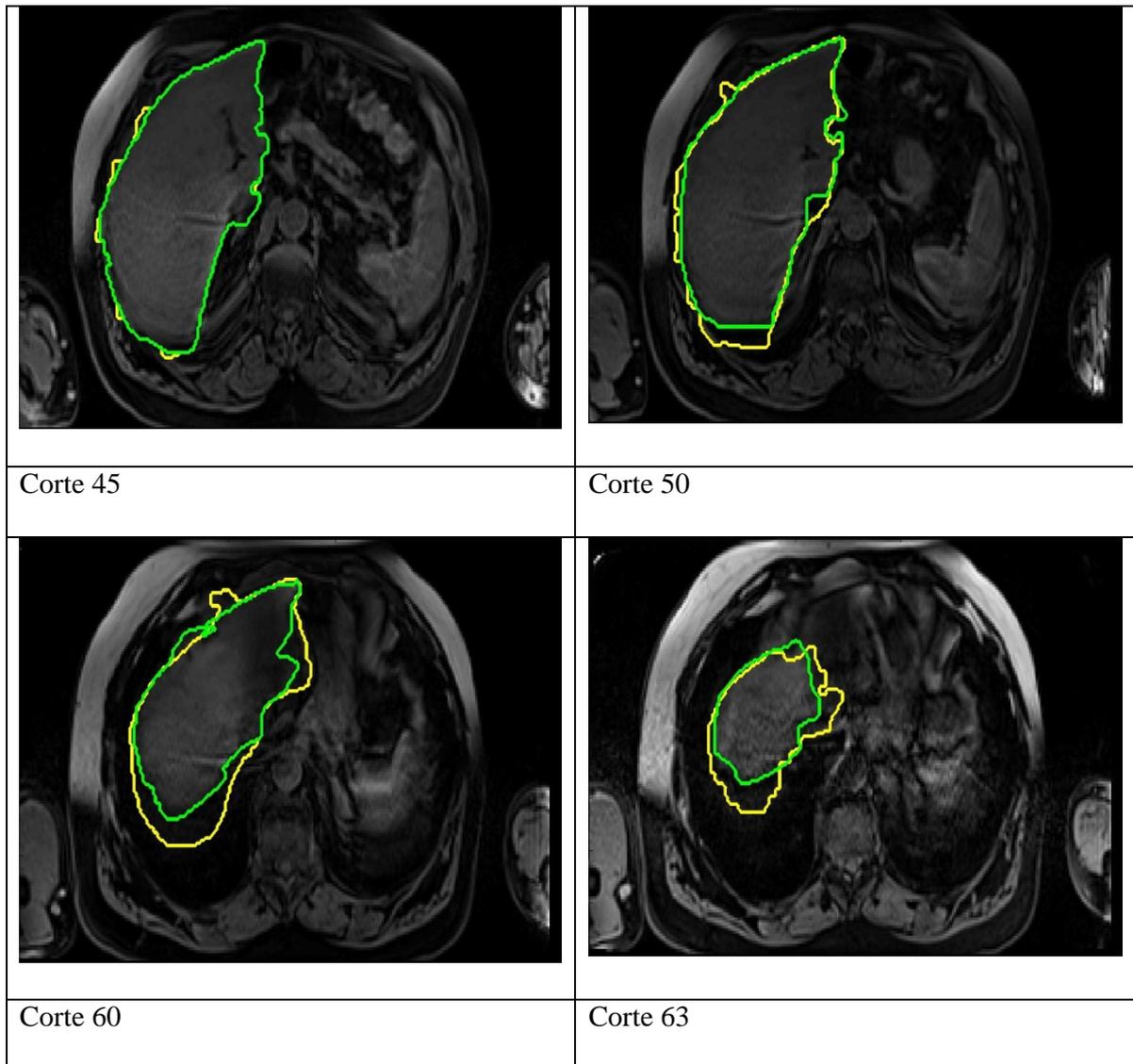


Figura 6.7. Imágenes de diferentes cortes de la secuencia VIBE 05 con hígado segmentado por método de minimización de contorno activo en 3D y entrada multisequencia (amarillo) y contorno de hígado de referencia (verde)

Es necesario comparar de manera rigurosa con el resto de métodos, y por ello, se evalúa de igual manera con las métricas establecidas.

En la tabla 6.8 se muestran los valores de las métricas obtenidas, con respecto a la segmentación de referencia generada manualmente sobre la secuencia VIBE 05. Estos valores son la media de las métricas obtenidas para los 6 estudios disponibles. Como en este apartado se está analizando el caso de entrada multisequencia, se ha validado el método sobre las 63 imágenes de cada una de las 5 secuencias en cada uno de los 6 estudios para pruebas disponibles, es decir 63 imágenes x 6 estudios x 5 secuencias / estudio, lo que hace un total de 1.890 imágenes. Se ha adoptado como valor de la métrica final, como se ha indicado, la media de los valores obtenidos para todos los cortes.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(mm)	(mm)	(mm)	(%)	(%)	(%)	(%)
2.08 ± 1.83	2.11 ± 1.93	0.93 ± 0.77	2.29 ± 1.69	10.72 ± 7.61	99.61 ± 0.12	0.24 ± 0.22	99.80 ± 0.19	98.59 ± 0.92

Tabla 6.9. Métricas finales de los resultados de segmentación por el método de minimización de contorno activo en 3D y entrada multiseuencia

Los valores obtenidos de las métricas son muy buenos. En cuanto al tiempo de computación necesario, el manejo de la información multiseuencia y la consiguiente generación del mapa de distancias con respecto al modelo estadístico multivariable de hígado previamente generado no supone una carga adicional relevante, constituyendo sólo unos 10 segundos que añadir a la segmentación 3D.

Como imagen ilustrativa, se incluye una reconstrucción 3D del hígado segmentado conseguido con los métodos que obtienen una segmentación más precisa, esto es, los que emplean la entrada multiseuencia, tanto en la aproximación 2D como en 3D.

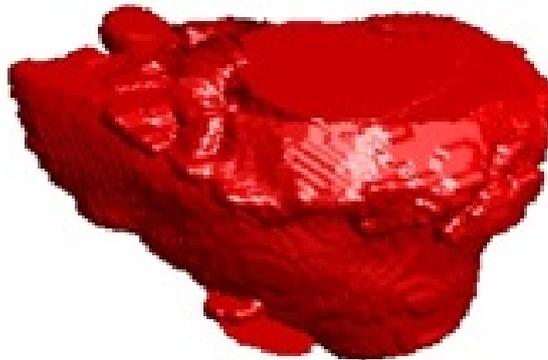


Figura 6.8. Reconstrucción 3D del volumen del hígado segmentado.

Se puede apreciar que la parte superior del hígado es ligeramente plana. Esto es debido a que el escaneo del paciente empieza justo en el límite de comienzo del hígado.

6.5. Análisis de resultados

Una vez se han mostrado los resultados obtenidos de la segmentación tras la aplicación de cada método en 2D y 3D es necesario analizarlos para extraer conclusiones e identificar futuras líneas de trabajo.

Se recoge en la tabla 6.9 una comparativa de las métricas obtenidas para cada uno de los métodos, resultantes del proceso de evaluación previamente descrito, junto con la desviación estándar de los valores en cada caso.

	VOE (%)	RVD (%)	ASD (MM)	RMSD (MM)	MSD (MM)	TPR (%)	FPR (%)	ACC (%)	DSC (%)
2D MONO (sin mapa)	39.38 ± 0.79	41.78 ± 28.66	14.05 ± 1.28	20.32 ± 3.64	60.09 ± 1.69	66.18 ± 2.11	0.22 ± 0.36	96.31 ± 3.06	73.21 ± 7.71
2D MONO (con mapa)	12.41 ± 6.11	7.89 ± 6.31	3.04 ± 1.24	4.93 ± 1.95	19.33 ± 7.36	94.00 ± 5.10	0.51 ± 0.32	98.97 ± 0.65	93.07 ± 3.54
2D MULTI (con mapa)	1.98 ± 1.82	2.01 ± 1.92	0.88 ± 0.73	2.27 ± 1.70	10.59 ± 7.77	99.62 ± 0.11	0.23 ± 0.22	99.81 ± 0.19	98.64 ± 0.92
3D MONO (con mapa)	11.86 ± 4.96	4.83 ± 3.00	3.07 ± 0.99	4.55 ± 1.82	17.91 ± 7.30	94.24 ± 1.90	0.48 ± 0.41	99.05 ± 0.47	93.48 ± 2.86
3D MULTI (con mapa)	2.08 ± 1.83	2.11 ± 1.93	0.93 ± 0.77	2.29 ± 1.69	10.72 ± 7.61	99.61 ± 0.12	0.24 ± 0.22	99.80 ± 0.19	98.59 ± 0.92

Tabla 6.10. Tabla comparativa de métricas para todos los métodos de segmentación y tipos de entradas propuestos

La información entre paréntesis “con mapa” o “sin mapa” hace referencia a la utilización del descriptor compacto propuesto para la generación del mapa de distancias con respecto al modelo estadístico multivariable de hígado previamente generado con el doble objetivo de manejar y combinar la información existente en las múltiples secuencias, en caso de entrada multisequencia, y al mismo tiempo, realizar un aumento del contraste de la imagen o volumen de entrada al método de minimización, para ambos tipos de entrada. Como se ha mostrado anteriormente, también se ha aplicado el método de minimización del contorno

activo sobre las imágenes de la secuencia VIBE tal cual son adquiridas por la máquina de resonancia magnética, sin obtener dicho mapa de distancias.

Como primera conclusión del análisis de la tabla 6.9 se extrae que los 4 métodos que generan el mapa de distancias mediante la aplicación del descriptor compacto propuesto proporcionan buenos resultados, siendo mejores aquellos que han contado con la entrada multisequencia. Esto es lógico porque ya se disponía del conocimiento de que las distintas secuencias contenían información complementaria que ayudaba al radiólogo en su tarea de diagnóstico y que se realizaba de manera manual en algunas ocasiones en la práctica clínica. Este hecho permite validar la bondad del descriptor compacto que se ha propuesto en esta tesis.

Este descriptor contiene la información de las 5 secuencias y la ha transformado en un mapa de distancias al modelo de hígado, o lo que es lo mismo, un mapa de probabilidad de pertenencia a la clase hígado basándose en un modelo estadístico multivariable previo del hígado. La imagen de entrada resultante que combina la información ya ofrece una mejora de contraste cualitativa. Tras la realización de las pruebas se valida que esta apreciación visual de mejora de contraste se traduce en una mejora notable en la precisión en la segmentación del hígado. Además, como se ha comentado anteriormente, esta manera de representar la información multisequencia permite reducir la dimensión del problema en 5 órdenes de magnitud, lo cual hace posible la aplicación de modelos deformables sobre entradas multisequencia. De otra manera, tendría un tiempo de cómputo y necesidad de unos recursos de memoria extremadamente elevados y probablemente, encontrar en este espacio de 7 dimensiones el mínimo global y por tanto, la segmentación óptima, sería más complejo.

Otra conclusión importante que se puede extraer de la observación de los resultados, es que los métodos de segmentación tanto en dos como tres dimensiones tienen un comportamiento muy estable en todos los cortes de la secuencia, ya que la desviación estándar de los valores de cada métrica obtenidos sobre todos los cortes son muy bajos, revelando un comportamiento en la segmentación conservado en todos los puntos, lo cual también es un factor importante.

La extensión del método de 2D a 3D no ha proporcionado los resultados que se esperaban a priori. Se estimaba que la información existente en el eje z proporcionaría datos de continuidad en este tercer eje del espacio, y podría mejorar notablemente los resultados con respecto al 2D. El método extendido a 3D, y con entrada monosequencia, proporciona unos resultados ligeramente mejores que en 2D, y algo más estables, ya que la desviación típica es menor. En el caso de la extensión del método a 3D y entrada multisequencia, la extensión a 3D incluso proporciona resultados ligeramente peores. Este comportamiento se analiza en detalle en el apartado 6.5.2.

6.5.1. Contribución del descriptor compacto a los resultados de la segmentación

El descriptor compacto propuesto, como ya se ha indicado previamente, realiza una contribución doble en la mejora de los resultados del proceso de segmentación. En la entrada monosecuencia, el descriptor aumenta el contraste de la imagen en relación a la imagen de grises de la secuencia VIBE en vacío axial (la empleada como entrada monosecuencia en todos los casos), realzando los bordes de los distintos órganos lo que permite que el contorno activo converja mejor y se pegue a los bordes, proporcionando unos mejores resultados de segmentación.

En la entrada multisecuencia, el descriptor compacto posibilita combinar la información existente en todas las secuencias al mismo tiempo que realiza esta mejora de contraste entre tejidos ya mencionada. Este doble efecto se traduce en una mejora sustancial de los resultados de la segmentación.

La tabla 6.10 recoge las métricas obtenidas para la segmentación en 2D sobre los 3 tipos de entrada: a) la secuencia VIBE en vacío axial tal cual, b) el mapa de distancias obtenido como resultado de la aplicación del descriptor compacto sobre la secuencia VIBE en vacío axial; c) el mapa de distancias obtenido como resultado de la aplicación del descriptor compacto sobre la entrada multisecuencia.

	VOE (%)	RVD (%)	ASD (MM)	RMSD (MM)	MSD (MM)	TPR (%)	FPR (%)	ACC (%)	DSC (%)
2D MONO (sin mapa)	39.38 ± 0.79	41.78 ± 28.66	14.05 ± 1.28	20.32 ± 3.64	60.09 ± 1.69	66.18 ± 2.11	0.22 ± 0.36	96.31 ± 3.06	73.21 ± 7.71
2D MONO (con mapa)	12.41 ± 6.11	7.89 ± 6.31	3.04 ± 1.24	4.93 ± 1.95	19.33 ± 7.36	94.00 ± 5.10	0.51 ± 0.32	98.97 ± 0.65	93.07 ± 3.54
2D MULTI (con mapa)	1.98 ± 1.82	2.01 ± 1.92	0.88 ± 0.73	2.27 ± 1.70	10.59 ± 7.77	99.62 ± 0.11	0.23 ± 0.22	99.81 ± 0.19	98.64 ± 0.92

Tabla 6.11. Tabla comparativa de métricas para los métodos de segmentación en 2D y tipos de entradas con aplicación o no del descriptor compacto

Las métricas son claramente mejores si se aplica el descriptor compacto sobre la entrada. Para ilustrar este efecto, se ha recogido en las figuras 6.9 y 6.10 una comparativa de los resultados de la segmentación en varios cortes diferentes de una misma secuencia sobre los que se ha utilizado diferentes tipos de entrada:

Cada fila de las figuras 6.9 y 6.10 representa un corte de la secuencia sobre el que se ha aplicado el mismo método de segmentación en 2D y donde solo varía el tipo de entrada. En la primera columna (a) la entrada es tal cual se recibe de la máquina de resonancia, no se aplica el descriptor propuesto. En la segunda columna (b) se aplica el descriptor propuesto sobre la entrada monosecuencia, y en la tercera columna (c) se aplica el descriptor propuesto sobre la entrada multisecuencia.

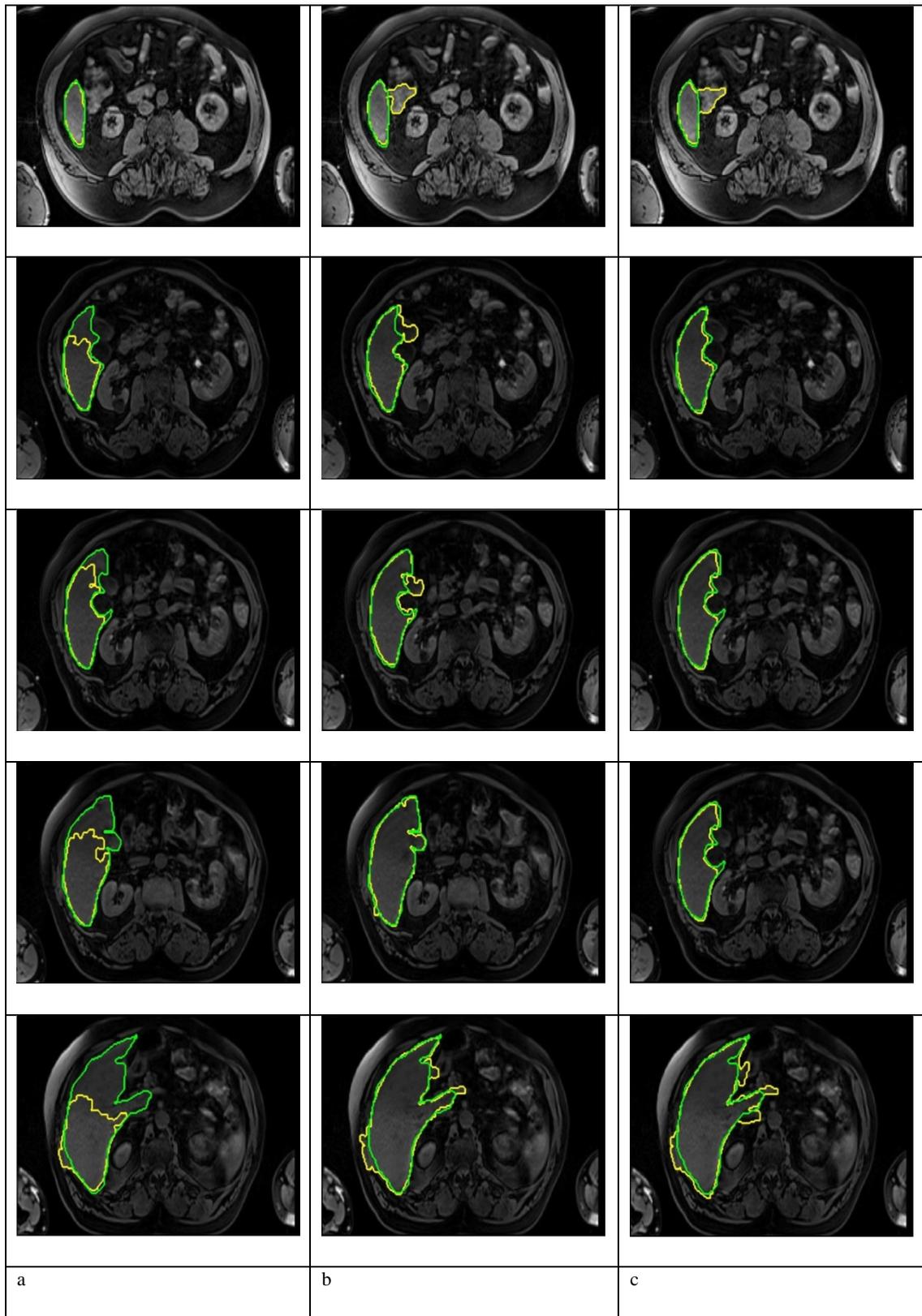


Figura 6.9. Comparativa (I) de los resultados obtenidos en 2D sobre entrada monosecuencia en diferentes cortes y con diferentes tipos de entrada. a) segmentación en 2D sobre entrada mapa de grises secuencia VIBE en vacío axial en 2D; b) segmentación en 2D sobre mapa de distancias tras aplicación del descriptor sobre entrada monosecuencia; c) segmentación en 2D sobre mapa de distancias tras aplicación del descriptor sobre entrada multisecuencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

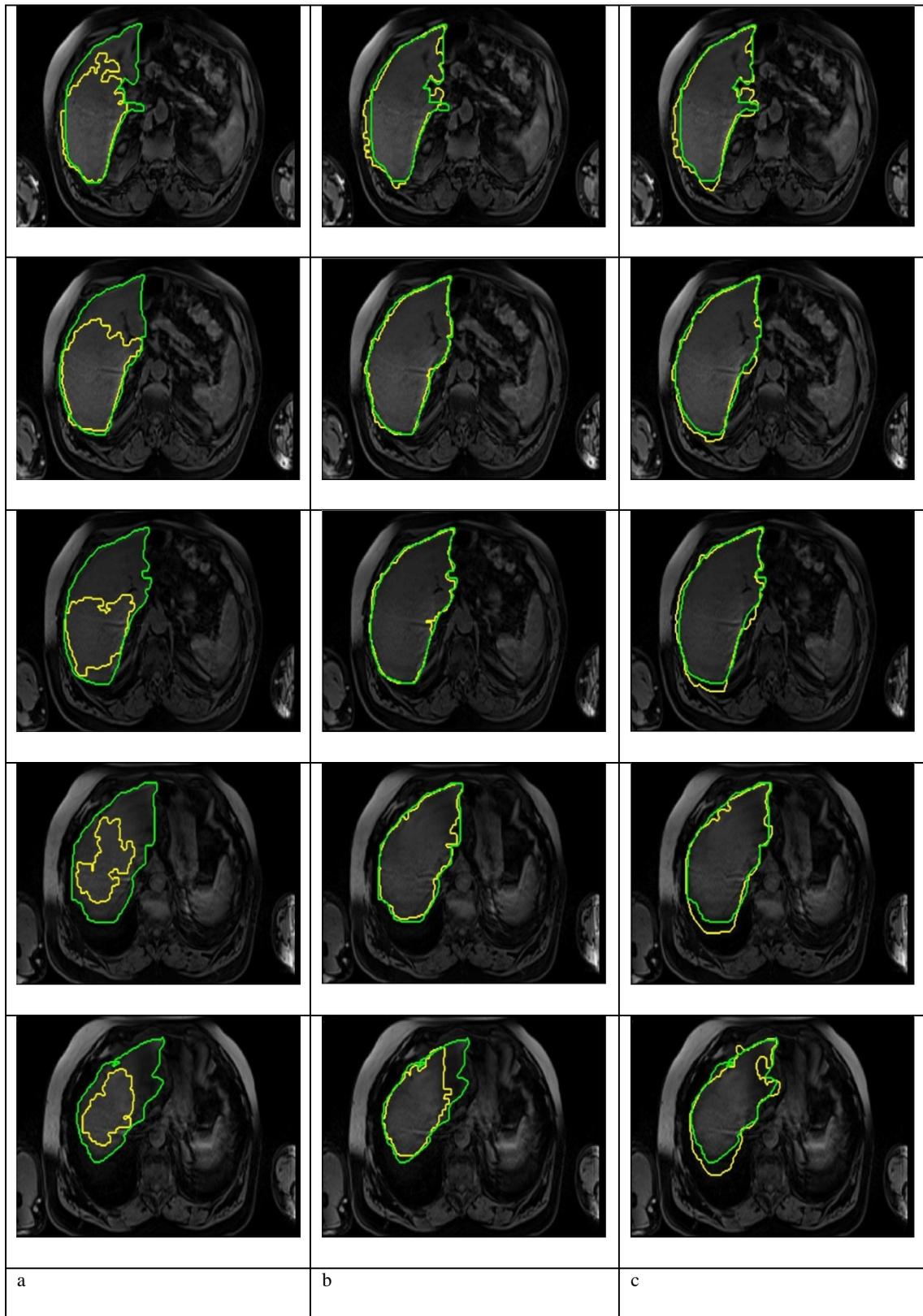


Figura 6.10. Comparativa (II) de los resultados obtenidos en 2D sobre entrada monosecuencia en diferentes cortes y con diferentes tipos de entrada. a) segmentación en 2D sobre entrada mapa de grises secuencia VIBE en vacío axial en 2D; b) segmentación en 2D sobre mapa de distancias tras aplicación del descriptor sobre entrada monosecuencia; c) segmentación en 2D sobre mapa de distancias tras aplicación del descriptor sobre entrada multisecuencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

En las figuras anteriores se han expuesto los resultados de la segmentación. El hígado segmentado automáticamente ha sido señalado con contorno en amarillo. El contorno verde se corresponde con el hígado de la segmentación de referencia.

De la observación de los resultados se pueden concluir los efectos que el descriptor compacto proporciona en los resultados de segmentación.

El descriptor compacto aplicado sobre la entrada monosecuencia mejora el contraste de la imagen y permite una mejor localización del hígado. Este efecto se confirma en todos los casos mostrados de manera notable (columnas a) y b)). El hecho de que la escena no aparezca uniformemente iluminada en la imagen original, con ligero aumento de contraste en su mitad inferior, hace que el contorno activo no delimite correctamente los límites del hígado y segmente la mitad de su región. El mapa de distancias generado a partir de la aplicación del descriptor propuesto corrige parcialmente este efecto, y mejora el contraste, permitiendo una segmentación más precisa del contorno del hígado.

El descriptor compacto aplicado sobre la entrada multisecuencia, además de manejar la información de los múltiples canales y no aumentar la dimensión del problema a resolver, mejora también el contraste de la nueva entrada, esto es, el mapa de distancias generado. En algunos ejemplos se produce una mejora notable en los resultados de la segmentación, como en el corte mostrado en las filas 2, 3 y 4. Se añade en la siguiente figura un ejemplo claro de mejora.

En la figura 6.11, se muestra el mismo corte en Fig.11a) y en la Fig.11b) con el hígado segmentado después de la ejecución del método en 2D sobre la entrada monosecuencia y la entrada multisecuencia, respectivamente. Sobre ambas entradas se había aplicado el descriptor propuesto. Se puede apreciar que la segmentación de la derecha no incluye la vesícula en la región segmentada, mientras que la segmentación de la izquierda no consigue discriminar ambos tejidos.

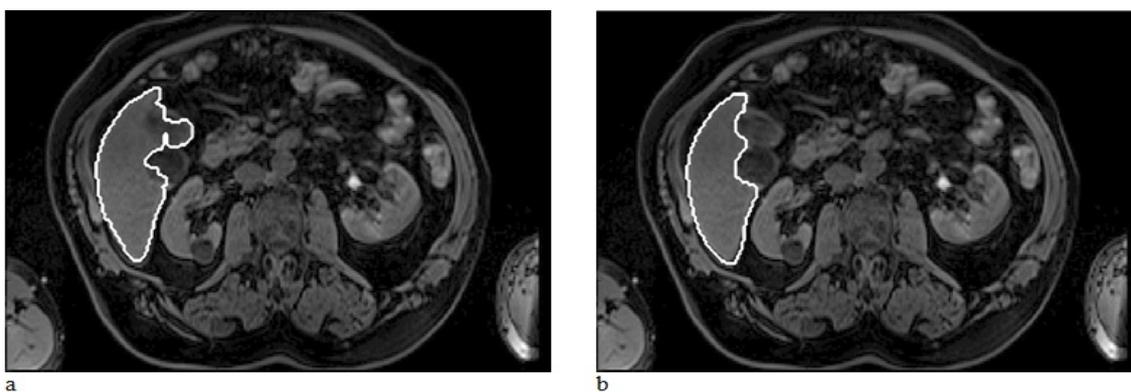


Figura 6.11. a) Hígado segmentado en un corte de la secuencia VIBE 05 en vacío axial mediante contorno activo minimizado en 2D y entrada monosecuencia; b) Hígado segmentado en el mismo corte de la misma secuencia mediante contorno activo minimizado en 2D y entrada multisecuencia

Esto confirma el hecho de que la información mutisequencia proporciona información complementaria para el procesamiento de las imágenes y el diagnóstico, lo que ya sabíamos por la información clínica recibida. Lo que se confirma con esta prueba es que el descriptor compacto propuesto para tratar y procesar automáticamente esta información múltiple es valioso y eficiente en cuanto a contribuir a unos mejores resultados de la segmentación.

Sin embargo, cabe añadir que el uso de la entrada mutisequencia también presenta algunos inconvenientes, como que añade un error debido a los problemas de registro no corregidos adecuadamente. Este efecto puede apreciarse en las filas 3,4 y 5 de la figura 6.10, donde el contorno de hígado segmentado ha sido ligeramente ampliado hacia abajo, empeorando el resultado. Claramente ha considerado el valor de alguna de las secuencias que no haya quedado tan bien registrada con respecto a la VIBE en vacío axial.

6.5.2. Contribución de la extensión a 3D a los resultados de la segmentación

La segmentación realizada sobre un volumen (3D) en lugar de sobre una imagen (2D) permite incorporar un eje z adicional, que proporciona información sobre la continuidad del elemento segmentado en el tercer eje del espacio. A priori, este hecho permite segmentaciones más precisas que eliminen picos y suavicen la superficie segmentada.

Se va a analizar la contribución de la extensión a 3D para la entrada monosequencia y la entrada mutisequencia. La tabla 6.11 recoge las métricas obtenidas para la segmentación en 2D y 3D sobre la entrada monosequencia, esto es, el mapa de distancias obtenido como resultado de la aplicación del descriptor compacto sobre la secuencia VIBE en vacío axial.

	VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
	(%)	(%)	(MM)	(MM)	(MM)	(%)	(%)	(%)	(%)
2D									
MONO	12.41	7.89 ±	3.04 ±	4.93 ±	19.33	94.00	0.51 ±	98.97	93.07
(con	± 6.11	6.31	1.24	1.95	± 7.36	± 5.10	0.32	± 0.65	± 3.54
mapa)									
3D									
MONO	11.86	4.83 ±	3.07 ±	4.55 ±	17.91	94.24	0.48 ±	99.05	93.48
(con	± 4.96	3.00	0.99	1.82	± 7.30	± 1.90	0.41	± 0.47	± 2.86
mapa)									

Tabla 6.12. Tabla comparativa de métricas para los métodos de segmentación en 2D y 3D sobre entrada monosequencia

Se ilustra este comportamiento en las figuras 6.12 y 6.13 para la entrada monosecuencia, y en las figuras 6.14 y 6.15 para la entrada multisecuencia.

En dichas figuras se han expuesto los resultados de la segmentación. Cada fila se corresponde con un corte de la secuencia, que ha sido segmentado por diferentes métodos. En todos los casos se ha calculado el descriptor propuesto para la generación del mapa de distancias que constituye la nueva entrada al método de segmentación. Para su visualización se han superpuesto los contornos del hígado sobre los cortes de la secuencia VIBE 05 en vacío axial, que es la referencia para la segmentación manual. El hígado segmentado automáticamente ha sido señalado con contorno en amarillo. El contorno verde se corresponde con el hígado de la segmentación de referencia.

En la primera columna (a) de las figuras 6.12 y 6.13 se muestran los resultados de la segmentación en 2D sobre la entrada monosecuencia. En la segunda columna (b) se muestran los resultados de la segmentación sobre el mismo corte de la secuencia en la aproximación 3D, también sobre la entrada monosecuencia.

De la observación de las imágenes en ambas tablas se puede concluir la contribución que la extensión a 3D proporciona en los resultados de la segmentación sobre la entrada monosecuencia. En general, se aprecia que los contornos se han regularizado, se suavizan, eliminándose las formas irregulares. La mejora en 3D en la delimitación de la región del hígado es poca, ya de por sí buena en 2D, pero es mejor al ser los contornos más suaves y regulares, los cuales se corresponden más con la forma natural del hígado.

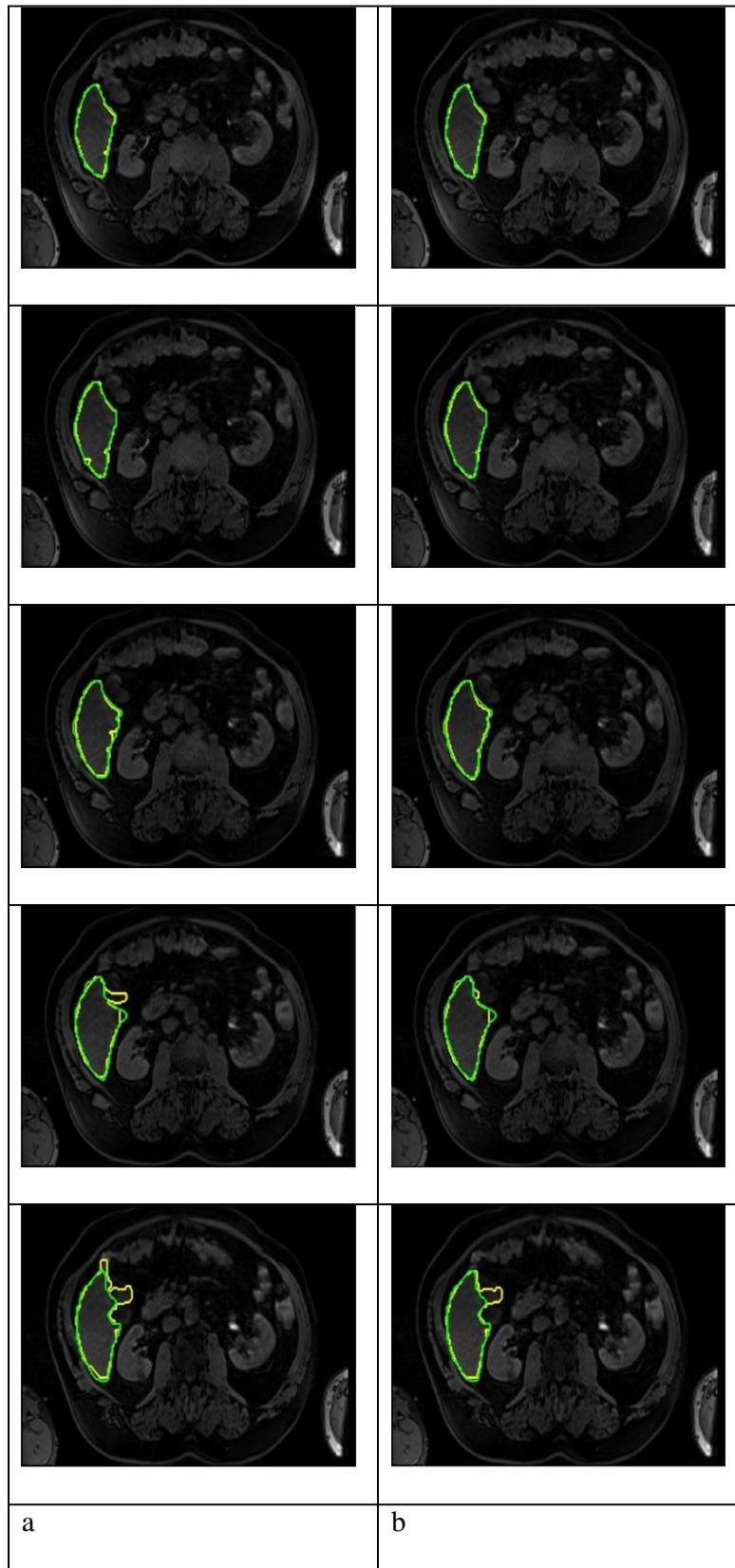


Figura 6.12. Comparativa (I) de los resultados obtenidos en 2D y 3D sobre entrada monosecuencia en diferentes cortes. a) segmentación en 2D sobre entrada monosecuencia; b) segmentación en 3D sobre entrada monosecuencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

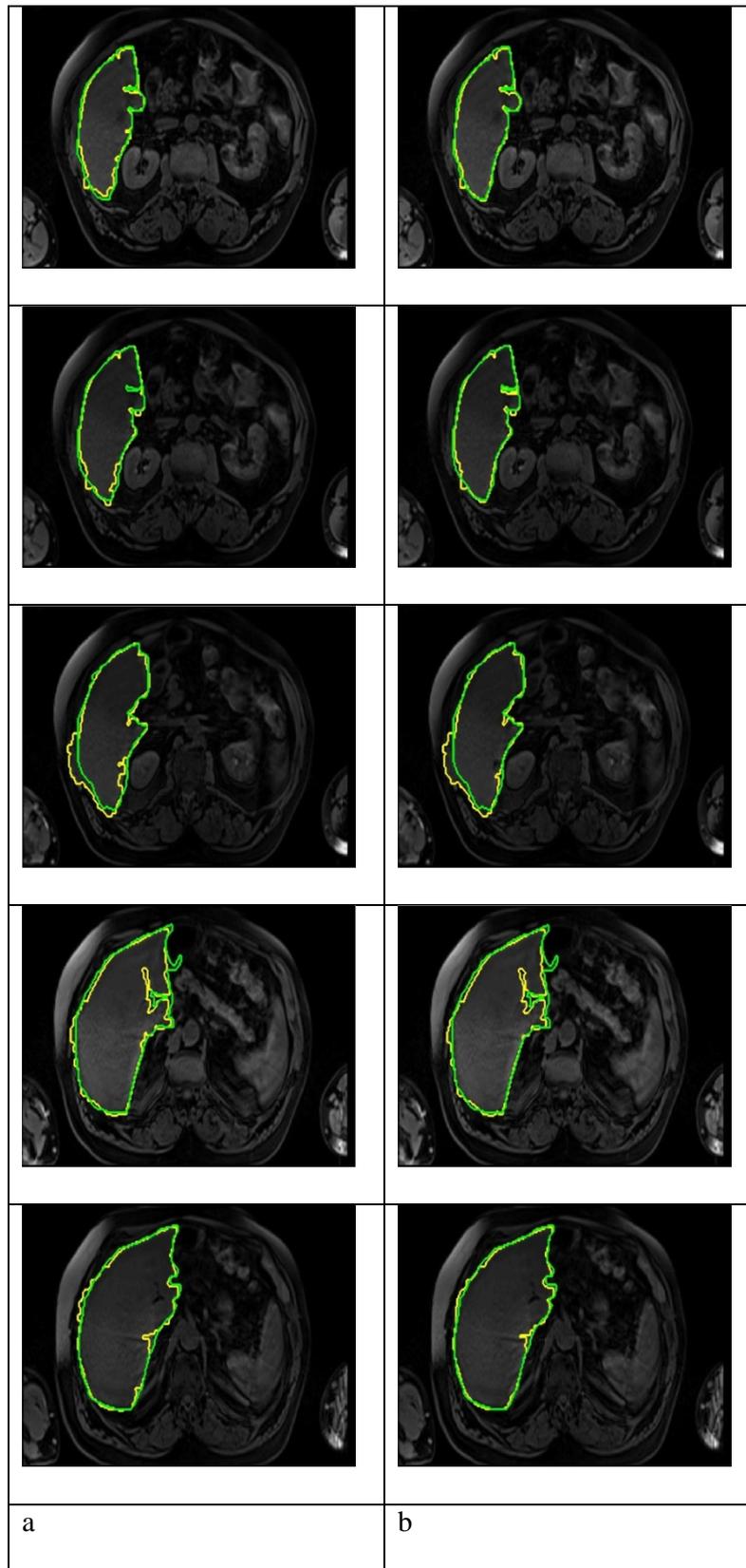


Figura 6.13. Comparativa (II) de los resultados obtenidos en 2D y 3D sobre entrada monosecuencia en diferentes cortes. a) segmentación en 2D sobre entrada monosecuencia; b) segmentación en 3D sobre entrada monosecuencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

La tabla 6.12 recoge las métricas obtenidas para la segmentación en 2D y 3D sobre la entrada multisequencia, esto es, el mapa de distancias obtenido como resultado de la aplicación del descriptor compacto sobre la entrada multisequencia.

	VOE (%)	RVD (%)	ASD (MM)	RMSD (MM)	MSD (MM)	TPR (%)	FPR (%)	ACC (%)	DSC (%)
2D MULTI (con mapa)	1.98 ± 1.82	2.01 ± 1.92	0.88 ± 0.73	2.27 ± 1.70	10.59 ± 7.77	99.62 ± 0.11	0.23 ± 0.22	99.81 ± 0.19	98.64 ± 0.92
3D MULTI (con mapa)	2.08 ± 1.83	2.11 ± 1.93	0.93 ± 0.77	2.29 ± 1.69	10.72 ± 7.61	99.61 ± 0.12	0.24 ± 0.22	99.80 ± 0.19	98.59 ± 0.92

Tabla 6.13. Tabla comparativa de métricas para los métodos de segmentación en 2D y 3D sobre entrada multisequencia

En las figuras 6.14 y 6.15 se muestra la contribución de la extensión a 3D del método de segmentación sobre la entrada multisequencia. De manera análoga, cada fila muestra el mismo corte de la secuencia con los resultados de la segmentación automática (contorno amarillo) y la segmentación de referencia (contorno verde). En la primera columna (a) de las figuras 6.13 y 6.14 se muestran los resultados de la segmentación en 2D sobre la entrada multisequencia. En la segunda columna (b) se muestran los resultados de la segmentación sobre el mismo corte de la secuencia en la aproximación 3D, también sobre la entrada multisequencia.

En este caso, la segmentación en 3D sobre la entrada multisequencia no proporciona una gran mejora. Se observa cierto suavizado en los contornos segmentados, de igual manera a como se ha indicado en el caso anterior. Si bien el descriptor compacto permite manejar la información de los múltiples canales, aumenta el contraste en la nueva imagen de mapa de distancias generada, y permite la mejora en la delimitación del hígado con respecto a la entrada monosequencia (como se ha ilustrado en el apartado anterior), las limitaciones existentes con el registro de las diferentes secuencias aporta un ligero ruido en la segmentación e imposibilita que la precisión de los resultados aumente. De hecho, el contorno segmentado es agrandado en algunas zonas, debido a los ligeros desajustes en el registro de las secuencias. Este hecho se puede observar en las últimas 3 filas de la figura 6.15.

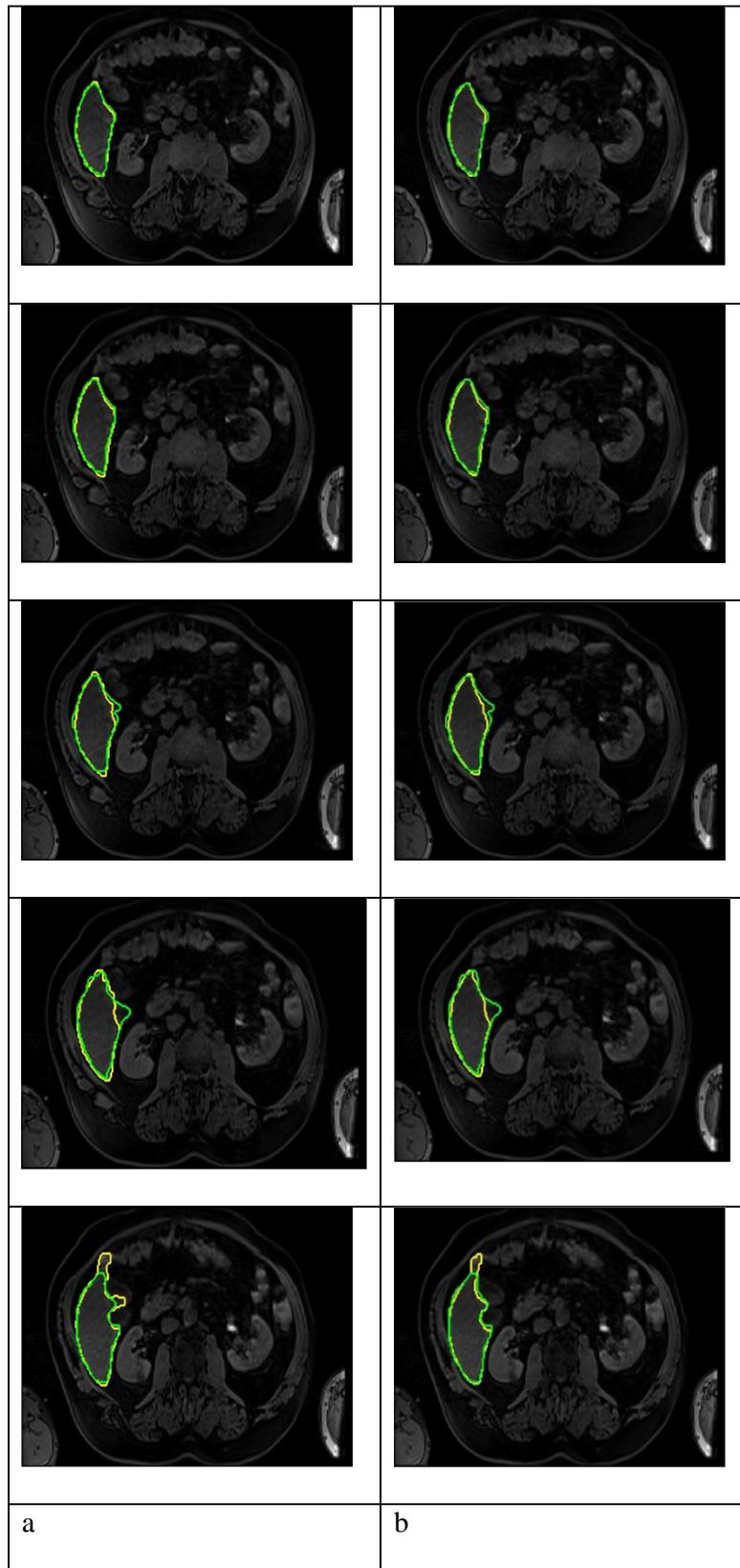


Figura 6.14. Comparativa (I) de los resultados obtenidos en 2D y 3D para entrada multisequencia en diferentes cortes. a) segmentación en 2D sobre entrada multisequencia; b) segmentación en 3D sobre entrada multisequencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

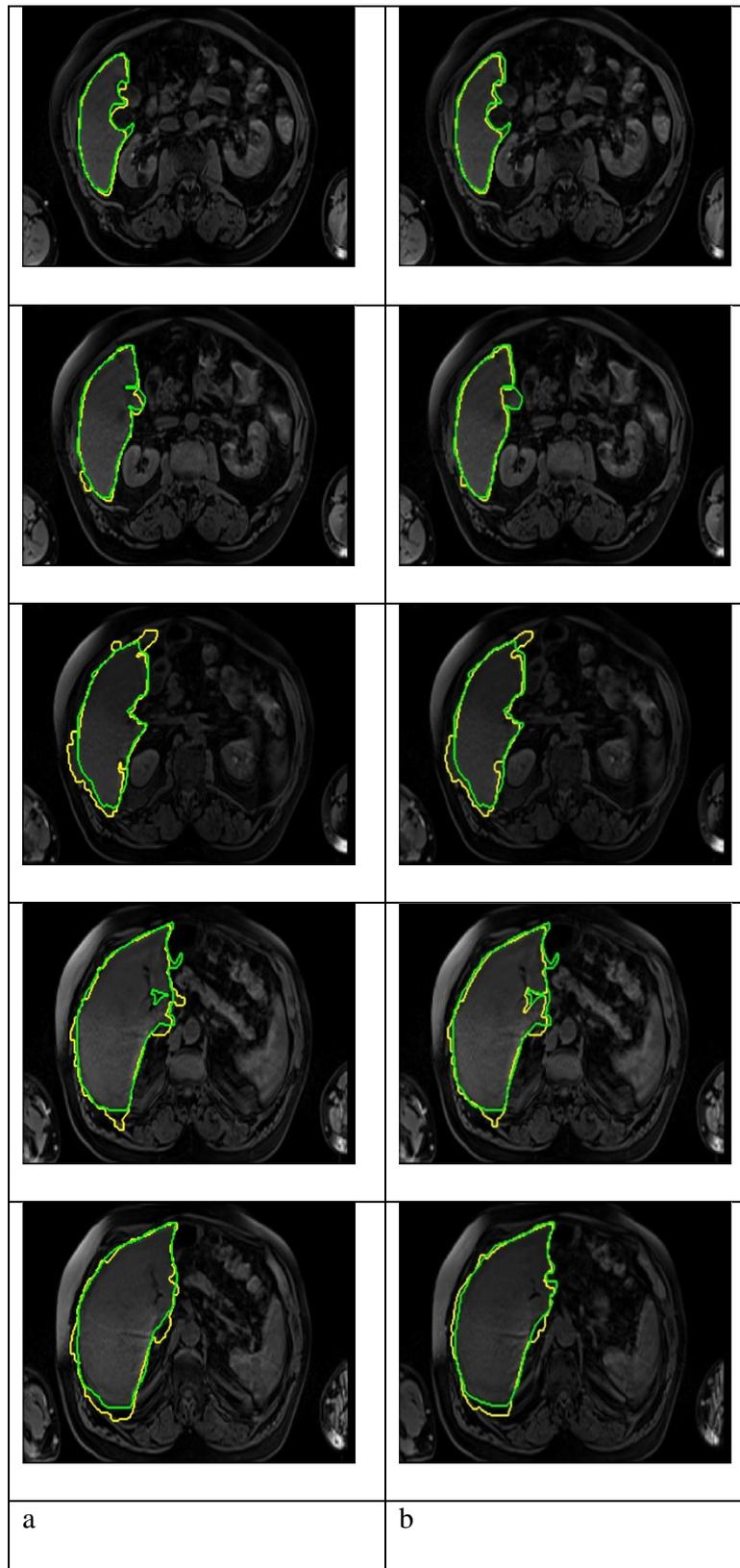


Figura 6.15. Comparativa (II) de los resultados obtenidos en 2D y 3D para entrada multisequencia en diferentes cortes. a) segmentación en 2D sobre entrada multisequencia; b) segmentación en 3D sobre entrada multisequencia. El contorno amarillo señala el hígado segmentado y el contorno verde representa el hígado de referencia.

La influencia del problema de registro en los resultados de la segmentación ha sido evaluada. Se obtuvieron los valores de las métricas para comparar las posiciones iniciales entre la secuencia T2 BLADE y el grupo de secuencias VIBE. Para ello se realizó una segmentación manual de la secuencia T2 BLADE y se comparó con la segmentación de referencia utilizada en toda la fase de validación. Las diferencias existentes de alineamiento, es decir, registro entre ambas se recogen en la siguiente tabla.

VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(%)	(%)	(MM)	(MM)	(MM)	(%)	(%)	(%)	(%)
19.19 ± 6.49	9.40 ± 8.02	5.07 ± 1.13	6.57 ± 1.86	21.57 ± 9.52	92.99 ± 2.51	1.12± 0.48	98.45 ± 0.61	89.18 ± 4.01

Tabla 6.14. Tabla de métricas obtenidas para comparación del alineamiento (registro) inicial entre la secuencia T2 BLADE y la secuencia VIBE.

Los resultados de la segmentación en 3D sobre la entrada multisequencia con las secuencias después de aplicar el método de registro y sin aplicarlo se recoge en la tabla 6.14.

	VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
	(%)	(%)	(MM)	(MM)	(MM)	(%)	(%)	(%)	(%)
3D MULTI (sin registro)	5.12 ± 10.14	7.14 ± 17.49	1.46 ± 1.86	3.31 ± 3.72	13.83 ±12.50	99.60 ± 0.12	0.33 ± 0.37	99.72 ± 0.35	96.71 ± 6.39
3D MULTI (con registro)	2.08 ± 1.83	2.11 ± 1.93	0.93 ± 0.77	2.29 ± 1.69	10.72 ± 7.61	99.61 ± 0.12	0.24 ± 0.22	99.80 ± 0.19	98.59 ± 0.92

Tabla 6.15. Métricas obtenidas para la segmentación en 3D sobre entrada multisequencia con las secuencias antes y después de aplicar el método de registro.

Se verifica que el buen registro de los múltiples canales de entrada es muy relevante para obtener buenos resultados en la segmentación.

6.6. Comparación con el estado del arte

Para poder valorar si los resultados obtenidos son buenos o no, es necesario revisar el estado del arte. Esta comparación presenta bastante dificultad ya que muchas publicaciones no muestran resultados de los métodos de segmentación de manera rigurosa, sino que la mayoría de las veces se muestran imágenes aisladas, lo cual puede dar una vista muy parcial de lo que pasa en conjunto. Además, se suelen emplear 1 o 2 métricas, y cada estudio las elige arbitrariamente.

Otra dificultad existente para comparar los resultados con otros estudios previos es que la naturaleza de las imágenes es distinta. En esta tesis se ha trabajado sobre imágenes de resonancia magnética, de MRI, mientras que la mayor parte de los estudios presentes en la literatura, son sobre imágenes de tomografía axial computerizada, CT. Este hecho hace que no sean estrictamente comparables dado que las imágenes de CT son distintas que las de MRI, presentan mucho menos ruido y mucho más contraste entre los órganos presentes en la imagen, con lo que las condiciones son más ventajosas para que los algoritmos de segmentación proporcionen mejores resultados sobre ellas. Una comparación estricta exige comparar imágenes de MRI con imágenes de MRI, y específicamente, con métodos automáticos que trabajen sobre MRI para segmentar la región del hígado.

Un problema adicional para realizar una comparativa con diferentes métodos del estado del arte es que no existen bases de datos públicas con imágenes MRI abdominales con las que validar los algoritmos. Es habitual que cada grupo de investigación trabaje en colaboración con un hospital o servicio de radiología que es quien le proporciona el conjunto de datos de prueba y validación. Cada conjunto de datos es propio y son diferentes entre sí, y en consecuencia la segmentación de referencia generada sobre ese conjunto de datos también es propia y realizada por diferentes expertos, añadiendo además la variabilidad en la segmentación manual dependiente del radiólogo que la realiza. En conclusión, la comparativa que se puede hacer con otros métodos del estado del arte es aproximada.

Dado que los algoritmos presentados en esta tesis pertenecen al conjunto de los considerados como “automáticos”, ya que no se precisa de ninguna interacción por parte del usuario, sus resultados serían comparables a los métodos que se exponen a continuación en la tabla 6.14. Estos métodos han sido descritos en el capítulo 4. Todos ellos son métodos automáticos de segmentación de hígado sobre imágenes abdominales de MRI.

	VOE	RVD	ASD	RMSD	MSD	TPR	FPR	ACC	DSC
(Reska et al., 2012)	12.10	-6.89	3.07	-	-	-	-	-	-
(Yuan et al., 2010)	7.0	4.0	-	2.6	-	-	-	-	-
(Huynh et al., 2014)	-	-	-	-	-	-	-	99.4 ± 0.14	93.6 ± 1.7
(Oh et al., 2014)	-	-	-	-	-	92.5 ± 3.1	4.5 ± 3.2	-	93.9 ± 1.6
(Masoumi et al., 2012)	-	-	-	-	-	-	-	94.0	-
(Chen et al., 2009)	-	-	-	-	-	93.2 ± 5.3	0.94 ± 0.04	-	-
(López-Mir et al., 2014)	-	-	-	-	33.58 ± 6.1	-	-	-	95.0 ± 1.0
(Göçeri et al., 2014)	-	-	-	-	-	-	-	96.0	-
(Gloger et al., 2010)	-	-	-	-	20.35 ± 8.66	-	-	-	94.0 ± 2.0
2D MONO	12.41 ± 6.11	7.89 ± 6.31	3.04 ± 1.24	4.93 ± 1.95	19.33 ± 7.36	94.00 ± 5.10	0.51 ± 0.32	98.97 ± 0.65	93.07 ± 3.54
3D MONO	11.86 ± 4.96	4.83 ± 3.00	3.07 ± 0.99	4.55 ± 1.82	17.91 ± 7.30	94.24 ± 1.90	0.48 ± 0.41	99.05 ± 0.47	93.48 ± 2.86
2D MULTI	1.98 ± 1.82	2.01 ± 1.92	0.88 ± 0.73	2.27 ± 1.70	10.59 ± 7.77	99.62 ± 0.11	0.23 ± 0.22	99.81 ± 0.19	98.64 ± 0.92
3D MULTI	2.08 ± 1.83	2.11 ± 1.93	0.93 ± 0.77	2.29 ± 1.69	10.72 ± 7.61	99.61 ± 0.12	0.24 ± 0.22	99.80 ± 0.19	98.59 ± 0.92

Tabla 6.16. Tabla comparativa de métricas obtenidas con nuestro método y otros métodos de segmentación del estado del arte

En la tabla anterior se ha mostrado una extensa comparación de las métricas obtenidas por los métodos implementados en esta tesis y las métricas de los trabajos comparables del estado del arte. Si bien los resultados de la segmentación obtenidos por el método tanto en 2D como 3D para entrada monosecuencia pueden ser comparables a la gran mayoría de los métodos, las métricas obtenidas para la entrada multisecuencia con la propuesta de descriptor compacto generado a partir de las 5 secuencias disponibles aumenta notablemente la calidad de la segmentación. Esto sucede tanto en la aproximación 2D como 3D. De la observación de los resultados se concluye que los métodos presentados en esta tesis están alineados e incluso superan los métodos del estado del arte, identificándose como el más cercano al método de (Huynh et al., 2014) que obtiene unos valores de precisión similares en el proceso de segmentación. Se recuerda que la comparación es aproximada ya que el conjunto de datos para prueba y validación es distinto en todos los casos así como la segmentación de referencia generada por diferentes radiólogos.

Otro factor interesante de comparación es el tiempo de cómputo entre los diferentes métodos. De igual manera a la comparativa realizada en cuanto a la calidad de los resultados, el análisis comparativo del tiempo de cómputo de los métodos propuestos a los que se hace referencia en el estado del arte presenta dificultades debido principalmente al hardware utilizado y los conjuntos de datos de prueba y validación que varían en los distintos trabajos.

El método propuesto en esta tesis lleva a cabo la segmentación de un conjunto de 63 imágenes de 160 x 256 píxeles de tamaño en unos $40,0 \pm 2,5$ segundos en un ordenador Dell Optiplex 780, Intel core 2 quad CPU Q9650 con procesador de 3GHz y 64 bits.

Huynh et al. refieren la ejecución de su método en $1,03 \pm 0,13$ min en un PC Intel Xeon 2.66 GHz. En él utilizan bases de datos de alrededor de 88 cortes o imágenes de diferentes tamaños, 256x256, 384x385, 512x512 píxeles.

López-Mir [14] utiliza 16 conjuntos de datos que contienen entre 76 y 104 imágenes cada uno de 512x512 píxeles de tamaño. Su método necesita 7-11 min para extraer el volumen del hígado. No refiere las características del PC utilizado.

Oh et al. utilizan un CPU Xeon X5355 de Intel @ 2.66 GHz con 16 GB de memoria principal en un sistema operativo Linux. El proceso de segmentación necesitó 13 minutos para un conjunto de datos de 128 imágenes de tamaño 256x256 píxeles. Masoumi et al. emplean imágenes de tamaño 256 x 256 píxeles y no proporcionan ninguna información sobre el tiempo de cálculo.

Finalmente, Gloger et al. realizaron pruebas sobre los conjuntos de datos de hígados sanos y con grasa. El análisis se llevó a cabo en un procesador de 1,8 GHz Intel Core 2 Duo con 3 GB de RAM, y en tiempo de ejecución de 11.22 ± 2.78 min para hígados sanos y de 15.37 ± 4.96 min para hígados grasos. El tamaño del conjunto de datos es de 64 imágenes por secuencia de tamaño 176 x 256 píxeles.

Tras esta revisión general del tiempo de cómputo necesario en los métodos referidos del estado del arte indica que el que precisa de un tiempo menor de procesamiento para segmentar todo el hígado parece ser Huynh, pero incluso si la comparación no es totalmente real, nuestro método propuesto puede ser considerado más rápido que la media y en la línea del más rápido, el propuesto por Huynh et al.

6.7. Conclusiones

En esta tesis se propone un nuevo método para la segmentación automática del hígado en imágenes de resonancia magnética, tanto en 2D como en 3D y para entrada monosecuencia y multisecuencia. El enfoque propuesto extiende un modelo de contorno activo 2D existente a una superficie activa 3D. El proceso de minimización del funcional de energía descrito por el contorno activo y la superficie activa se realiza por medio de la aproximación dual al problema variacional propuesto por Chambolle, que también se ha extendido en el espacio 3D. La superficie activa 3D propuesta integra de manera natural la regularización volumétrica a través del modelo.

Otra de las aportaciones del trabajo es la representación de los datos multisecuencia por medio de un descriptor compacto único a través de un modelo estadístico multivariable del hígado generado previamente. Este descriptor compacto es capaz de combinar la variabilidad de la información visual y de textura multicanal. Este enfoque evita el aumento de la dimensionalidad del problema de segmentación y permite la incorporación de la información multisecuencia a través del descriptor compacto, lo que hace posible una segmentación más precisa y rápida.

La extensión a 3D del modelo propuesto de contorno activo y su proceso de minimización junto con la aplicación del descriptor compacto para la generación del mapa de distancias de entrada proporciona resultados más precisos que en la segmentación 2D sobre una sola entrada. Además, la inclusión de varias secuencias de MRI y su combinación mediante el descriptor compacto propuesto para generar el mapa de distancias que constituirá la nueva entrada al método de segmentación también ha quedado validado como elemento que aumenta en gran medida la precisión en los resultados de segmentación del hígado.

En la tabla 6.15 se ha realizado un compendio de los métodos comparables más relevantes del estado del arte en segmentación automática de hígado sobre imágenes de resonancia magnética para su comparación con el nuestro. Como se ha explicado previamente, la comparación real es difícil ya que cada método posee sus propios conjuntos de datos, con una segmentación de referencia y unas métricas diferentes. Una vez se han mostrado y analizado los resultados obtenidos a partir de los métodos propuestos en esta tesis, se puede concluir que dicho método es un método válido y prometedor para resolver el problema de la segmentación automática de hígado sobre imágenes abdominales obtenidas por resonancia magnética. Los resultados obtenidos por los diferentes enfoques propuestos están en línea con los métodos comparables del estado del arte e incluso son ligeramente superiores en algunos casos.

Se debe tener en cuenta que si bien el uso simultáneo de múltiples secuencias puede beneficiar los resultados de la segmentación, también presenta algunos inconvenientes, tales como el tiempo de adquisición adicional que se requiere o el problema de registro que surge. De hecho, el registro es un problema relevante sobre todo en órganos deformables, como es el hígado. Es necesario conseguir una alineación razonable para poder considerar múltiples canales al mismo tiempo.

La transformación de las cinco secuencias de entrada en una imagen de mayor contraste y calidad que se convierte en la nueva entrada de la superficie activa 3D implica una reducción de la dimensionalidad del problema que impacta en el coste computacional necesario para ejecutar el proceso de segmentación. Además, este proceso de segmentación es totalmente automático en detrimento de la mayoría de los métodos del estado del arte que son interactivos, y por lo tanto, son más lentos y requieren del conocimiento y participación del usuario.

El método propuesto puede ser aplicado a otros protocolos de adquisición con diferentes máquinas de resonancia magnética y con mayor o menor número de secuencias. El comportamiento del hígado en las diferentes secuencias elegidas es almacenado en el modelo estadístico multivariable de hígado generado inicialmente. El descriptor compacto permite el manejo y combinación de la información de todas las secuencias disponibles, sin aumentar la dimensión de la entrada, y permitiendo la generación del mapa de distancias de hígado con respecto al modelo generado del mismo en el protocolo de adquisición definido.

La metodología propuesta en esta tesis es aplicable a la segmentación de cualquier otro órgano en imagen de resonancia magnética con posibilidad de entrada multisequencia.

Este modelo estadístico multivariable del órgano es susceptible de ser mejorado, añadiéndole mayor número de variables e información según se considere relevante y pudiendo, de esta manera, caracterizar mejor el tejido objeto de estudio.

Capítulo 7: CONCLUSIONES,
APORTACIONES Y LÍNEAS
FUTURAS

7.1. Comentarios sobre el trabajo realizado

El trabajo realizado en esta tesis ha sido extenso. Este documento recoge sólo aquella parte del trabajo que supone una aportación sobre el estado del arte. Se ha mostrado la aproximación teórica del modelo propuesto junto con las pruebas realizadas y los resultados obtenidos. Sin embargo, se exploraron previamente otros caminos antes de llegar a este punto.

El procesamiento de la imagen médica es complejo, como ya se ha explicado en varios puntos de esta memoria, y abordar un problema difícil como es la segmentación de un órgano deformable, el hígado, en imágenes de resonancia magnética, con poco contraste y nitidez y además, con entrada multicanal, ha necesitado de un proceso largo de aprendizaje, valoración y pruebas.

Esta memoria recoge dicha trayectoria. Inicialmente se investigaron los comienzos del procesamiento de imagen médica, las primeras herramientas, las diferentes tecnologías de adquisición disponibles, la problemática de cada tipo de imágenes y las diferencias entre ellas. Se abordó la problemática concreta referida a resonancia magnética, en particular aplicada a segmentación de hígado, y sobre entrada multisequencia.

En las primeras aproximaciones al problema, se probó la bondad de las librerías ITK y VTK para la segmentación de imágenes, codificando algunos métodos, basados en umbralización, o crecimiento de regiones, entre otros. Estas primeras pruebas permitieron comprobar la dificultad que se ponía de manifiesto en la literatura. Se implementaron algoritmos basados en contornos activos de forma y de apariencia, dada la naturaleza deformable del órgano de estudio.

La disponibilidad y en principio, ventaja, de la entrada multisequencia debía aprovecharse. Se debía gestionar la entrada múltiple de manera que los resultados de la segmentación se vieran mejorados. Por ello, se exploró inicialmente la posibilidad de generar un vector de características a partir de un conjunto de descriptores extendidos en procesamiento de imagen (ver capítulo 3) obtenidos en las diferentes secuencias para los diferentes tejidos y órganos presentes en la imagen abdominal, lo que constituía la obtención de una firma espectro-espacial de cada órgano, entendiendo cada secuencia como una banda de dicho "espectro". Una vez obtenido dicho vector de características se probó tanto con métodos de clasificación no supervisado como supervisado del estado del arte (ver capítulo 4). Con este método no se consiguió una segmentación del hígado satisfactoria. Sin embargo, este enfoque inicial implicó un gran estudio e investigación de todos los descriptores susceptibles de ser aplicados sobre imagen médica, y de los métodos de segmentación con mayor potencial. Esto ha dado lugar a la identificación de una extensa cantidad de información sobre el estado del arte que ha sido recogida en los capítulos 3 y 4, que abordan los descriptores y métodos de segmentación, respectivamente.

Finalmente, el descubrimiento de las técnicas variacionales como posible solución al problema de minimización de un contorno activo, y herramienta de regularización tanto en

2D como 3D, las propuestas de nuevos modelos de contorno activo junto con la disponibilidad de la multisequencia permitieron esbozar lo que serían los desarrollos propuestos en esta tesis y en definitiva, la aportación científica.

Se exponen a continuación las conclusiones, las aportaciones de la tesis y las líneas identificadas de trabajo futuro.

7.2. Conclusiones

A partir del trabajo realizado en esta tesis y del análisis de los resultados mostrados en el capítulo anterior se extraen las siguientes conclusiones:

- El procesamiento y análisis de la imagen médica es un problema complejo que requiere de un conocimiento especializado. La imagen médica presenta muchos artefactos, ruidos y efectos de iluminación no homogénea en la imagen, causados por las tecnologías de adquisición, que hacen necesario el uso de técnicas avanzadas para la segmentación correcta y precisa de los tejidos presentes en las imágenes.
- La segmentación automática del hígado en imágenes de resonancia magnética es una tarea complicada. Las imágenes de MRI aportan mucha información para el diagnóstico, pero son imágenes de baja calidad, con ruidos, artefactos de adquisición y movimiento, efecto de volumen parcial, y en general, baja nitidez y escaso contraste entre órganos, de manera que el límite entre los diferentes tejidos suele ser confuso. El hígado, que es un órgano que ocupa buena parte de la cavidad abdominal, es adyacente al riñón, intestinos, pulmón, vesícula, o grasa abdominal, algunos de los cuales presentan niveles de gris semejantes a los del hígado. En estas circunstancias, la localización y segmentación automática del hígado en imágenes de MRI se convierte en una tarea difícil.
- El método de segmentación propuesto, consistente en un modelo de contorno activo en 2D (o superficie activa en 3D) que evoluciona sobre una imagen generada en base a un nuevo descriptor propuesto, y acorde a un proceso de minimización mediante aproximación dual al problema variacional, proporciona buenos resultados en la segmentación. Los valores obtenidos de las métricas de validación están alineadas (y son incluso mejores en algunos casos) con los de otros métodos comparables del estado del arte.
- Las técnicas avanzadas de minimización de modelos deformables para segmentación, como son las técnicas variacionales y en concreto la aproximación dual propuesta por Chambolle, permiten segmentaciones más precisas, con curvaturas flexibles, que otros métodos más tradicionales no consiguen. Además, no dependen de parámetros externos que delimiten el punto de corte de la función de energía minimizadora sino que, como se ha visto previamente, a partir del análisis de la distribución probabilística entre la zona interior y exterior de dicha superficie o volumen delimitado por la función de energía, es capaz de encontrar el punto mínimo, y por tanto, el borde o zona de

separación entre regiones homogéneas, es decir, los objetos segmentados en la imagen. Son independientes de la inicialización y pueden segmentar regiones interiores.

- La extensión del método de minimización por aproximación dual al problema variacional de dos dimensiones a tres dimensiones, propuesto e implementado en esta tesis, aporta un valor limitado. En este caso de aplicación, la minimización en tres dimensiones (sobre el volumen que constituye una secuencia) mejora los resultados con respecto a la realizada sobre dos dimensiones (sobre cada corte de la secuencia). La superficie activa realiza una mejor regularización, obteniendo un volumen segmentado con curvaturas más suaves y constituyendo una solución ligeramente más estable, en cuanto a que la desviación estándar de los valores de las métricas en el conjunto de validación es menor. El contorno activo en 3D se denomina superficie activa, y se ha partido de la hipótesis de que la tercera dimensión añade información sobre la continuidad del volumen segmentado en dicho tercer eje, lo cual facilitaría la localización del mínimo global, y por tanto, de la solución óptima y más estable, ya que no sólo se segmenta a través de imágenes 2D, sino tratando cada caso como el volumen 3D que representan. La extensión 3D ha mejorado los resultados ligeramente.
- La información disponible existente en las diferentes 5 secuencias es complementaria y aporta valor para una segmentación más precisa.
- La combinación y el manejo de la información multisequencia a partir del descriptor compacto propuesto en esta tesis ha sido validada con éxito. La generación de una imagen que representa un mapa de distancias al modelo estadístico multivariable de hígado a partir del descriptor compacto propuesto es valiosa y mejora el resultado de la segmentación con respecto a la aplicación del mismo método sobre una sola secuencia. La imagen resultante presenta mayor contraste entre los tejidos presentes, con lo que aumenta la calidad de la imagen. Se recuerda en este punto que para poder emplear de manera efectiva la información multisequencia, todas las secuencias deben estar convenientemente registradas, esto es, que cada vóxel de la secuencia, cada píxel de cada corte 2D de cada una de las N secuencias esté referido a un mismo punto físico. En este caso, el registro conseguido es aceptable (como se ha mostrado en el capítulo 6 de pruebas), lo cual ha permitido la obtención de estos resultados.
- El descriptor compacto propuesto reduce notablemente el tiempo de cómputo necesario y permite la reducción de la dimensión del problema. En este caso, las 5 secuencias de entrada han sido transformadas en una única imagen que representa un mapa de distancias de cada píxel de la imagen con respecto a un modelo estadístico multivariable de hígado previamente generado. Si se aplica la misma metodología a cualquier otro caso, se puede transformar una entrada de N canales en una sola entrada. Esto es de gran importancia para reducir la dimensión del problema en caso de disponer de múltiples secuencias del mismo estudio, y facilita la aplicación de técnicas de modelos deformables o superficies activas sobre esta entrada única que contiene la información de todos los canales.

- Los valores de las métricas de validación obtenidas a partir de las segmentaciones realizadas ponen de manifiesto una mejora en los resultados obtenidos por otras técnicas del estado del arte. Esto ratifica que los métodos propuestos en el marco de esta tesis son válidos y demuestran su aplicabilidad y eficiencia en el proceso de segmentación de imagen médica, y en concreto, de segmentación hepática sobre imagen de resonancia magnética.
- El método propuesto puede ser aplicado a otros protocolos de adquisición diferentes procedentes de distintas máquinas de resonancia magnética. La respuesta del hígado a un determinado protocolo de adquisición se almacenaría en el modelo de hígado generado previamente. El mapa de distancias que integra la información de los múltiples canales de entrada se calcula en base a este modelo. La metodología propuesta puede ser extendida a cualquier otro órgano.

Sin embargo, aunque los métodos propuestos en esta tesis proporcionan buenos resultados de segmentación, se necesitan pruebas más exhaustivas de validación para su uso en la práctica clínica.

7.3. Aportaciones de la tesis

Las **aportaciones de esta tesis son:**

- **Análisis exhaustivo del estado del arte de descriptores y de técnicas avanzadas de segmentación de imagen médica**, aplicables a la **segmentación hepática** en imagen abdominal de resonancia magnética, tanto en 2D, 3D, con múltiples canales de entrada, habiéndose analizado desde los métodos más sencillos basados en análisis de histograma y selección de umbral hasta la resolución de complejas ecuaciones en derivadas parciales de las técnicas variacionales para la minimización de los funcionales de energía que representan los modelos deformables, como son los contornos activos.
- **Aplicación y validación del método de minimización de un contorno activo mediante aproximación dual al problema variacional para la segmentación automática de hígado en imágenes 2D**, que se corresponden con los cortes de una secuencia de estudio. Los resultados de la segmentación han sido buenos y son comparables con los de otros métodos identificados en el estado del arte.
- **Extensión y validación de dicho método de dos a tres dimensiones**. Esta extensión que se ha propuesto e implementado en el marco de esta tesis es viable técnicamente y válida para la segmentación de hígado sobre imágenes de MRI. La aplicación del método en 3D mejora los resultados de segmentación ligeramente, incluso sobre una sola secuencia en la entrada. La mejora se traduce en un volumen de hígado segmentado con contornos más suavizados y un aumento de la estabilidad de la solución, cuantificada en un valor menor de la desviación estándar del conjunto de valores de las métricas.

- **La propuesta de uso simultáneo de varias secuencias obtenidas en el mismo estudio** a partir de diferentes configuraciones del campo electromagnético como información complementaria que mejore la segmentación ha sido considerada, y se ha validado que mejora los resultados de la segmentación.
- **Nuevo descriptor compacto basado en modelo estadístico multivariable.** El descriptor propuesto tiene una doble función. Por una parte permite aumentar el contraste en la imagen de entrada (tanto monosecuencia como multisequencia) y además, en la entrada multisequencia, permite manejar y combinar la información de los múltiples canales. Mediante la aplicación de dicho descriptor, se **genera un mapa de distancias por cada corte de la secuencia**, de manera que el valor de cada píxel en la imagen 2D se corresponde con la media de la distancia al modelo gaussiano multivariante del valor de gris de cada píxel en cada una de las secuencias con el valor de gris del modelo de hígado generado previamente y con respecto a la desviación estándar. De esta manera se ha transformado la información contenida en 5 secuencias, es decir, en 5 imágenes 2D por corte, en un mapa de distancias resultante por corte que será la entrada al proceso de segmentación definido. Se ha validado, vistos los resultados, que esta aproximación realizada contribuye a aumentar la precisión de la segmentación.
- **Reducción del tiempo de cómputo del proceso de segmentación.** El descriptor compacto propuesto, es decir, el que permite generar este mapa de distancias que aglutina la información de las múltiples secuencias en base a modelo posibilita una **reducción notable del tiempo de cómputo**. El vector de entrada de N componentes, en este caso 5, ha quedado reducido a 1. Adicionalmente, esta reducción en la dimensión del vector de entrada hace **viable y comprensible la aplicación de modelos deformables** como el presentado, que son minimizados sobre una única entrada.
- **Metodología completa de pruebas para la validación de los algoritmos de segmentación** propuestos, seleccionando e implementando **9 métricas diferentes**, que han sido calculadas para todos los métodos. La validación se ha realizado sobre 6 casos clínicos. Se ha realizado la segmentación manual de los mismos para establecer la segmentación de referencia. En definitiva, la metodología establecida ha permitido una validación real de los algoritmos sobre un conjunto de casos clínicos, que aun siendo limitada, aporta conclusiones sobre la bondad de los métodos desarrollados.
- **Comparativa exhaustiva** de los resultados de segmentación obtenidos mediante los métodos implementados en esta tesis **con otros métodos de igual propósito y comparables**, de segmentación automática de hígado sobre imagen de resonancia magnética. Esta comparativa ha sido realizada a través de los valores de las métricas, que, no pueden ser comparadas en su totalidad ya que los trabajos presentes en el estado del arte dan una información parcial y heterogénea de dichas métricas, sobre bases de datos distintas entre sí, y en consecuencia segmentaciones de referencia diferentes realizadas por radiólogos dispares. Para poder comparar el trabajo realizado con todos ellos, se han implementado las 9 métricas más frecuentes.

Como aportación de la tesis, se cuenta con dos publicaciones:

- A. Bereciartua, A. Picon, A. Galdran, P. Iriondo, Automatic 3D model-based method for liver segmentation in MRI based on active contours and total variation minimization, *Biomed. Signal Process. Control.* 20 (2015) 71–77. doi:10.1016/j.bspc.2015.04.005
- A. Bereciartua, A. Picon, A. Galdran, P. Iriondo, 3D active surfaces for liver segmentation in multisequence MRI images, *Comput. Methods Programs Biomed.* 132 (2016) 149–160. doi:10.1016/j.cmpb.2016.04.028

Estas revistas se encuentran en los cuartiles Q2 y Q1 respectivamente en el año 2015.

7.4. Líneas de trabajo futuras

Las **líneas de trabajo futuro identificadas**, entendidas como aquellas que podrían contribuir a mejorar aún más los resultados de la segmentación hepática, en concreto, o la segmentación de imagen médica en general, son:

- **Validación de los métodos propuestos sobre un conjunto de casos clínicos más amplio.** De esta manera, se verificará la fiabilidad de los resultados y se valorará la robustez de los métodos ante hígados sanos o con patologías.
- **Configuración de nuevas secuencias en la máquina de resonancia magnética.** La configuración de nuevos protocolos de adquisición pretende que el campo electromagnético generado provoque respuestas diversas sobre los órganos del cuerpo, ocasionando que la información recogida en las imágenes de las secuencias sea diferente pero complementaria. Esta tarea debe realizarse con ayuda de un radiólogo.
- **Registro óptimo de las secuencias.** Las técnicas y procesos de segmentación trabajados sobre información multisequencia serán tanto más eficientes cuanto mejor esté registrado el conjunto de secuencias. En las secuencias perfectamente registradas, cada píxel se refiere al mismo punto físico del espacio. De esta manera se aprovecha al máximo la información complementaria existente en ellas. El proceso de registro en sí mismo constituye una importante línea de trabajo, siendo problemático en órganos como el hígado. El hígado pertenece al grupo de órganos denominados “deformables”, ya que pueden variar su forma a lo largo del tiempo que dura el proceso de adquisición. Además, la forma redondeada del hígado añade gran complejidad al proceso al no poder establecerse unos puntos de interés, fácilmente identificables o fijos en el espacio.
- **Análisis de secuencias determinantes en la segmentación.** En esta tesis se han empleado las cinco secuencias disponibles en el estudio acorde al protocolo de adquisición establecido por criterio clínico, y se ha supuesto que todas aportaban valor, si bien no se ha verificado con métodos específicos. Es posible que algunas de ellas ofrezcan información redundante que, lejos de aportar valor, empeoren el resultado. No

se ha analizado esto. Podrían aplicarse métodos de Análisis de Componentes Principales (PCA o Principal Component Analysis) u otros.

- **Mejora de la calidad de las imágenes capturadas.** La falta de homogeneidad presente en la iluminación de la imagen y la falta de nitidez y contraste en algunos puntos debidas a los artefactos de captura de las máquinas de resonancia magnética, la respuesta de los órganos blandos del cuerpo, la deformación y los efectos de la baja resolución en algunos casos, han sido parcialmente superadas por el empleo de los métodos propuestos y sobre todo, por el uso de la entrada multisequencia. En la medida en que se apliquen filtros o métodos que reduzcan dichos problemas y mejoren la calidad de la imagen adquirida, los resultados de la segmentación sobre estas imágenes van a ser más precisos.
- **Definición de un modelo de contorno activo más complejo,** que añada más términos al funcional de energía además de la distribución probabilística de la región y de las consideraciones de gradientes para detección de bordes, como pudieran ser términos referidos a la forma, de manera que pueda realizar segmentaciones más finas.
- **Desarrollo de nuevos métodos de segmentación basado de aprendizaje profundo o *Deep Learning*.** Las técnicas de aprendizaje profundo están teniendo mucho auge en el ámbito de procesamiento de imagen, y están comenzando a ser aplicadas como herramienta de segmentación. No está aún desarrollado para segmentación 3D y existen muy pocas referencias en procesamiento de imagen médica. La limitación de estas técnicas es que se precisan muchos ejemplos para la fase de aprendizaje, del orden de miles de imágenes y esto puede ser un problema. Dicha técnica es capaz de identificar y extraer los descriptores más relevantes contenidos en las imágenes para la caracterización de los tejidos y realizar la clasificación de cada píxel en base a ello. La selección de descriptores se realiza de manera automática. Esta es una vía importante a explorar.
- **Generación de un modelo de hígado más complejo.** En las pruebas preliminares acometidas se implementaron una serie de descriptores estadísticos de primero (media y varianza), segundo (a partir de matrices de coocurrencia, energía, contraste...) y tercer orden (*skewness* y *kurtosis*), además de probar con otros descriptores como *Local Binary Pattern* (LBP) hasta un conjunto de 11 valores para describir el hígado. Como conclusión de aquellas pruebas iniciales se extrajo que el descriptor que más aportaba era el nivel de gris. Sin embargo, la definición e inclusión de nuevos descriptores en la generación del modelo de hígado permitiría caracterizarlo de manera más robusta y podría mejorar su identificación en la imagen.
- **Aplicación de la aproximación metodológica propuesta en esta tesis para segmentar otros elementos,** por ejemplo, las propias patologías contenidas en el hígado, tras la generación de un modelo que las caracterice. También se puede aplicar para segmentar otros órganos tras la adecuación del modelo y selección de secuencias correspondientes en el protocolo de adquisición.

Referencias

- Adams, R., Bischof, L., 1994. Seed Region Growing. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 641–647. doi:10.1109/34.295913
- Alomari, R.S., Kompalli, S., Chaudhary, V., 2008. Segmentation of the Liver from Abdominal CT Using Markov Random Field Model and GVF Snakes. 2008 Int. Conf. Complex, Intell. Softw. Intensive Syst. 293–298. doi:10.1109/CISIS.2008.135
- Amini, A., Weymouth, T., Jain, R., 1990. Using dynamic programming for solving variational problems in vision. *Pattern Anal. Mach. Intell.* 12, 855–867.
- Ankerst, M., Kastenmüller, G., Kriegel, H.-P., Seidl, T., 1999. 3D shape histograms for similarity search and classification in spatial databases, in: *International Symposium of Spatial Databases*. pp. 207–226.
- Antonidoss, Kaliyamurthie, K.P., 2014. Segmentation from Images Using Adaptive Threshold. *Middle-East J. Sci. Res.* 20, 479–484. doi:10.5829/idosi.mejsr.2014.20.04.21037
- Bao, P., Zhang, L., 2003. Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Trans. Med. Imaging* 22, 1089–99. doi:10.1109/TMI.2003.816958
- Basset, O., Sun, Z., Mestas, J.L., Giménez, G., 1993. Texture analysis of ultrasonic images of the prostate by means of co-occurrence matrices. *Ultrason. Imaging* 15, 218–237.
- Bayes, T., 1984. An essay toward solving a problem in the doctrine of chances. *Philos. Trans. R. Soc. London* 53, 370–418.
- Beichel, R., Mitchell, S., Sorantin, E., 2001. Shape-and appearance-based segmentation of volumetric medical images. *ICIP* 2, 589–592.
- Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1798–828. doi:10.1109/TPAMI.2013.50
- Bezdek, J., Ehrlich, R., Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* 10, 191–203.
- Bovik, A.C., Clark, M., Geisler, W.S., 1990. Multichannel Texture Analysis using Localized Spatial Filters. *Biomed. Tech. Biomed. Eng.* 55, 6440.
- Boykov, Y., Funka-Lea, G., 2006. Graph Cuts and Efficient N-D Image Segmentation. *Int. J. Comput. Vis.* 70, 109–131. doi:10.1007/s11263-006-7934-5
- Boykov, Y., Jolly, M., 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. *Int. Conf. Comput. Vis.* 1, 105–112.
- Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 26, 1124–37. doi:10.1109/TPAMI.2004.60
- Bresson, X., Chan, T., 2007. Active contours based on chambolle’s mean curvature motion. *Image Process. 2007. ICIP 2007. IEEE Int. Conf. On.*

- Bresson, X., Esedoğlu, S., Vandergheynst, P., Thiran, J.-P., Osher, S., 2007. Fast Global Minimization of the Active Contour/Snake Model. *J. Math. Imaging Vis.* 28, 151–167. doi:10.1007/s10851-007-0002-0
- Buades, A., Coll, B., Morel, J., 2005. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.* 2, 490–530.
- Bushong, S.C., 2003. An overview of Magnetic Resonance Imaging, in: *Magnetic Resonance Imaging. Physical and Biological Principles*. Mosby, pp. 3–18.
- Campadelli, P., Casiraghi, E., Esposito, A., 2009. Liver segmentation from computed tomography scans: a survey and a new algorithm. *Artif. Intell. Med.* 45, 185–96. doi:10.1016/j.artmed.2008.07.020
- Canny, J., 1986. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 8, 679–698.
- Cao, G., Shi, P., Hu, B., 2005. Liver fibrosis identification based on ultrasound images captured under varied imaging protocols. *J. Zhejiang Univ. Sci. B* 6, 1107–1114. doi:10.1631/jzus.2005.B1107
- Caselles, V., Catta, F., Coll, T., 1993. A Geometric Model for Active Contours in Image Processing. *Numer. Math.* 66, 1–31.
- Caselles, V., Chambolle, A., Novaga, M., 2011. Total variation in imaging. *Handb. Math. Methods Imaging* 1.
- Caselles, V., Kimmel, R., Sapiro, G., 1997. Geodesic Active Contours. *Int. J. Comput. Vis.* 22, 61–79. doi:10.1109/83.951533
- Chambolle, A., 2004. An Algorithm for Total Variation Minimization and Applications. *J. Math. Imaging Vis.* 20, 89–97. doi:10.1023/B:JMIV.0000011321.19549.88
- Chambolle, A., 2004. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*
- Chambolle, A., Caselles, V., 2010. An introduction to total variation for image analysis. *Theor. Found. Numer. Methods Sparse Recover.*
- Chan, T.F., Esedoglu, S., Nikolova, M., 2006. Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *SIAM J. Appl. Math.* 66, 1632–1648. doi:10.1137/040615286
- Chan, T.F., Sandberg, B.Y., Vese, L. a., 2000. Active Contours without Edges for Vector-Valued Images. *J. Vis. Commun. Image Represent.* 11, 130–141. doi:10.1006/jvci.1999.0442
- Chan, T.F., Vese, L. a., 2001. Active contours without edges. *IEEE Trans. Image Process.* 10, 266–77. doi:10.1109/83.902291
- Chen, D.-R., Chang, R.-F., Kuo, W.-J., Chen, M.-C., Huan, Y.-L., 2002. Diagnosis of breast tumours with sonographic texture analysis using wavelet transform and neural networks. *Ultrasound Med. Biol.* 28, 1301–1310.
- Chen, G., Gu, L., Qian, L., Xu, J., 2009. An Improved Level Set for Liver Segmentation and Perfusion Analysis in MRIs. *IEEE Trans. Inf. Technol. Biomed.* 13, 94–103.

doi:10.1109/TITB.2008.2007110

- Chen, X., Udupa, J.K., Bagci, U., Zhuge, Y., Yao, J., 2012. Medical image segmentation by combining graph cuts and oriented active appearance models. *IEEE Trans. Image Process.* 21, 2035–46. doi:10.1109/TIP.2012.2186306
- Chen, Y., Tagare, H., 2002. Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vis.* 50, 315–328.
- Chen, Y., Wanga, Z., Hua, J., Zhaoa, W., Wub, Q., 2012. The domain knowledge based graph-cut model for liver CT segmentation. *Biomed. Signal Process. Control* 7, 591–598.
- Chen, Y.-W., Tsubokawa, K., Foruzan, A., 2010. Liver Segmentation from Low Contrast Open MR Scans Using K-Means Clustering and Graph-Cuts, in: Zhang, L., Lu, B.-L., Kwok, J. (Eds.), *Advances in Neural Networks - ISNN 2010 SE - 21*, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 162–169. doi:10.1007/978-3-642-13318-3_21
- Cheng, K., Gu, L., Xu, J., 2008. A novel shape prior based level set method for liver segmentation from MR Images. 2008 *Int. Conf. Technol. Appl. Biomed.* 144–147. doi:10.1109/ITAB.2008.4570544
- Chung, S., Logeswaran, R., 2007. Evaluation of Texture Analysis Techniques for Characterization of Multimode-Based Liver Images Using SGLCM and FOS, in: Ibrahim, F., Osman, N., Usman, J., Kadri, N. (Eds.), *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006 SE - 68*, IFMBE Proceedings. Springer Berlin Heidelberg, pp. 262–266. doi:10.1007/978-3-540-68017-8_68
- Clarke, L., Velthuizen, R., Cmacho, M., Heine, J., Vaidyanathan, M., Hall, L., Thatcher, RW, Silbiger, M., 1995. MRI segmentation: methods and applications. *Magn. Reson. Imaging* 13, 343–368.
- Cohen, L., 1991. On active contour models and balloons. *CVGIP Image Underst.* 53, 211–218.
- Cohen, L., Cohen, I., 1993. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *Pattern Anal. Mach. Intell.* ... 15, 1131–1147.
- Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 603–619.
- Cootes, T., Edwards, G., Taylor, C., 1998. Active Appearance Models. *Comput. Vis.* — ECCV'98. *Lect. Notes Comput. Sci.* 2, 484–498.
- Cootes, T., Hill, A., Taylor, C., Haslam, J., 1994. Use of active shape models for locating structures in medical images. *Image Vis. Comput.* 12, 355–366. doi:10.1016/0262-8856(94)90060-4
- Cootes, T., Taylor, C., 2001. Statistical models of appearance for medical image analysis and computer vision. *Med. Imaging* 2001 236–248.
- Cootes, T., Taylor, C., Cooper, D., Graham, J., 1995. Active shape models-their training and application. *Comput. Vis. Image Underst.* 61, 38–59.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297.

- Couinaud, C., 1957. *Le Foie: Études anatomiques et chirurgicales* [The Liver: Anatomical and Surgical Studies]. Paris: Masson.
- Damadian, R., 1971. Tumor detection by nuclear magnetic resonance. *Science* (80-). 171, 1151–1153.
- Danciu, M., Gordan, M., Florea, C., Vlaicu, A., 2012. 3D DCT supervised segmentation applied on liver volumes, in: 35th International Conference on Telecommunications and Signal Processing. pp. 779 – 783.
- Daugman, J.G., 1980. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res.* 20, 847–856.
- Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. ...* 39, 1–38.
- Duncan, J., Ayache, N., 2000. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 85–106.
- Fang, F., Li, F., Yang, X., Shen, C., Zhang, G., 2010. Single image dehazing and denoising with variational method. 2010 Int. Conf. Image Anal. Signal Process. 219–222. doi:10.1109/IASP.2010.5476126
- Fisher, R., 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7, 179–188.
- Fix, E., Hodges, J., 1951. Discriminatory analysis nonparametric discrimination: consistency properties. USAF School of Aviation Medicine, Randolph Field, Texas.
- Foruzan, A., Yen-Wei, C., 2013. Segmentation of Liver in Low-Contrast Images Using K-Means Clustering and Geodesic Active Contour Algorithms. *IEICE Trans. INF SYST.* E96-D, 798–807.
- Furukawa, D., Akinobu, S., Kobakate, H., 2007. Automatic liver segmentation method based on maximum a posterior probability estimation and level set method. *Proc. MICCAI Work. 3-D Segmentat. Clin. A Gd. Chall.* 117–124.
- Gabor, D., 1946. Theory of communication. *J. Inst. Electr. Eng.* 93, 429–457.
- Gambino, O., Vitabile, S., Re, G. Lo, Tona, G. La, Librizzi, S., Pirrone, R., Ardizzone, E., Midiri, M., 2010. Automatic Volumetric Liver Segmentation Using Texture Based Region Growing. *Complex, Intell. Softw. Intensive Syst. (CISIS)*, 2010 Int. Conf. pp.146–152.
- Ginneken, B. Van, Frangi, A., 2002. Active shape model segmentation with optimal features. *IEEE Trans. Med. Imaging* 21, 924–933.
- Gloger, O., Kühn, J., Stanski, A., Völzke, H., Puls, R., 2010. A fully automatic three-step liver segmentation method on LDA-based probability maps for multiple contrast MRI images. *Magn. Reson. Imaging* 28, 882–897. doi:http://dx.doi.org/10.1016/j.mri.2010.03.010
- Gloger, O., Toennies, K., Kuehn, J.-P., 2011. Fully Automatic Liver Volumetry Using 3D Level Set Segmentation for Differentiated Liver Tissue Types in Multiple Contrast MR Datasets, in: Heyden, A., Kahl, F. (Eds.), *Image Analysis SE - 48, Lecture Notes in Computer Science*. Springer Berlin Heidelberg, pp. 512–523. doi:10.1007/978-3-

- Göçeri, E., Gürcan, M.N., Dicle, O., 2014. Fully automated liver segmentation from SPIR image series. *Comput. Biol. Med.* 53, 265–278. doi:10.1016/j.combiomed.2014.08.009
- Göçeri, E., Ünlü, M., Dicle, O., 2015. A comparative performance evaluation of various approaches for liver segmentation from SPIR images. *Turkish J. Electr. Eng. Comput. Sci.* 23, 741–768. doi:10.3906/elk-1304-36
- González Ballester, M., Zisserman, A.P., Brady, M., 2002. Estimation of the partial volume effect in MRI. *Med. Image Anal.* 6, 389–405.
- Gulrud, T.O., Kjode, S., 1996. Optimal filter for detection of stellate lesions and circumscribed masses in mammograms. *SPIE Vis. Commun. Image Process. I*, 430–440.
- Gunn, S.R., Nixon, M.S., 1997. A robust snake implementation; a dual active contour. *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 63–68. doi:10.1109/34.566812
- Haralick, R., 1973. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* SMC-3, 610–621.
- Haralick, R.M., 1979. Statistical and structural approaches to texture. *Proc. IEEE* 67, 786–804.
- Heimann, T., Meinzer, H., Wolf, I., 2007. A Statistical Deformable Model for the Segmentation of Liver CT Volumes Using Extended Training Data. *Proc. MICCAI Work. 3-D Segmentat. Clin. A Gd. Challenge*, 161–166.
- Heimann, T., Meinzer, H.-P., 2009. Statistical shape models for 3D medical image segmentation: a review. *Med. Image Anal.* 13, 543–63. doi:10.1016/j.media.2009.05.004
- Heimann, T., van Ginneken, B., Styner, M. a, Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., Bello, F., Binnig, G., Bischof, H., Bornik, A., Cashman, P.M.M., Chi, Y., Cordova, A., Dawant, B.M., Fidrich, M., Furst, J.D., Furukawa, D., Grenacher, L., Hornegger, J., Kainmüller, D., Kitney, R.I., Kobatake, H., Lamecker, H., Lange, T., Lee, J., Lennon, B., Li, R., Li, S., Meinzer, H.-P., Nemeth, G., Raicu, D.S., Rau, A.-M., van Rikxoort, E.M., Rousson, M., Rusko, L., Saddi, K. a, Schmidt, G., Seghers, D., Shimizu, A., Slagmolen, P., Sorantin, E., Soza, G., Susomboon, R., Waite, J.M., Wimmer, A., Wolf, I., 2009. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans. Med. Imaging* 28, 1251–65. doi:10.1109/TMI.2009.2013851
- Hirokawa, Y., Isoda, H., Maetani, Y.S., Arizono, S., Shimada, K., Togashi, K., 2008. Evaluation of Motion Correction Effect and Image Quality With the Periodically Rotated Overlapping Parallel Lines With Enhanced Reconstruction (PROPELLER) (BLADE) and Parallel Imaging Acquisition Technique in the Upper Abdomen. *J. Magn. Reson. Imaging* 28, 957–962. doi:10.1002/jmri.21538
- Huynh, H.T., Karademir, I., Oto, A., Suzuki, K., 2014. Computerized liver volumetry on MRI by using 3D geodesic active contour segmentation. *AJR. Am. J. Roentgenol.* 202, 152–9. doi:10.2214/AJR.13.10812
- Iakovidis, D.K., Savelonas, M. a., Karkanis, S. a., Maroulis, D.E., 2006. Segmentation of

- Medical Images with Regional Inhomogeneities. 18th Int. Conf. Pattern Recognit. 976–979. doi:10.1109/ICPR.2006.1036
- Ibrahim, H., Petrou, M., Wells, K., Doran, S., Olsen, Ø., 2010. Automatic Volumetric Liver Segmentation from MRI Data. *Int. J. Comput. Theory Eng.* 2, 176–179. doi:10.7763/IJCTE.2010.V2.136
- Iftekharruddin, K.M., Zheng, J., Islam, M. a., Ogg, R.J., Lanningham, F., 2006. Brain Tumor Detection in MRI: Technique and Statistical Validation. 2006 Fortieth Asilomar Conf. Signals, Syst. Comput. 1983–1987. doi:10.1109/ACSSC.2006.355112
- Ivins, J., Porril, J., 1995. Active Region Models for segmenting textures and colours. *Image Vis. Comput.* 13, 431–437.
- Jackson, A., Kassner, A., Annesley-Williams, D., Reid, H., Zhu, X.-P., Li, K.-L., 2002. Abnormalities in the recirculation phase of contrast agent bolus passage in cerebral gliomas: comparison with relative blood volume and tumor grade. *AJNR. Am. J. Neuroradiol.* 23, 7–14.
- James, A.P., Dasarathy, B. V., 2014. Medical image fusion: A survey of the state of the art. *Inf. Fusion* 19, 4–19. doi:10.1016/j.inffus.2013.12.002
- Jolliffe, I., 2005. Choosing a Subset of Principal Components or Variables, in: *Principal Component Analysis*. Springer, pp. 111–130.
- Kainmüller, D., Lange, T., Lamecker, H., 2007. Shape Constrained Automatic Segmentation of the Liver based on a Heuristic Intensity Model. *Proc. MICCAI Work. 3-D Segmentat. Clin. A Gd. Chall.* 109–116.
- Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active contour models. *Int. J. Comput. Vis.* 321–331.
- Kazhdan, M., Funkhouser, T., Rusinkiewicz, S., 2003. Rotation invariant spherical harmonic representation of 3 D shape descriptors. *Symp. Geom. Process.* 156–165.
- Kirsch, R., 1971. Computer determination of the constituent structure of biological images. *Comput. Biomed. Res.* 4, 315–328.
- KISTLER, J.P., HOCHBERG, F.H., BROOKS, B.R., RICHARDSON, E.P., NEW, P.F.J., SCHNUR, J., 1975. Computerized axial tomography: Clinicopathologic correlation. *Neurol.* 25, 201. doi:10.1212/WNL.25.3.201
- Lai, K.F., Kin, R.T., 1994. On regularization, initialization and extraction of the active contour model (snakes), in: *Proceedings of the 1st Asian Conference on Computer Vision*. pp. 542–545.
- LAUTERBUR, P.C., 1973. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature* 242, 190–191.
- Lee, J., Kim, N., Lee, H., Seo, J.B., Won, H.J., 2007. Efficient Liver Segmentation exploiting Level-Set Speed Images with 2 . 5D Shape Propagation. *Proc. MICCAI Work. 3-D Segmentat. Clin. A Gand, Challenge*, 189– 196.
- Li, B.N., Chui, C.K., Chang, S., Ong, S.H., 2011. Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. *Comput. Biol. Med.* 41, 1–10. doi:10.1016/j.combiomed.2010.10.007

- Lin, J., Cheng, K., Mao, C., 1996. Multispectral magnetic resonance images segmentation using fuzzy Hopfield neural network. *Int. J. Biomed. Comput.* 42, 205–214.
- Linguraru, M.G., Li, Z., Shah, F., Chin, S., Summers, R.M., 2009. Automated Liver Segmentation using a Normalized Probabilistic Atlas. *Proc. SPIE 7262, Med. Imaging 2009 Biomed. Appl. Mol. Struct. Funct. Imaging*, 7262, 72622R–72622R–8. doi:10.1117/12.810938
- Lloyd, S., 1982. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* 28, 129–137. doi:10.1109/TIT.1982.1056489
- Loncaric, S., 1998. A survey of shape analysis techniques. *Pattern Recognit.* 31, 983–1001. doi:10.1016/S0031-2023(97)00122-2
- López-Mir, F., Naranjo, V., Angulo, J., Alcañiz, M., Luna, L., 2014. Liver segmentation in MRI: A fully automatic method based on stochastic partitions. *Comput. Methods Programs Biomed.* 114, 11–28. doi:10.1016/j.cmpb.2013.12.022
- Luo, S., Hu, Q., He, X., Li, J., Jin, J.S., Park, M., 2009. Automatic liver parenchyma segmentation from abdominal CT images using support vector machines. 2009 ICME *Int. Conf. Complex Med. Eng.* 1–5. doi:10.1109/ICCME.2009.4906625
- Luo, S., Li, X., Li, J., 2014. Review on the Methods of Automatic Liver Segmentation from Abdominal Images. *J. Comput. Commun.* 2, 1–7.
- Luo, S., Li, X., Li, J., 2013. Improvement of Liver Segmentation by Combining High Order Statistical Texture Features with Anatomical Structural Features. *Engineering* 05, 67–72. doi:10.4236/eng.2013.55B014
- Malladi, R., Sethian, J. a., Vemuri, B.C., 1995. Shape modeling with front propagation: a level set approach. *IEEE Trans. Pattern Anal. Mach. Intell.* 17, 158–175. doi:10.1109/34.368173
- Mandelbrot, B., 1982. *The fractal geometry of nature*. Freeman, New York.
- Manjón, J. V, Coupé, P., Buades, A., Louis Collins, D., Robles, M., 2012. New methods for MRI denoising based on sparseness and self-similarity. *Med. Image Anal.* 16, 18–27. doi:10.1016/j.media.2011.04.003
- Marr, D., Hildreth, E., 1980. Theory of edge detection. *Proc. R. Soc. London. Ser. B, Biol. Sci.* 207, 187–217.
- Masoumi, H., Behrad, A., Pourmina, M.A., Roosta, A., 2012. Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network. *Biomed. Signal Process. Control* 7, 429–437. doi:10.1016/j.bspc.2012.01.002
- Massotier, L., Casciaro, S., 2007. Fully automatic liver segmentation through graph-cut technique. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2007, 5243–6. doi:10.1109/IEMBS.2007.4353524
- Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W., 2003. PET-CT image registration in the chest using free-form deformations. *IEEE Trans. Med. Imaging* 22, 120–128. doi:10.1109/TMI.2003.809072
- Mayerhoefer, M.E., Schima, W., Trattnig, S., Pinker, K., Berger-Kulemann, V., Ba-Ssalamah, A., 2010. Texture-based classification of focal liver lesions on MRI at 3.0

- Tesla: A feasibility study in cysts and hemangiomas. *J. Magn. Reson. Imaging* 32, 352–359. doi:10.1002/jmri.22268
- Mortensen, E., Barrett, W., 1995. Intelligent scissors for image composition, in: *Proceedings of the 22nd Annual Conference ...* pp. 191–198.
- Mumford, D., Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.* 42, 577–685.
- Nishiura, T., 2005. Ultrasound evaluation of the fibrosis stage in chronic liver disease by the simultaneous use of low and high frequency probes. *Br. J. Radiol.* 78, 189–197. doi:10.1259/bjr/75208448
- Oh, J., Martin, D.R., Hu, X., 2014. Partitioned edge-function-scaled region-based active contour (p-ESRAC): Automated liver segmentation in multiphase contrast-enhanced MRI. *Med. Phys.* 41, 041914. doi:10.1118/1.4867865
- Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* 29.
- Okada, T., Shimada, R., Hori, M., Nakamoto, M., Chen, Y.-W., Nakamura, H., Sato, Y., 2008. Automated Segmentation of the Liver from 3D CT Images Using Probabilistic Atlas and Multilevel Statistical Shape Model. *Acad. Radiol.* 15, 1390–1403. doi:http://dx.doi.org/10.1016/j.acra.2008.07.008
- Oliveira, D.A., Feitosa, R.Q., Correia, M.M., 2011. Segmentation of liver, its vessels and lesions from CT images for surgical planning. *Biomed. Eng. Online* 10, 30. doi:10.1186/1475-925X-10-30
- Osher, S., Sethian, J. a, 1988. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.* 79, 12–49. doi:10.1016/0021-9991(88)90002-2
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. man Cybern.* 9, 62–66.
- Paganelli, C., Peroni, M., Pennati, F., Baroni, G., 2012. Scale Invariant Feature Transform as feature tracking method in 4D imaging: A feasibility study. 2012 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 6543 – 6546.
- Pal, N., Pal, S., 1993. A review on image segmentation techniques. *Pattern Recognit.* 26, 1277–1294.
- Parzen, E., 1962. On estimation of a probability density function and mode. *Ann. Math. Stat.* 33, 1065–1076.
- Peng, B., Zhang, L., Yang, J., 2010. Iterated graph cuts for image segmentation. *Comput. Vision–ACCV 2009* 677–686.
- Peng, B., Zhang, L., Zhang, D., 2013. A survey of graph theoretical approaches to image segmentation. *Pattern Recognit.* 46, 1020–1038. doi:10.1016/j.patcog.2012.09.015
- Peng, J., Kong, D., Dong, F., 2011. A new adaptive variational model for liver segmentation with region appearance propagation. Preprint.
- Peng, J., Wang, J., Kong, D., 2012. A New Convex Variational Model for Liver

- Segmentation. *Int. Conf. Pattern Recognition(ICPR)*, 3754–3757.
- Peng, J., Wang, Y., Kong, D., 2014. Liver segmentation with constrained convex variational model. *Pattern Recognit. Lett.* 43, 81–88. doi:10.1016/j.patrec.2013.07.010
- Pentland, A.P., 1984. Fractal based description of natural scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* 6, 661–672.
- Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 629–639.
- Pham, D.L., Xu, C., Prince, J.L., 1998. A survey of current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–337.
- Pipe, J.G., 1999. Periodically Rotated Overlapping Parallel Lines with Enhanced Reconstruction (PROPELLER) MRI ; Application to Motion Correction [WWW Document]. *Int. Soc. Magn. Reson. Med. Website*. URL <http://cds.ismrm.org/ismrm-1999/PDF1/242.pdf>
- Platero, C., Gonz, P., Tobar, M.C., Poncela, J.M., Sanguino, J., Asensio, G., Santos, E., 2008a. Automatic method to segment the liver on multi-phase MRI. *Comput. Assist. Radiol. Surg.* 22nd Int. Congr. Exhib. 2008. 1–4.
- Platero, C., Poncela, J., Gonzalez Manchon, P.M., Tobar Puente, M. del C., Sanguino Botella, F.J., Asensio Madrid, G., Santos, E., 2008b. Liver segmentation for hepatic lesions detection and characterisation. *IEEE Int. Symp. Biomed. Imaging - From Nano to Macro*, 2008.
- Platero, C., Tobar, M., Sanguino, J., 2011. Level set segmentation with shape and appearance models using affine moment descriptors. *Pattern Recognit. Image Anal.* 109–116.
- Prewitt, J., 1970. Object enhancement and extraction. *Pict. Process. Psychopictorics*.
- Punia, R., Singh, S., 2013. Review on Machine Learning Techniques for Automatic Segmentation of Liver Images. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* 3, 666–670.
- Qu, W., Huang, X., Jia, Y., 2008. Segmentation in noisy medical images using PCA model based particle filtering. p. 69143I–69143I–7.
- Reska, D., Boldak, C., Kretowski, M., 2012. Fast 3D Segmentation of Hepatic Images Combining Region and Boundary Criteria. *Image Process. Commun.* 17, 31–38. doi:10.2478/v10248-012-0026-y
- Reynolds, D., Rose, R., 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. *Speech Audio Process. IEEE Trans.* 3, 72 –83.
- Rofsky, N.M., Lee, V.S., Laub, G., Pollack, M.A., Krinsky, G.A., Thomasson, D., Ambrosino, M.M., Weinreb, J.C., 1999. Abdominal MR Imaging with a Volumetric Interpolated Breath-hold Examination. *Radiology* 212, 876–884.
- Rokach, L., Maimon, O., 2005. Decision trees, in: *Data Mining and Knowledge Discovery Handbook*. Springer US, TelAviv, pp. 165–192.
- Rosenkrantz, A.B., Patel, J.M., Babb, J.S., Storey, P., Hecht, E.M., 2010. Liver MRI at 3 T using a respiratory-triggered time-efficient 3D T2-weighted technique: impact on

- artifacts and image quality. *AJR. Am. J. Roentgenol.* 194, 634–41.
doi:10.2214/AJR.09.2994
- Rother, C., Kolmogorov, V., Blake, A., 2004. “GrabCut”: Interactive Foreground Extraction Using Iterated Graph Cuts. *ACM Trans. Graph.* 23, 309–314.
doi:10.1145/1015706.1015720
- Rousson, M., Cremers, D., 2005. Efficient kernel density estimation of shape and intensity priors for level set segmentation. *Med. Image Comput. Comput. Assist. Interv.* 447–460.
- Rusko, L., Bekes, G., 2007. Fully automatic liver segmentation for contrast-enhanced CT images. *Int. Conf. Med. Image Comput. Comput. Interv. Segmentation Clin. a Gd. challenge, 2007* 143–150.
- Schmidt, G., 2007. Update on Cognition Network Technology for a Fully Automated 3D Segmentation of Liver 2007, 80339.
- Seghers, D., Slagmolen, P., Lambelin, Y., Hermans, J., Maes, F., Suetens, P., 2007. Landmark based liver segmentation using local shape and local intensity models, in: *Proc. MICCAI Workshop 3-D Segmentat. Clinic: A Grand Challenge.* pp. 135–142.
- Sethian, J., 1999a. Fast Marching methods. *SIAM Rev.* 41.
- Sethian, J., 1999b. Advancing interfaces: level set and fast marching methods. *Proc. Int. Conf. Ind. Appl. Math.*
- Sharma, N., Aggarwal, L.M., 2010. Automated medical image segmentation techniques. *J. Med. Phys.* 35, 3–14.
- Siewert, R., Schnapauff, D., 2010. Automatic Liver Segmentation in Contrast-enhanced MRI. *Bild. für die ...* 4–8.
- Sinha, R., Marcinczak, J.M., Grigat, R.-R., 2014. Automatic seed selection for segmentation of liver cirrhosis in laparoscopic sequences. *Proc. SPIE.* doi:10.1117/12.2043025
- Sobel, I., Feldman, G., 1973. A 3x3 Isotropic Gradient Operator for Image Processing, in: *Pattern Classification and Scene Analysis.* Jon Winley and sons, pp. 271–272.
- Sonka, M., Hlavac, V., Boyle, R., 1993. *Image Processing, Analysis, and Machine Vision Second Edition,* Chapman an. ed.
- Strzelecki, M., Szczypinski, P., Materka, A., Klepaczko, A., 2013. A software tool for automatic classification and segmentation of 2D/3D medical images. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 702, 137–140. doi:http://dx.doi.org/10.1016/j.nima.2012.09.006
- Susomboon, R., 2007. A hybrid approach for liver segmentation. *Proc. MICCAI Work. 3-D Segmentat. Clin. A Gd. Challenge, i,* 151–160.
- Suzuki, K., Huynh, H.T., Liu, Y., Calabrese, D., Zhou, K., Oto, A., Hori, M., 2013. Computerized segmentation of liver in hepatic CT and MRI by means of level-set geodesic active contouring. *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* 2013, 2984–7.
doi:10.1109/EMBC.2013.6610167

- Tankyevych, O., 2006. Automated abdominal tissue segmentation of multicontrast magnetic resonance images. Upsala University.
- Tomoshigea, S., Oosta, E., Shimizua, A., Watanabea, H., Nawanob, S., 2014. A conditional statistical shape model with integrated error estimation of the conditions; Application to liver segmentation in non-contrast CT images. *Med. Image Anal.* 18, 130–143.
- Tuceryan, M., Jain, A., 1998. Texture analysis. *Handb. pattern Recognit. Comput. Vis.* (2nd Ed. 1–41.
- Unay, D., Ekin, A., Cetin, M., Jasinschi, R., Ercil, A., 2007. Robustness of local binary patterns in brain MR image analysis. *Conf. Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 1, 2098–2101. doi:10.1109/IEMBS.2007.4352735
- Vannier, M., Pilgram, T., Speidel, C., Neumann, L., Rickman, D., Schertz, L., 1991. Validation of magnetic resonance imaging (MRI) multispectral tissue classification. *Comput. Med. imaging Graph.* 15, 217–23.
- Vovk, U., Pernus, F., Likar, B., 2007. A Review of Methods for Correction of Intensity Inhomogeneity in MRI. *Med. Imaging, IEEE Trans.* doi:10.1109/TMI.2006.891486
- Waite, J.B., Welsh, W.J., 1990. Head boundary location using snakes. *British Telecom J.* 8, 127–136.
- Wang, L., He, D.-C., 1990. Texture classification using texture spectrum. *Pattern Recognit.* 23, 1905–1910.
- Williams, D., Shah, M., 1992. A fast algorithm for active contours and curvature estimation. *CVGIP Image Underst.* 55, 14–26.
- Wimmer, A., Soza, G., Hornegger, J., 2009. A generic probabilistic active shape model for organ segmentation. *Med. Image Comput. Anal.* 26–33.
- Wu, C.M., Chen, Y.C., Hsieh, K.S., 1992. Texture features for classification of ultrasonic liver images. *IEEE Trans. Med. Imaging* 11, 141–152.
- Xie, J., Jiang, Y., Tsui, H.-T., 2005. Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans. Med. Imaging* 24, 45–57.
- Yuan, Z., Wang, Y., Yang, J., Liu, Y., 2010. A novel automatic liver segmentation technique for MR images. 2010 3rd Int. Congr. Image Signal Process. 1282–1286. doi:10.1109/CISP.2010.5647676
- Zhang, D., Lu, G., 2004. Review of shape representation and description techniques. *Pattern Recognit.* 37, 1–19. doi:10.1016/j.patcog.2003.07.008
- Zhang, J., Tan, T., 2002. Brief review of invariant texture analysis methods. *Pattern Recognit.* 35, 735–747. doi:10.1016/S0031-3203(01)00074-7
- Zhang, X., Fujita, H., Kanematsu, M., 2005. Improving the classification of cirrhotic liver by using texture features. *IEEE Eng. Med. Biol.*
- Zhang, X., Fujita, H., Qin, T., Zhao, J., 2008. CAD on liver using CT and MRI. *Med. Imaging Inform.* 279–288.
- Zhang, X., Tian, J., Deng, K., Wu, Y., Li, X., 2010. Automatic liver segmentation using a

statistical shape model with optimal surface detection. *IEEE Trans. Biomed. Eng.* 57, 2622–6. doi:10.1109/TBME.2010.2056369

Zhao, Y., Yunlong, Z., Wang, X., Li, G., 2010. Fuzzy C-means clustering-based multilayer perceptron neural network for liver CT images automatic segmentation, in: *Hinese Control and Decision Conference, CCDC 2010*. doi:10.1109/CCDC.2010.5498558

