

# Trabajo Fin de Grado

Grado en Ingeniería de las Tecnologías de Telecomunicación  
Especialidad en Imagen y Sonido

## Segmentación en 3D de huesos en imágenes TAC

Área de conocimiento: Teoría de la Señal y Comunicaciones

Autor: Francisco Javier Olías Sánchez

Tutor: Jose Antonio Pérez Carrasco

Dpto. Teoría de la Señal y Comunicaciones  
Escuela Técnica Superior de Ingeniería  
Universidad de Sevilla

Sevilla, 2014





Trabajo Fin de Grado  
Grado en Ingeniería de las Tecnologías de Telecomunicación  
Especialidad en Imagen y Sonido

# **Segmentación en 3D de huesos en imágenes TAC**

Autor:

Francisco Javier Olías Sánchez

Tutor:

Jose Antonio Pérez Carrasco

Dpto. de Teoría de la Señal y Comunicaciones

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2014

Trabajo Fin de Grado: Segmentación en 3D de huesos en imágenes TAC

Autor: Francisco Javier Olías Sánchez  
Tutor: Jose Antonio Pérez Carrasco

El tribunal nombrado para juzgar el Trabajo arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, a            de            2014

El Secretario del Tribunal,



# Índice

Índice de ilustraciones .....	8
Índice de tablas .....	10
Glosario.....	11
Capítulo 1: Introducción .....	12
Capítulo 2: Las imágenes médicas.....	15
2.1. Radiografía .....	16
2.2. Imágenes médicas por ultrasonidos.....	18
2.3. La resonancia magnética (RM) .....	19
2.4. Medicina nuclear.....	20
2.5. Tomografía Axial Computarizada .....	21
2.5.1. Artefactos en las imágenes TAC .....	25
Capítulo 3: Las imágenes digitales.....	28
3.1. Imagen digital .....	28
3.2. El estándar DICOM .....	28
3.1.1. Estructura de un fichero DICOM .....	29
Capítulo 4: Procesamiento de imágenes.....	31
4.1. Técnicas basadas en histograma.....	31
4.2. Matemáticas morfológicas .....	32
4.3. Detección de bordes .....	36
Capítulo 5: Técnicas de segmentación.....	40
5.1. Umbralización.....	40
5.2. Level-sets .....	40
5.3. Segmentación basada en Max-flow.....	43
5.4. Segmentación de estructuras óseas basada en métodos estadísticos y Max-flow .....	44
5.5. Método de segmentación por regiones.....	45
5.5.1. Pre-procesado de la imagen .....	46
5.5.2. Detección de bordes y rellenado .....	48
5.5.3. Clasificador de regiones.....	51
Capítulo 6: Resultados.....	52
6.1. Resultados obtenidos mediante umbralización.....	53
6.2. Resultados obtenidos con el método de Level-Sets .....	55
6.3. Resultados obtenidos con el método de Max-Flow.....	57

6.4. Resultados obtenidos con el método de segmentación de estructuras óseas basada en métodos estadísticos y Max-flow .....	61
6.5. Resultados obtenidos mediante el método de segmentación por regiones.....	63
Capítulo 7: Comparación de resultados y conclusiones .....	65
Capítulo 8: Bibliografía y referencias .....	67
Anexo A .....	69

# Índice de ilustraciones

Ilustración 1: Ejemplo de osteoporosis.....	13
Ilustración 2: Primera radiografía realizada en la historia.....	16
Ilustración 3: Generación de rayos X.....	17
Ilustración 4: Variación de los coeficientes de atenuación frente a kV aplicado.....	17
Ilustración 5: Movimientos de precesión y rotación de una peonza.....	19
Ilustración 6: Interpretación de la transformada de Radon.....	22
Ilustración 7: Primera imagen de TAC de la historia (1972).....	23
Ilustración 8: Esquema del funcionamiento de un TAC.....	23
Ilustración 9: Rango de valores Hounsfield de los distintos tejidos del cuerpo.....	24
Ilustración 10: Lightspeed16, de GE Health.....	25
Ilustración 11: Iquilion One, de Toshiba.....	25
Ilustración 12: Artefacto por endurecimiento del haz.....	26
Ilustración 13: Ejemplo de Streaking.....	26
Ilustración 14: Representación de una imagen de TAC junto con su histograma.....	31
Ilustración 15: Misma imagen que la Ilustración 14 tras transformación de histograma.....	32
Ilustración 16: Operaciones morfológicas básicas.....	33
Ilustración 17: Dilatación.....	33
Ilustración 18: Ejemplo de erosión.....	34
Ilustración 19: Ejemplo de apertura y cierre.....	35
Ilustración 20: Ejemplo de rellenado de regiones.....	35
Ilustración 21: Detección de bordes por cruces por cero.....	37
Ilustración 22: Detector de bordes Sobel.....	38
Ilustración 23: Detector de bordes Prewitt.....	38
Ilustración 24: Detector de bordes Roberts.....	38
Ilustración 25: Imagen para comparar detectores de bordes.....	38
Ilustración 26: Detector de bordes LoG.....	39
Ilustración 27: Detector de bordes Canny.....	39
Ilustración 28: Ejemplo de inicialización de level-sets.....	41
Ilustración 29: Resultado final de una segmentación usando level-sets.....	42
Ilustración 30: Inicialización de Max-flow. Figura extraída de [11].....	43
Ilustración 31: a) Imagen original, b) Imagen umbralizada, c) HDI: imagen distancia a histograma, d) CIT: imagen de costes.....	45
Ilustración 32: Diagrama de flujo del algoritmo desarrollado.....	45
Ilustración 33: Imagen original sobre la que explicaremos el algoritmo desarrollado.....	46
Ilustración 34: Diagrama de flujo de la función de ajuste.....	46
Ilustración 35: Histograma original e histograma expandido entre los valores 0.7373 y 0.8549.....	47
Ilustración 36: Ejemplo de uso de la función de ajuste de contraste automático en la Ilustración 33.....	47
Ilustración 37: Diagrama de flujo de la función de ajuste de contraste.....	48
Ilustración 38: Ilustración 36 tras haber detectado los bordes. Se marcan los bordes inconexos.....	49
Ilustración 39: Ilustración 38 tras haberle aplicado la dilatación. Los bordes inconexos se han cerrado.....	49
Ilustración 40: Ilustración 39 tras rellenarla y aplicarle una erosión.....	50

Ilustración 41: Imagen de la izquierda obtenida con máscara de erosión de tamaño 2x2. Imagen de la derecha obtenida con máscara de erosión 3x3 .....	50
Ilustración 42: Imagen de TAC del cráneo antes y después del clasificador de regiones .....	51
Ilustración 43: Imagen IM <sub>14</sub> segmentada mediante umbralización.....	54
Ilustración 44: Imagen IM <sub>2</sub> segmentada mediante umbralización .....	54
Ilustración 45: Ejemplo de aplicar Level-Sets sobre la imagen pre-procesada IM <sub>2</sub> .....	55
Ilustración 46: Ejemplo de aplicar Level-Sets sobre la imagen pre-procesada IM <sub>4</sub> .....	55
Ilustración 47: Level-sets con parámetros $\lambda=0.5$ y $\alpha=8$ (600 iteraciones) .....	56
Ilustración 48: Level-sets con parámetros $\lambda=0.5$ y $\alpha=1$ (600 iteraciones) .....	56
Ilustración 49: Level-sets con parámetros $\lambda=8$ y $\alpha=1$ (600 iteraciones).....	57
Ilustración 50: Umbralización más Level-sets (umbral=0) .....	57
Ilustración 51: Imagen original que usamos para encontrar los centroides óptimos .....	58
Ilustración 52: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.765 (hueso) y 0.55 (fuera del hueso). Tiempo de ejecución: 3.628579 s.....	58
Ilustración 53: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.4 (fuera del hueso). Tiempo de ejecución: 3.551244 s .....	58
Ilustración 54: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.43 (fuera del hueso). Tiempo de ejecución: 4.184355 s.....	59
Ilustración 55: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.51 (fuera del hueso). Tiempo de ejecución: 3.486082 s .....	59
Ilustración 56: Imagen ajustada .....	60
Ilustración 57: Segmentación con Max-flow a partir de la imagen ajustada y con centroides en 0.392 (hueso) y 0 (fuera del hueso). Tiempo de ejecución: 3.662611 s.....	60
Ilustración 58: Segmentación con Max-flow a partir de la imagen ajustada y con centroides en 0.784 (hueso) y 0 (fuera del hueso). Tiempo de ejecución: 3.621492 s.....	60
Ilustración 59: Comparación de técnicas de segmentación. De izquierda a derecha. Imagen original, segmentación por distancia a histograma + Max-flow, y segmentación mediante regiones .....	64

## Índice de tablas

---

Tabla 1: Resultados obtenidos mediante umbralización.....	53
Tabla 2: Resultados obtenidos mediante Level-sets .....	56
Tabla 3: Resultados obtenidos mediante Max-flow .....	61
Tabla 4: Resultados obtenidos mediante métodos estadísticos y Max-flow.....	62
Tabla 5: Resultados obtenidos mediante el método de segmentación por regiones .....	63
Tabla 6: Comparación de resultados obtenidos con los distintos métodos .....	65

## Glosario

<b>ACR</b>	American College of Radiology
<b>CT</b>	Tomografía Computerizada (en inglés, Computed Tomography)
<b>DICOM</b>	Digital Imaging and Communication in Medicine
<b>DRLSE</b>	Distance Regularized Level Set Evolution
<b>FN</b>	False Negative
<b>FOV</b>	Field of View
<b>FP</b>	False Positive
<b>IM</b>	Imagen Médica
<b>LoG</b>	Laplacian of Gaussian
<b>NEMA</b>	National Electrical Manufacturers Association
<b>PAC</b>	Pictures Archiving and Communication System
<b>PET</b>	Tomografía por emisión de positrones (en inglés, Positron Emission Tomography)
<b>PPV</b>	Positive Predicted Value
<b>RM</b>	Resonancia Magnética
<b>SPECT</b>	Tomografía Computerizada por Emisión de Fotón Simple (en inglés, Single-Photon Emission Computed Tomography)
<b>TAC</b>	Tomografía Axial Computerizada
<b>TN</b>	True Negative
<b>TP</b>	True Positive

# Capítulo 1: Introducción

---

En este trabajo se pretende hacer un estudio y comparación de las distintas técnicas que existen actualmente para segmentar automáticamente tejido óseo en imágenes de TAC.

El tejido óseo o hueso es un tejido firme, duro y resistente que forma parte del endoesqueleto de los vertebrados. Está compuesto a la vez por tejidos duros y blandos. Principalmente podemos distinguir dos tipos de tejidos óseos:

- **Hueso compacto:** aparecen como una masa sólida y continua cuya estructura solo se ve al microscopio óptico.
- **Hueso esponjoso o trabecular:** forman una estructura esponjosa dejando huecos llenos de médula ósea roja.

El conjunto total y organizado de las piezas óseas o huesos conforma el esqueleto. Cada pieza cumple una función en particular y de conjunto en relación con las piezas próximas a las que está articulada.

Además de la comparación de técnicas existentes, se ha desarrollado un algoritmo en lenguaje MatLab que segmenta huesos y se ha incluido en la comparación realizada.

Para la realización de este algoritmo, establecimos una serie de características que debía de cumplir:

- El algoritmo debe de ser totalmente autónomo y no debe de necesitar ni parámetros, ni recortes de regiones, ni interacción alguna con la persona que lo use.
- El algoritmo debe de ofrecer los mismos resultados independientemente de la parte del cuerpo que aparezca en las imágenes.
- El algoritmo debe de ser lo más rápido posible, ya que el objetivo es aplicárselo a un gran número de imágenes para poder obtener posteriormente un modelo 3D del cuerpo (o parte del cuerpo) que se esté estudiando.

Además de las características que impusimos, nos encontramos con un problema añadido que se escapa de nuestro control: las enfermedades óseas, que hacen especialmente difícil para un programa detectar el tejido óseo.

Entre estas enfermedades cabe destacar especialmente la **osteoporosis**. La osteoporosis es una enfermedad degenerativa de los huesos que causa una pérdida de densidad y masa ósea, lo que desemboca en un alto riesgo de fractura o micro fracturas.

En la osteoporosis, la densidad de mineral ósea se ve disminuida, las micro-estructuras de los huesos son debilitadas y la cantidad y variedad de proteínas del tejido óseo se

altera. La **densidad ósea** es considerada como una medida de la gravedad de la osteoporosis y se mide mediante una prueba médica que se llama densitometría ósea.

Esta enfermedad se suele dar en mujeres después o durante la menopausia debido a la disminución de los niveles de estrógeno. La pérdida de estrógeno ocasiona una rápida disminución de masa ósea que desemboca en una menor resistencia de los huesos.

Los pacientes con osteoporosis tienen una tasa de mortalidad aumentada debido a la mayor probabilidad de que se produzcan fracturas, sobre todo de cadera o vertebrales.

En las imágenes de TAC, los huesos de los pacientes con osteoporosis aparecen con un número de Hounsfield muy bajo en la parte esponjosa del hueso, conocida también como hueso trabecular. Definir el rango de número de Hounsfield en el que se encuentran es difícil y depende mucho de la modalidad de TAC que se use, pero en cualquier caso es común que los músculos, la grasa u otros tejidos lleguen a tener valores de número de Hounsfield mayores que estas zonas del hueso, lo que imposibilita hacer una segmentación basada en el nivel de gris. Esto se puede apreciar bien en la siguiente imagen, en donde mostramos tres valores de número de Hounsfield y resultan ser mayores en tejido blando que en el hueso trabecular o incluso que en el hueso compacto.

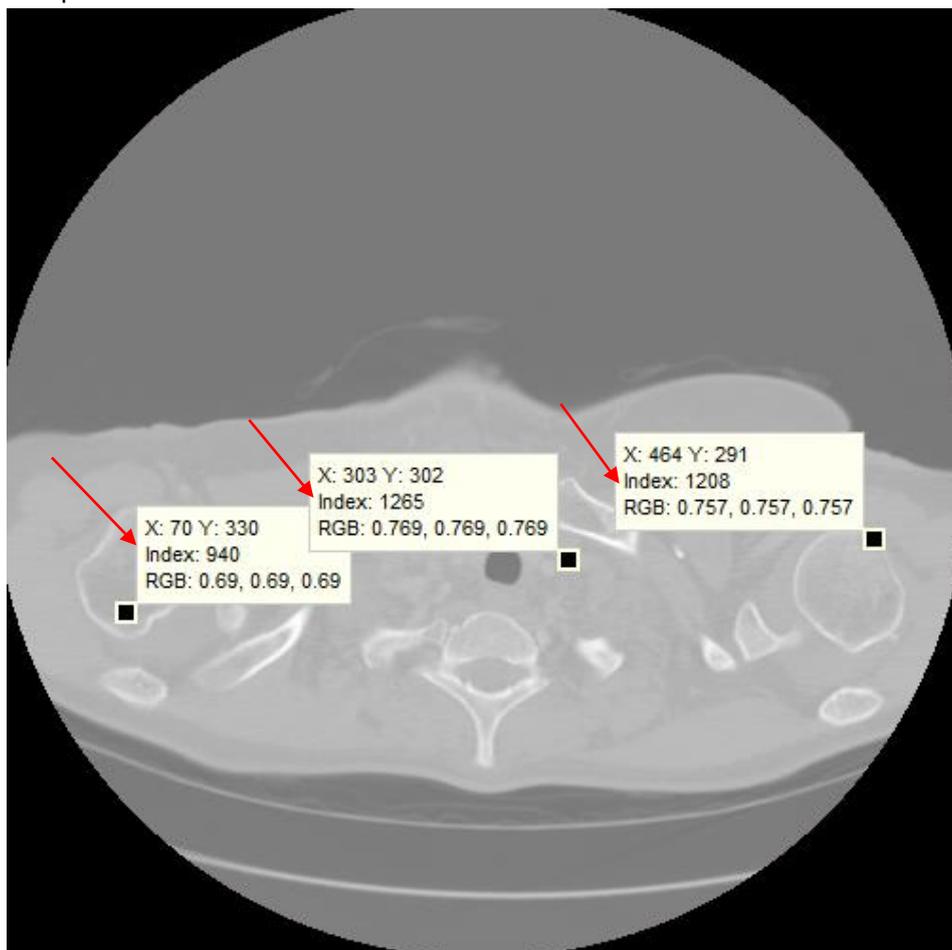


Ilustración 1: Ejemplo de osteoporosis

Existen otras patologías que dificultan enormemente la tarea de segmentación como la osteoartritis o la artritis reumatoide. Estas enfermedades afectan a las articulaciones por la pérdida de cartílago en esas zonas. La pérdida de cartílago conlleva un desgaste de los huesos en las zonas de las articulaciones, el desgaste afecta al hueso compacto que se encuentra en el exterior del hueso, lo que hace que no exista un borde pronunciado del hueso en la imagen.

Como conclusión, a las características ya mencionadas, hay que añadir que nuestro algoritmo debe abarcar un escenario lo más amplio posible, y debe detectar tanto huesos sanos, como los que no lo están.

## Capítulo 2: Las imágenes médicas

---

Las imágenes médicas son aquellas imágenes obtenidas a partir de un conjunto de técnicas y procesos usados para crear imágenes del cuerpo humano o partes de él con propósitos clínicos.

Tradicionalmente, las imágenes médicas se han venido usando para el diagnóstico de enfermedades con la finalidad de obtener información de partes del interior del cuerpo no visibles a simple vista, sin tener que someter al paciente a ninguna intervención quirúrgica. Estas imágenes supusieron un gran avance en la medicina y en el diagnóstico.

Actualmente con el auge de la telemedicina y la posibilidad de consultar con expertos que no están en el lugar del paciente, también se busca obtener imágenes médicas de zonas que sí son accesibles a la inspección visual. Esto facilita el diagnóstico a distancia, y la consulta de segundas opiniones.

Este tipo de imágenes, se pueden tomar con cámaras tradicionales o corrientes, sin la necesidad de tener ningún añadido especial, con el único requerimiento de que ofrezcan una calidad suficiente para el diagnóstico. En este grupo de imágenes se encuentran por ejemplo, las fotografías de quemaduras o fotografías del ámbito dermatológico. A pesar de que a veces nos baste con una cámara de fotos corriente, también existen imágenes dermatológicas que requieren de un equipo especializado como por ejemplo, el dermatoscopio que se usa para observar las lesiones pigmentadas de la piel.

Existen multitud de técnicas para la obtención de estas imágenes y se pueden clasificar en función de características distintas. Por ejemplo, podemos clasificarlas en función de si son imágenes funcionales (PET, RM) que muestran cómo funciona el cuerpo, por ejemplo, observando como digerimos algún alimento, o como un órgano metaboliza alguna proteína, o imágenes anatómicas (TAC) que muestran con gran detalle estructuras anatómicas del cuerpo.

Otra manera de clasificar las imágenes médicas, podría ser en función de las tecnologías que usan cada una de ellas, como puede ser la medicina nuclear, la radiología, las ecografías o la bioluminiscencia.

A continuación describiremos brevemente algunas de las técnicas más importantes en cuanto a la obtención de imágenes médicas.

## 2.1. Radiografía

Una radiografía es una técnica de diagnóstico radiológico. La imagen se obtiene al exponer al paciente a una fuente de radiación de alta energía, normalmente rayos X o rayos gamma.

El descubrimiento de los rayos X se produjo en 1895 a manos de Wilhelm Röntgen investigando las propiedades de los rayos catódicos. Aplicando los efectos de los rayos X a una placa fotográfica, obtuvo la primera radiografía de la historia, en la que aparece la mano de su esposa.



Ilustración 2: Primera radiografía realizada en la historia

Como ya se ha dicho, el fundamento de las radiografías son los rayos X o los rayos gamma. Vamos a centrarnos en los rayos X ya que son los más comúnmente usados. Los rayos X se generan cuando los electrones rápidos (de alta energía) interactúan con la materia. Cuando un electrón con suficiente energía interactúa con un electrón orbital de un átomo, se pueden generar rayos X característicos. Se denominan rayos X característicos porque su energía es determinada por la composición característica del átomo perturbado. Cuando electrones con suficiente energía interactúan con el núcleo de los átomos, se genera radiación Bremsstrahlung (también conocida como radiación de frenado) debida al exceso de energía de los electrones que han sido acelerados y posteriormente frenados. Cuando la energía de los electrones de bombardeo es lo suficientemente alta, esa radiación está en la región de rayos X del espectro electromagnético.

Una vez explicado el fundamento físico, vamos a explicar cómo se producen en un tubo de rayos X. El tubo de rayos X está formado por una carcasa de vidrio (en su interior hay vacío), y en su interior tenemos un alambre de wolframio. En primer lugar, se crea una diferencia de tensión entre dos puntos (ánodo y cátodo). Seguidamente se hace circular una corriente por el alambre de wolframio (situado en el cátodo) que, al alcanzar una determinada temperatura, comienza a emitir electrones. Estos electrones son atraídos por el lado positivo, y debido a la gran diferencia de potencial son fuertemente acelerados hacia el ánodo. Al chocar los electrones con gran energía cinética contra el ánodo se generan los rayos X.

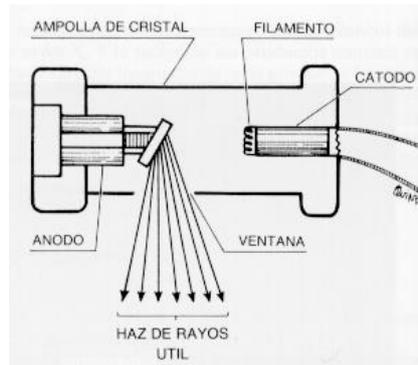


Ilustración 3: Generación de rayos X

Entre las propiedades de los rayos X destacan su efecto fotográfico y que se atenúan de manera distinta dependiendo del grosor, la densidad y la impedancia de la materia que atraviesan. De esta manera, si se han atenuado poco, quemaran la radiografía produciendo un color negro en ella y si se han atenuado mucho, llegarán débilmente a la placa fotográfica y ésta quedará con un color más claro. Este es el motivo por el que los huesos son representados en colores claros y los tejidos blandos con colores oscuros, ya que los huesos atenúan mucho más que los tejidos blandos.

Se puede variar el contraste de la radiografía variando los kilovoltios con los que los electrones son acelerados. En la siguiente gráfica podemos ver la variación del coeficiente de atenuación frente al aumento de kV. Podemos observar que se consigue un mayor contraste con menos kV.

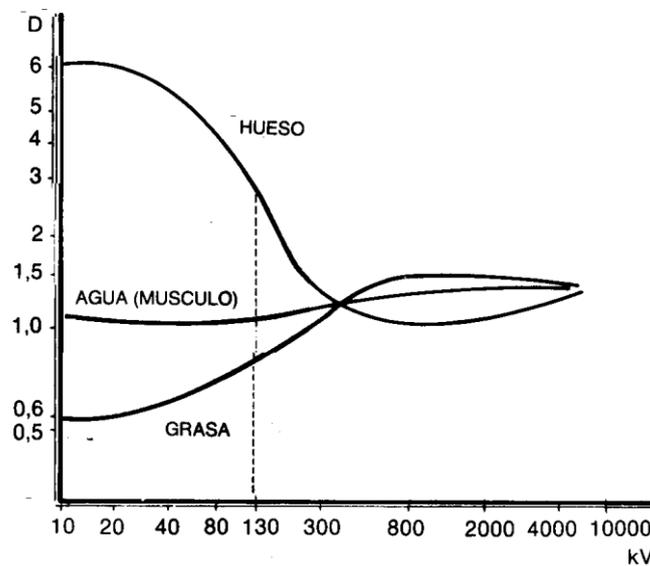


Ilustración 4: Variación de los coeficientes de atenuación frente a kV aplicado

No podemos ignorar el hecho de que los rayos X son rayos ionizantes, lo que quiere decir que alteran las cadenas de hidratos de carbono que tienen las enzimas, proteínas y membranas celulares. Por este motivo son nocivos para la salud, pudiendo provocar

muerte celular u oncogenes<sup>1</sup>. Los tejidos más afectados por los rayos X son la piel (por mayor exposición), la médula ósea, los órganos reproductores, el feto (si la mujer estuviese embarazada) y el aparato digestivo.

En base a la radiografía se han desarrollado distintas técnicas de adquisición de imagen, como las radiografías con contraste (contraste radiológico) en las que se suministra al paciente alguna sustancia radiopaca, como puede ser el yodo o el bario. Estas sustancias son toleradas por el organismo y permiten obtener imágenes funcionales del mismo. Entre las tecnologías más importantes surgidas a partir de las radiografías, se encuentra el TAC, el cual veremos con mayor nivel de detalle en el apartado 2.5 de este capítulo.

## 2.2. Imágenes médicas por ultrasonidos

Las imágenes médicas obtenidas por ultrasonidos, o más comúnmente conocidas como ecografías, es un procedimiento de diagnóstico que emplea el ultrasonido<sup>2</sup> para crear imágenes bidimensionales o tridimensionales. Actualmente también podemos encontrarnos ecografías de cuatro dimensiones (altura, ancho, profundo y tiempo).

Las ecografías se obtienen gracias a un instrumento llamado transductor, que se encarga de emitir ultrasonidos hacia el área del cuerpo humano en estudios, y también es el encargado de recibir el eco o las reflexiones de estas ondas ultrasónica. Una computadora es la encargada de convertir el eco en una imagen.

La ecografía es un procedimiento sencillo, no invasivo (no se emplea radiación) y es por esta razón por la cual es la prueba por excelencia realizada en mujeres embarazadas, ya que las células en división de los fetos son muy susceptibles a todo tipo de radiaciones.

Las ecografías no necesitan de personal técnico especializado, y el equipamiento es relativamente barato en comparación con otras modalidades, lo que hace posible que muchas pequeñas consultas dispongan de ecógrafos.

Aunque están muy avanzadas no ofrecen la resolución espacial que son capaces de proporcionar otras pruebas como los TAC o las RM. Además las ecografías tienen el inconveniente de que los huesos crean sombras tras ellos, por lo que los mejores resultados se obtienen sobre todo en tejidos blandos.

Para más información consultar [1].

---

<sup>1</sup> Un oncogén es un gen anormal o activado que procede de la mutación de un alelo de un gen normal. Los oncogenes son los responsables de la transformación de una célula normal en una maligna que desarrollará un determinado tipo de cáncer.

<sup>2</sup> El ultrasonido es una onda acústica o sonora cuya frecuencia está por encima del umbral de audición del oído humano (aproximadamente 20.000 Hz). En las ecografías se usan frecuencias de entre 2 y 20 MHz.

### 2.3. La resonancia magnética (RM)

La imagen por resonancia magnética (RM) o también conocida como tomografía por resonancia magnética o imagen por resonancia magnética nuclear, es una técnica que utiliza el fenómeno de la resonancia magnética nuclear para obtener información sobre la estructura y composición del cuerpo a analizar. La información obtenida es procesada por ordenadores y transformada en una imagen tomográfica del interior de lo que se está procesando.

Esta técnica se usa principalmente para obtener información de los tejidos y detectar cáncer y otras patologías. No solo tiene uso en el ámbito de la medicina, sino que también es usada industrialmente para analizar la estructura de materiales (orgánicos e inorgánicos).

El fenómeno de la resonancia magnética nuclear se consigue gracias a que en nuestro organismo tenemos millones de dipolos magnéticos. Estos dipolos magnéticos están en todos los átomos con un número impar de protones o neutrones, no obstante, se eligió el hidrógeno para su uso sanitario, por su alta presencia en el cuerpo humano y por su elevada señal de resonancia. El núcleo del hidrógeno (o protón), presenta un movimiento de espín magnético: gira sobre sí mismo, al igual que lo hace la tierra. Todos los núcleos de hidrógeno giran arbitrariamente y de una manera completamente desorganizada, lo que lleva a que el campo magnético producido por todos ellos sea nulo.

El procedimiento que se emplea para obtener la imagen por RM es por tanto el siguiente. En primer lugar, se aplica un campo magnético exterior lo suficientemente fuerte como para hacer que estos dipolos se orienten en la misma dirección que él. Una vez que esto ocurre, se crea por los dipolos otro campo magnético, al que llamaremos  $M_0$ , con la misma dirección que tiene el campo exterior.

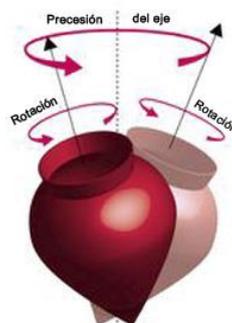


Ilustración 5: Movimientos de precesión y rotación de una peonza

Los dipolos que se encuentran orientados a causa del campo externo, se ponen a precesar<sup>3</sup> como si de una peonza se tratase. La frecuencia a la que giran viene dada mediante la ecuación de Larmor:

$$\omega = \gamma \cdot B \quad (1)$$

A partir de esta ecuación, si conocemos la frecuencia a la que giran, es posible crear un nuevo campo magnético perpendicular al anterior, y si se hace por impulsos de la misma frecuencia que la dada por la ecuación de Larmor, estos giran cada vez más rápido y además giran en fase (de aquí viene el nombre de RM), creando un campo magnético lo suficientemente potente como para ser medido por los receptores. Este campo magnético que llega a los receptores, no es el que nos proporciona toda la información para poder obtener las imágenes. Lo que hace que se pueda diferenciar cada uno de los tejidos del cuerpo humano es el tiempo de relajación, es decir, el tiempo que tarda en desaparecer el campo magnético en cada una de las partes del cuerpo.

La principal ventaja que tiene la RM es que es un método de adquisición no invasivo ya que al usar los campos magnéticos no estamos usando radiación ionizante. Otra ventaja a parte de la seguridad del paciente, es que tiene mejor resolución espacial que las imágenes TAC con lo que a tejidos blandos se refiere, haciendo mucho más fácil por ejemplo, la diferenciación en el cerebro entre masa gris y masa blanca. En contraposición a estas ventajas, podemos decir que la RM es mucho más lenta que otras modalidades de IM, favoreciendo así la aparición de artefactos por movimientos del paciente. Además, los pacientes con cualquier implante metálico, como puede ser un marcapasos, no pueden someterse a este tipo de pruebas.

Esta ha sido una explicación muy breve y se recomienda al lector que tenga más interés sobre este tema que consulte las referencias [2] [3].

## 2.4. Medicina nuclear

La medicina nuclear es una especialidad de la medicina actual en la que se le suministra al paciente una determinada sustancia radioactiva, ya sea por vía oral o intravenosa (entre otras) para poder diagnosticar y determinar la gravedad y tratamiento de varias enfermedades, incluyendo cáncer, enfermedades cardiacas o gastrointestinales, y muchas otras más.

---

<sup>3</sup> La **precesión** o **movimiento de precesión nutación**, es el movimiento asociado con el cambio de dirección en el espacio, que experimenta el eje instantáneo de rotación de un cuerpo. Este es además uno de los cinco movimientos que hace la tierra.

La medicina nuclear ha supuesto una revolución en el estudio funcional de muchos órganos, especialmente el cerebro. Los mecanismos físicos para generar las imágenes son relativamente sencillos de comprender, y pasaremos a explicarlos a continuación.

Como ya hemos comentado antes, se usa una sustancia radioactiva, generalmente conocida como radiofármaco que está formado a su vez por un fármaco transportador y un isótopo radioactivo. Existen varios tipos de radiofármacos (algunos incluso hay que generarlos en un acelerador de partículas) ya que dependiendo de la sustancia a la que se van a adherir, necesitaremos un isótopo radioactivo u otro y esto, a su vez depende del órgano a estudiar.

La imagen se forma detectando la radiación emitida por el radiofármaco y al igual que antes, el tipo de máquina en cuestión depende de la radiación emitida. Actualmente los más usados son:

- **Gammacámara:** Una gammacámara es un dispositivo capaz de detectar la radiación gamma y de formar una imagen bidimensional de la proyección de la distribución espacial de la actividad radiológica. Tiene la desventaja de que se superponen estructuras.
- **SPECT:** Se trata de una gammacámara que gira alrededor del paciente o que simplemente lo rodea y un ordenador capaz de formar una imagen de un corte axial del paciente. El concepto de tomografía se explicará más detalladamente cuando se hable del TAC.
- **PET:** Los conceptos son similares a los del SPECT, pero se diferencia en el tipo de radiofármaco que se usa. Los radiofármacos utilizados en el PET se deben crear en un ciclotrón, mientras que los del SPECT se comercializan. El PET tiene menos resolución espacial, pero tiene la principal ventaja de que además de información funcional, ofrece información metabólica.

Este tipo de imágenes tienen muy poca resolución espacial, por eso se suelen combinar con imágenes radiológicas, como TAC o radiografías.

Para más información consultar [4].

## 2.5. Tomografía Axial Computarizada

Debido a que en este trabajo se ha basado principalmente en imágenes de TAC, esta modalidad la vamos a desarrollar más profundamente que las anteriores.

La tomografía axial computarizada (TAC) o simplemente CT es una técnica de imagen médica que utiliza la radiación X para obtener cortes o secciones de objetos anatómicos con el fin de obtener un diagnóstico. El TAC nos permite crear una imagen transversal del cuerpo humano en vez de una imagen proyección como era el caso de las radiografías. Es decir, en una definición menos formal y para que lo pueda entender

cualquier persona, los TACs nos permiten abrir al cuerpo humano como si fuese un libro y mirar dentro.

Las fórmulas matemáticas para reconstruir una imagen tridimensional a partir de múltiples imágenes axiales planas fueron desarrolladas por el físico Johann Radon, nacido en Austria en 1887. El TAC está basado en la transformada de este científico, la transformada de Radon, que viene dada por la siguiente fórmula:

$$\mathcal{R} = [f](\theta, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos\theta + y \sin\theta - s) dx dy \quad (2)$$

Esta sería la transformada en un caso bidimensional. La interpretación matemática de la transformada de Radon es que nos da la proyección de la función bidimensional sobre la recta que se especifica por el ángulo que tiene con respecto al eje de abscisas  $\theta$  y por la distancia al origen  $s$ .

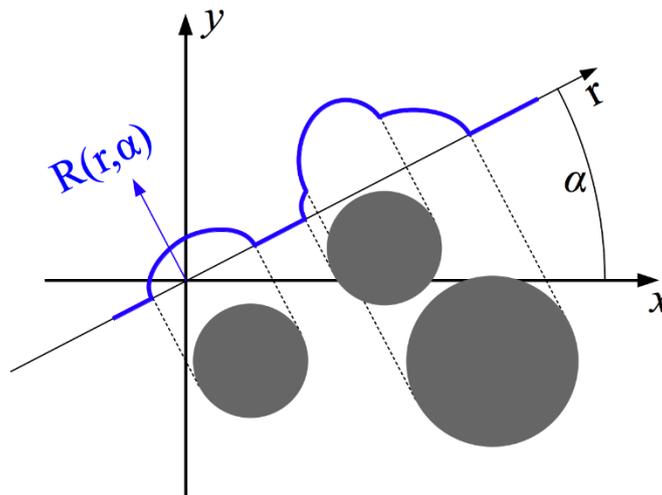


Ilustración 6: Interpretación de la transformada de Radon

Esta transformada en su versión bidimensional y tridimensional, fueron introducidas por primera vez en un artículo en el año 1917. La transformada inversa de radón se calcula mediante el método de las retroproyecciones filtradas.

A pesar de disponer de las transformadas de Radon, no existía el equipo de rayos X ni la máquina capaz de hacer los cálculos, por lo que la medicina tuvo que esperar hasta mediados de los años sesenta, cuando Godfrey Newbold Hounsfield consiguió el primer prototipo aplicable de TAC. Allan M. Cormack publica sus trabajos sobre la TAC en 1967, basándose en el trabajo de Hounsfield, por lo que en 1979 recibieron de forma conjunta el Premio Nobel de Fisiología o Medicina. En honor a Hounsfield, las unidades que definen los distintos valores de densidad de los tejidos en las imágenes de TAC se denominan unidades Hounsfield.

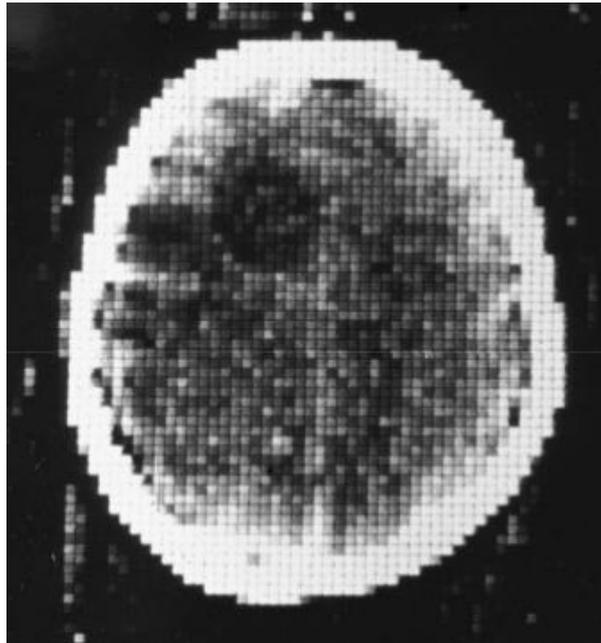


Ilustración 7: Primera imagen de TAC de la historia (1972)

Existe un algoritmo inverso de la transformada de Radon, que se llama algoritmo de las retroproyecciones filtradas. Este algoritmo es computacionalmente eficiente y es capaz de dar valor a un conjunto finito de píxeles, a partir de un número finito de proyecciones, pero que por su complejidad no se detalla en esta memoria.

Una vez explicados los fundamentos matemáticos en los que se basan los TACs, entender el funcionamiento del dispositivo que logra esto es bastante simple.

En un dispositivo de imágenes TAC podemos diferenciar tres elementos principales: el tubo de rayos X, el computador encargado de realizar la reconstrucción, y los detectores encargados de transformar la radiación recibida en señales digitales que el computador pueda procesar.

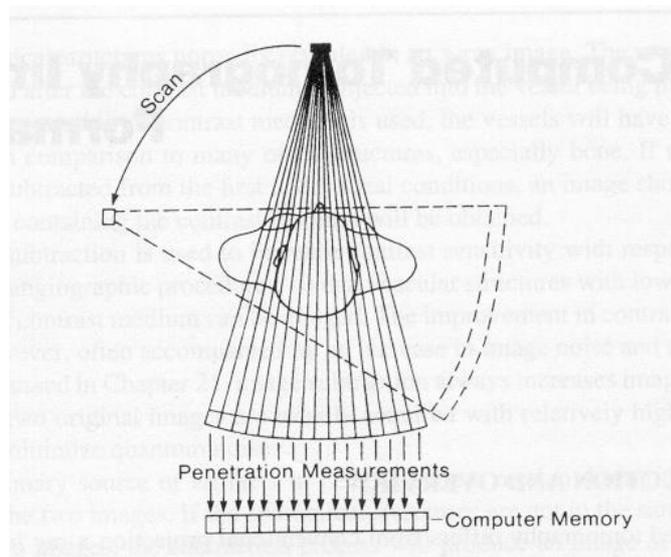


Ilustración 8: Esquema del funcionamiento de un TAC

El tubo emisor de rayos X gira alrededor del paciente. En el lado opuesto se encuentran los detectores que forman un perfil de la radiación que logra atravesar el cuerpo. La radiación que llega a los detectores depende de las densidades de los órganos atravesados, a mayor densidad mayor atenuación.

A continuación mostramos un gráfico en el que se muestran el rango de valores (en unidades Hounsfield) con el que aparecen los distintos tejidos del cuerpo humano en una imagen de TAC.

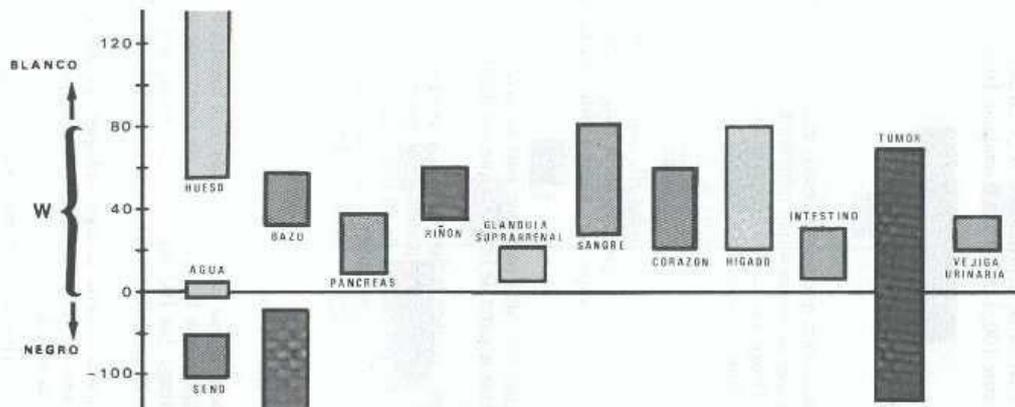


Ilustración 9: Rango de valores Hounsfield de los distintos tejidos del cuerpo

Cada píxel de la imagen representa un elemento de volumen del cuerpo llamado voxel. El tamaño del voxel que representa cada píxel está determinado por el diámetro de la imagen y el tamaño de la matriz.

En la actualidad, la mayoría de los TAC se realizan de manera helicoidal, es decir, el tubo de rayos X va girando de manera continua mientras la mesa sobre la que está el paciente se va desplazando, de manera que al final el tubo de rayos X ha seguido un movimiento en forma de hélice con respecto al paciente. Para conseguir un plano axial se interpolan los datos por encima y por debajo del plano seccionado.

La realización de los TAC helicoidales supone una disminución en el tiempo de adquisición y una disminución de la radiación recibida por el paciente. Además, al reducir el tiempo de adquisición conseguimos que aparezcan menos artefactos por movimientos del paciente. En el apartado 2.5.1 explicaremos con más detalle qué son los artefactos.

Al igual que en el caso de las radiografías, los TACs también se pueden realizar con contraste si se pretende localizar algún tipo de órgano, vías urinarias o vasos sanguíneos, por ejemplo. Durante el desarrollo de este trabajo hemos descartado todo tipo de imágenes que se hayan obtenido con contraste, ya que los contrastes aparecen con niveles de Hounsfield muy altos, por encima incluso que los huesos, por lo que sería

muy difícil que un programa informático diferenciase entre las zonas con contraste y los huesos.

En el desarrollo de este trabajo, y para los resultados obtenidos, se han usado imágenes de TAC adquiridas en el Hospital Virgen del Rocío (Sevilla, España) y por estas dos máquinas en concreto:



Ilustración 10: Lightspeed16, de GE Health



Ilustración 11: Iquilion One, de Toshiba

### 2.5.1. Artefactos en las imágenes TAC

Debido a la frecuencia de aparición de los artefactos en imágenes TAC se ha considerado conveniente describir lo que son y los diferentes tipos que nos podemos encontrar.

Los artefactos en las imágenes de TAC se definen como distorsiones en la imagen radiológica. A la hora de realizar programas automatizados, es muy importante conocer los distintos tipos de artefactos que se pueden dar, ya que así podemos descartar las imágenes que contengan artefactos que imposibiliten la labor del programa, aunque también existen artefactos con los que sí es posible trabajar aunque empeoren los resultados obtenidos al final.

Los artefactos se manifiestan en las imágenes de distintas formas, pudiendo aparecer desde rayas o ruido hasta la distorsión de las formas esperadas. Hay muchos tipos de artefactos: artefactos por movimientos, de "streaking", en cebra, por endurecimiento de haz y artefactos con efecto Cone-Beam. A continuación haremos una breve descripción de algunos de ellos.

El artefacto por **endurecimiento del haz** suele ser el más común y sucede cuando el haz de rayos X atraviesa estructuras densas, que filtran los rayos menos energéticos. Se

pueden identificar porque aparecen líneas negras, que se deben a que no coinciden las lecturas realizadas por los distintos detectores.

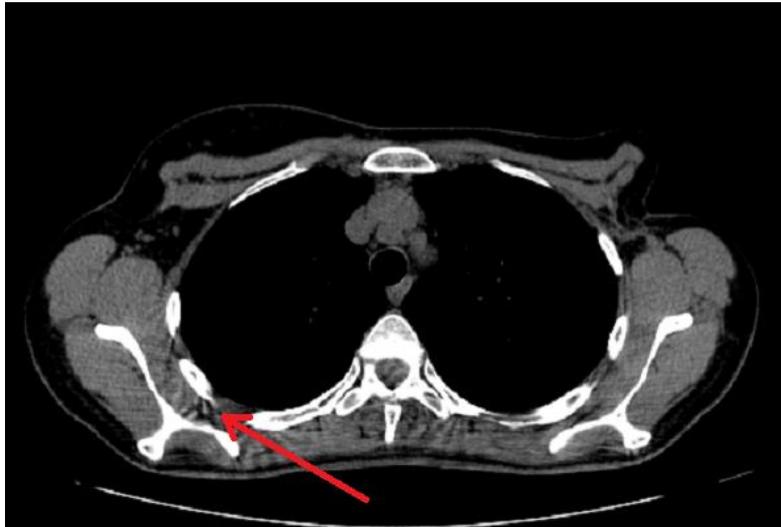


Ilustración 12: Artefacto por endurecimiento del haz

Este tipo de artefactos se pueden evitar colocando adecuadamente al paciente o endureciendo previamente el haz. Otra posibilidad es corregirlo posteriormente mediante software especializado.

El artefacto de **"streaking"** se da cuando existen objetos metálicos en el plano axial del cuerpo escaneado. Este tipo de artefactos son fatales para nuestro detector de huesos, ya que se basa en la detección de bordes, y el resultado de los artefactos de "streaking" son zonas con niveles de Hounsfield muy superiores a los de los huesos, creándonos bordes en forma de estrella. Esto se puede apreciar bien en la siguiente imagen:



Ilustración 13: Ejemplo de Streaking

<sup>4</sup> Streaking: proviene del inglés streak, que significa mancha.

En imágenes con este tipo de artefactos también solemos encontrar artefactos por endurecimiento del haz, que se pueden apreciar en la Ilustración 13.

Los **artefactos en cebra** se identifican porque aparece una especie de ruido de forma escalonada. Es debido a un excesivo grosor del corte en el eje Z. Para nuestro programa este artefacto no será crítico, además de que en los TAC modernos tiende a desaparecer.

Por último, solo nos queda describir los artefactos por movimientos, que normalmente son producidos por el propio paciente al moverse. Dependiendo de la gravedad del artefacto, nuestro programa será capaz de segmentar el tejido óseo o no, pero por lo general vamos a considerar que tenemos imágenes sin artefactos.

## Capítulo 3: Las imágenes digitales

Antes de ahondar en las distintas técnicas de segmentación que actualmente existen, creemos conveniente definir ciertos conceptos que son elementales para entender cómo funcionan y en que se basan las distintas técnicas.

### 3.1. Imagen digital

Una imagen digital es una matriz de datos de tamaño  $(n \times m \times 1)$  si la imagen está en escala de grises o de  $(n \times m \times 3)$  si está en color, donde  $n$  es el número de filas de la matriz,  $m$  el número de columnas y la tercera dimensión representa los planos de color de la imagen, uno, en el caso de imágenes en escala de grises, y tres para las imágenes en color (RGB, HSV, Lab). Cada celda de la matriz  $n \times m$  representa un pixel. Como en este proyecto se trabaja con imágenes en escala de grises no se considera de interés ahondar en el concepto de imágenes en color y sus distintas representaciones. Si el lector desea más información sobre las imágenes digitales se recomiendan las siguientes referencias [5] [6].

El rango de valores que puede tomar cada pixel diferencia varios tipos de imágenes. Si el formato de la imagen es de *uint8*<sup>5</sup>, significa que cada pixel se representa con 8 bits, es decir, cada pixel puede tomar 256 valores distintos de gris. Así mismo, existe el formato *uint16* en el que cada pixel se representa con 16 bits, lo que quiere decir que cada pixel puede tomar 65536 valores distintos. Además, existen dos tipos especiales de imágenes en escala de grises: las **imágenes binarias**, en las que los pixeles toman valores iguales a cero o a uno y las imágenes de tipo *double*, en las que los pixeles se expresan con valores entre cero y uno y con decimales.

### 3.2. El estándar DICOM

El estándar DICOM es el estándar reconocido mundialmente para el intercambio de pruebas médicas, pensado para su manejo, visualización, almacenamiento, impresión y transmisión. La cabecera de este formato, permite almacenar información sobre el paciente, las condiciones en las que se tomó la imagen y el formato interno de ésta, lo que permite que no se desvincule la imagen de la información.

Con la introducción de imágenes médicas digitales en la década de los setenta y el uso de ordenadores para el procesamiento de estas imágenes una vez adquiridas, la mayoría de los dispositivos almacenaban las imágenes en un formato propietario y transferían estos formatos a través de una red o dispositivos de almacenamiento portátiles. Esto ocasionaba en multitud de ocasiones problemas de compatibilidad, por lo que el ACR y el NEMA formaron un comité conjunto para crear un método estándar para la transmisión de imágenes médicas y la información asociada a ella. Este comité,

---

<sup>5</sup> Formatos empleados en el entorno de trabajo MATLAB, que será el que usemos en este trabajo.

formado en 1983, publicó en 1985 el estándar ACR-NEMA [7]. El nombre de DICOM surgió con la versión 3.0 de este estándar. Actualmente se encuentra por la versión PS3.1 2014a [8].

DICOM permite la integración de escáneres, servidores, estaciones de trabajo, impresoras, y hardware de red de múltiples proveedores dentro de un sistema de almacenamiento y comunicación de imágenes, por lo que el estándar no es sólo un formato de fichero para imágenes médicas sino que pretende ser un estándar completo que cubra todas las necesidades de un PACS: almacenamiento, pero también transmisión, comunicaciones en general e impresión. Para lograr esto el estándar especifica:

- Conjunto de protocolos para las comunicaciones de red que deben ser seguidos por los dispositivos conformes al estándar.
- Sintaxis y semántica de los comandos y la información asociada que puede ser intercambiada usando estos protocolos.
- En cuanto a la comunicación de datos, se define un conjunto de servicios para el almacenamiento de datos que deben ser seguidos por los dispositivos conformes al estándar, así como un formato de fichero y una estructura de directorio médico para facilitar el acceso a las imágenes y a la información relacionada.

### 3.1.1. Estructura de un fichero DICOM

El formato de fichero DICOM es muy complejo, ya que en la cabecera se especifican una gran cantidad de campos. Además, existen varios tipos de cabecera y la imagen se puede grabar en multitud de formatos distintos.

El fichero DICOM se puede dividir en cuatro partes diferenciadas:

- Preámbulo y prefijo identificativo del fichero
- Meta-cabecera.
- Cabecera.
- Imagen (aunque desde el punto de vista del formato, la imagen es un elemento más de la cabecera).

El preámbulo tiene un tamaño fijo de 128 bytes, y está pensado para tener un uso definido por la implementación. El preámbulo puede contener información sobre el nombre de la aplicación usada para crear el fichero, por ejemplo. En cuanto al prefijo consiste en cuatro bytes que contienen la cadena de caracteres DICOM. Esta cadena debe estar codificada siempre con las letras en mayúscula y usando el repertorio de caracteres ISO 8859 Go. El propósito de este prefijo es permitir a las implementaciones diferenciar si un fichero es DICOM o no.

En cuanto a la cabecera y la meta-cabecera de un fichero DICOM, consisten en una serie de campos con toda la información necesaria sobre la imagen en cuestión, incluyendo la propia imagen.

Entre estos campos podemos encontrar datos sobre el paciente (nombre, sexo), y sobre el tipo de imagen entre otros. Al conjunto de toda la información codificada sobre un campo se le conoce con el nombre de Elemento de Datos (Data Element).

En un elemento de datos podemos diferenciar los siguientes campos:

- Etiqueta del Elemento de Datos (Data Element Tag): sirve para identificar cada elemento de datos de forma unívoca. Esta etiqueta está constituida por un número de grupo y un número de elemento.
- Representación del Valor (Value Representation, abreviado VR): indica la forma en que se codifica el valor del elemento. Por ejemplo el valor puede estar codificado como una cadena de caracteres o un entero sin signo.
- Longitud del Valor (Value Length): como su nombre indica, es la longitud del campo valor.
- Valor (Value): es el valor del elemento de datos, codificado según el campo VR y con la longitud que indica el campo longitud del valor.

## Capítulo 4: Procesamiento de imágenes

En este apartado se van a detallar las técnicas usadas a lo largo de este trabajo. Para una descripción más detallada de otras técnicas existentes se recomienda al lector acudir a [5] [6] [9].

### 4.1. Técnicas basadas en histograma

En estadística, un histograma es una representación gráfica de una variable en forma de barras, donde cada barra es proporcional a la frecuencia de los valores representados, ya sea de forma diferencial o acumulada.

Por lo tanto, el histograma de una imagen nos muestra con qué frecuencia aparece cada valor presente en la imagen. Se puede hacer una representación gráfica del histograma o simplemente evaluarlo numéricamente, pero es de gran utilidad a la hora de hacer transformaciones en los valores de grises de la imagen. Veamos un ejemplo de una imagen TAC junto con su histograma:

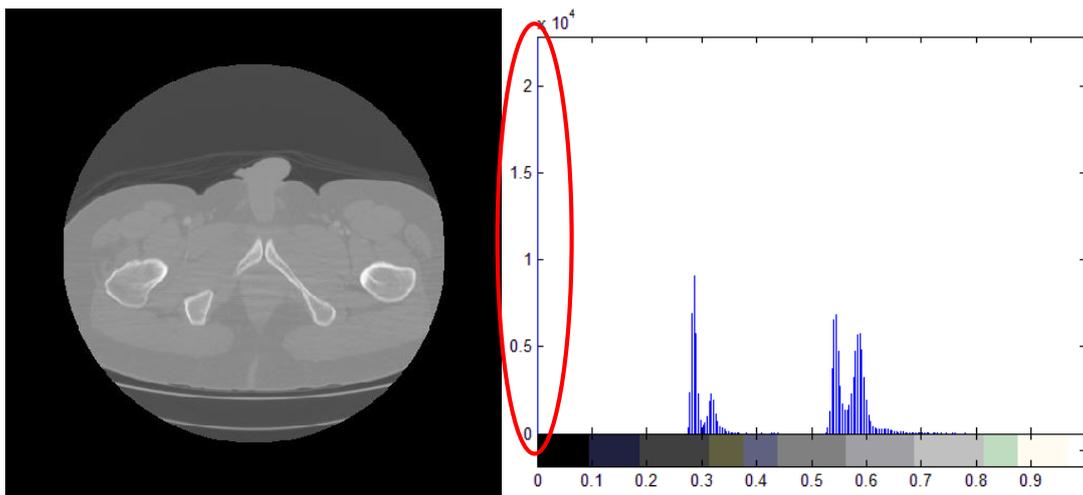


Ilustración 14: Representación de una imagen de TAC junto con su histograma

Como vemos, el valor que más veces se da en la imagen es el cero, ya que corresponde a todos los píxeles que están fuera del FOV.

A partir del histograma podemos determinar que transformaciones aplicar a la imagen, por ejemplo, modificar el contraste de la imagen para apreciar más diferencias en distintos niveles de gris, lo cual nos va a resultar de gran utilidad en el tratamiento de imágenes.

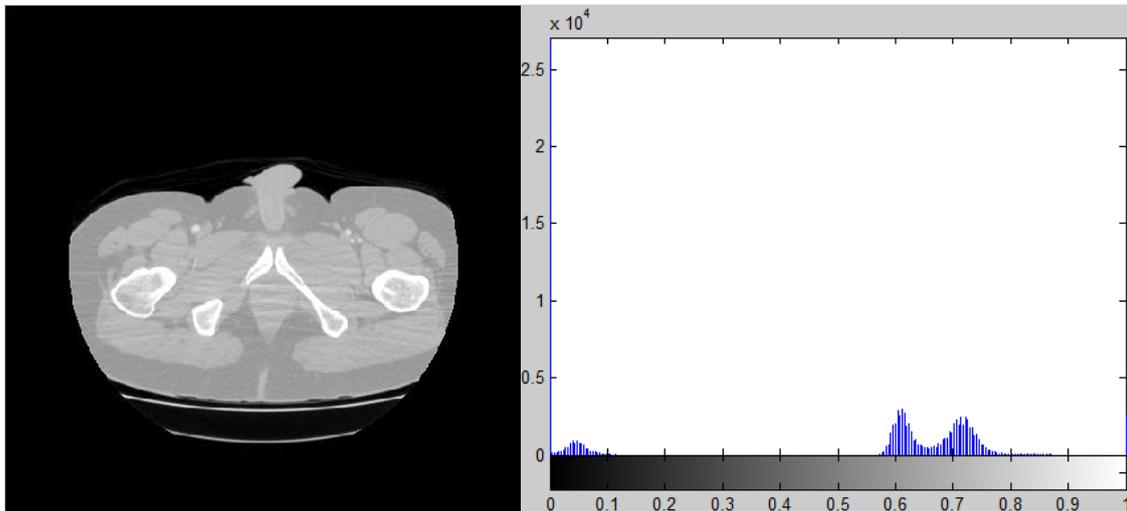


Ilustración 15: Misma imagen que la Ilustración 14 tras transformación de histograma

## 4.2. Matemáticas morfológicas

La Morfología matemática es una teoría y técnica para el análisis y tratamiento de las estructuras geométricas, basada en la teoría de conjuntos, teoría de retículos, topología, y funciones aleatorias. Las matemáticas morfológicas se aplican normalmente a imágenes digitales, pero también puede ser empleada en gráficos, mallas poligonales, sólidos y otras estructuras espaciales.

Las matemáticas morfológicas engloban un sinnúmero de operaciones, pero en este apartado solo nos centraremos en un pequeño grupo de operaciones que consideramos de gran interés y entre las cuales, algunas han sido de utilidad en el desarrollo del programa para detectar hueso. Estas operaciones se pueden aplicar tanto a imágenes en escala de grises como a imágenes binarias. En este trabajo solo se realizan operaciones morfológicas a imágenes binarias, por lo que nos ceñiremos a explicar el efecto de las operaciones únicamente en este caso.

Antes de explicar en qué consisten la dilatación, erosión, apertura y cierre, es conveniente nombrar que estas operaciones más complejas están basadas en operaciones **básicas**, que son las siguientes: NOT, AND, OR y XOR.

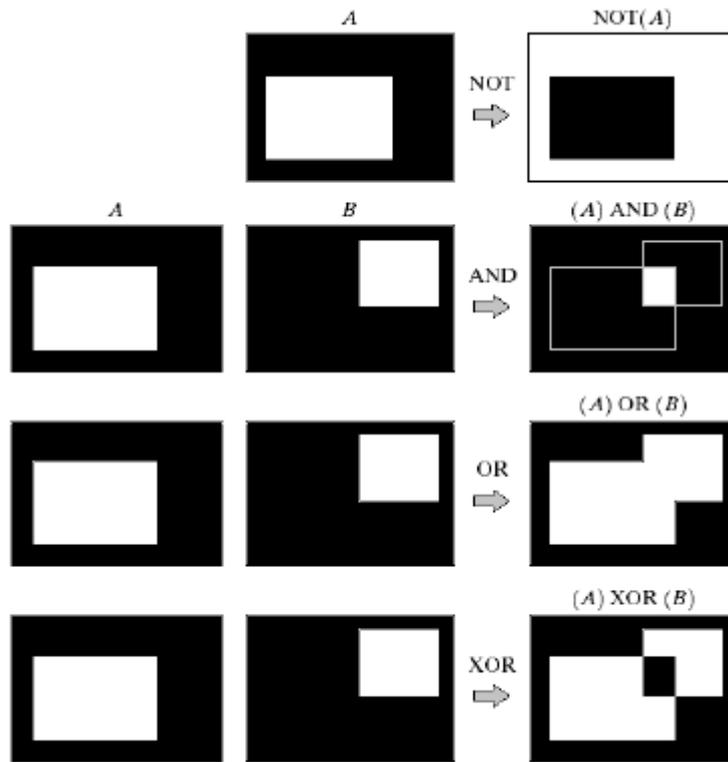


Ilustración 16: Operaciones morfológicas básicas

- Dilatación

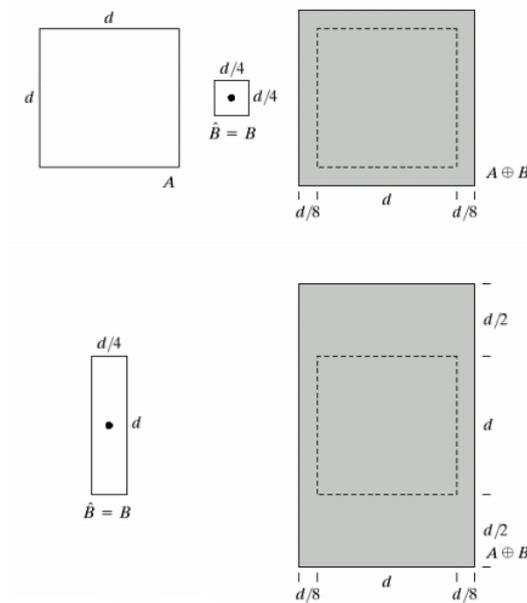


Ilustración 17: Dilatación

Para realizar una operación de dilatación necesitamos dos elementos: por un lado la imagen binaria (a la que llamaremos A), y por otro lado la máscara (que llamaremos B). La máscara B se desplaza por toda la imagen A, colocándose sobre cada uno de los píxeles. En cada desplazamiento se realiza una operación AND, y si en el resultado de

esta operación AND hay algún elemento a uno, el pixel a evaluar pasa a valer uno. Dependiendo de la máscara que se use, se pueden conseguir distintos efectos. Se denota como  $A \oplus B$ , y como su nombre indica dilata los objetos de la imagen.

- **Erosión**

Al igual que en el caso de la dilatación, necesitamos la imagen binaria (A) y la máscara (B). La máscara se desplaza por toda la imagen A, colocándose sobre cada uno de los píxeles. En cada desplazamiento se realiza una operación OR, y si en el resultado de esta operación OR hay algún elemento a cero, el pixel a evaluar pasa a valer cero. Aquí también dependemos de la máscara que se use, y en función de esto, se pueden conseguir distintos efectos. Se denota como  $A \ominus B$ , y como su nombre indica erosiona los objetos de la imagen.

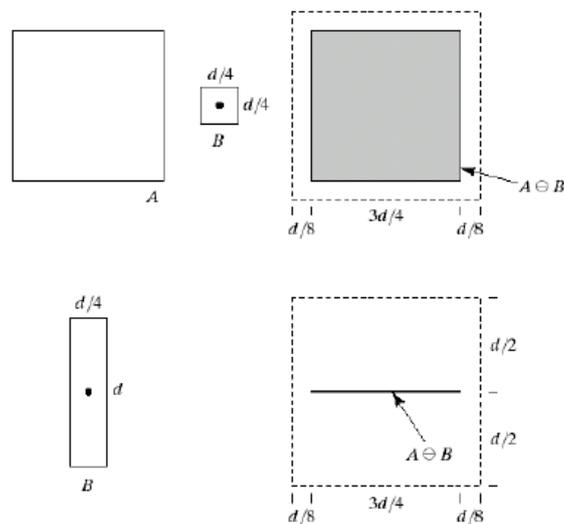


Ilustración 18: Ejemplo de erosión

- **Apertura**

Se trata de una erosión, seguida de una dilatación. Tiene la propiedad de romper uniones estrechas y eliminar salientes.

- **Cierre**

Se trata de una dilatación, seguida de una erosión. Tiene la propiedad eliminar entrantes y pequeños huecos, y de fusionar separaciones estrechas.

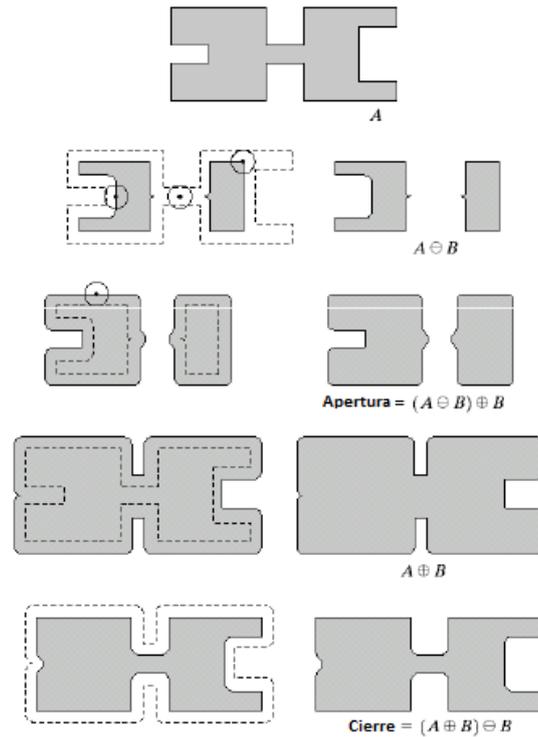


Ilustración 19: Ejemplo de apertura y cierre

**- Rellenado de regiones**

Aunque el relleno de regiones no es una operación morfológica como tal, sino un algoritmo morfológico, es una parte fundamental del programa que se desarrolla en este proyecto, por lo que consideramos que es necesaria su explicación.

- Se parte de un conjunto A de puntos 8-conectados<sup>6</sup> que encierran una región que se desea rellenar.
- Se asigna el valor uno a un punto  $p$  interior al contorno A ( $X_0 = p$ )
- Se realiza la dilatación condicional dada por la fórmula siguiente con un elemento estructurante  $3 \times 3$  de conectividad-4 (cruz):  $X_k = (X_{k-1} \oplus B) \cap A^c$
- El algoritmo termina cuando  $X_k = X_{k-1}$



Ilustración 20: Ejemplo de relleno de regiones

<sup>6</sup> 8-conectados se refiere a que en una vecindad de 9 píxeles ( $3 \times 3$ ) el píxel central está conectado con cada uno de los píxeles vecinos (conectado con 8 en total).

### 4.3. Detección de bordes

La detección de bordes tiene como objetivo la identificación de puntos en una imagen digital que tiene discontinuidades y es fundamental en el tratamiento de imágenes, por lo que es importante conocer tanto los sistemas más simples como otros algoritmos más complejos.

Existen dos principales familias de detectores de bordes: por un lado están los que se basan en el gradiente, es decir la primera derivada; y por otro lado, los que se basan en la Laplaciana, que se corresponde con la segunda derivada.

La diferencia entre los distintos detectores de borde la marca la máscara usada para calcular las derivadas. De ese modo nos encontramos con que los detectores que se basan en el gradiente tienen una máscara para calcular el gradiente sobre el eje X y otra para calcularlo sobre el eje Y, mientras que los detectores que se basan en la Laplaciana usan una única máscara.

A los detectores se les puede incluir además un umbral. Este umbral marca el valor por debajo del cual se considera que el pixel que se está evaluando no forma parte de un borde. Esto puede resultar de mucha utilidad para eliminar falsos positivos, que son pixeles que se consideran como bordes pero que realmente no lo son, y para eliminar ruido en el resultado final.

A continuación, vamos a explicar con más detalles los distintos tipos de máscaras que se usan en el cálculo del gradiente:

- **Sobel:** este detector se puede ver como la combinación de un filtro de suavizado de ruido con un operador de gradiente. Utiliza una máscara de 3x3, que se mueve píxel a píxel, calculando el valor del gradiente para cada uno de ellos. Como ya se ha explicado anteriormente, una vez obtenido el valor del gradiente se decide si es un borde o no en función de un umbral prefijado. Las máscaras usadas para el cálculo del gradiente en el eje x y en el eje y son, respectivamente:

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (3)$$

- **Prewitt:** Es un operador similar al de Sobel, pero se diferencia en que Prewitt no enfatiza los píxeles cercanos al centro de la máscara. Las máscaras para calcular el gradiente en el eje x, y el gradiente en el eje y, son las siguientes:

$$G_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (4)$$

- **Robbets:** Este es el operador más simple, aunque no ofrece tan buenos resultados como los anteriores, pero a cambio ofrece menor carga computacional. Las máscaras que se usan en el cálculo del gradiente en el eje x y en el eje y son:

$$G_x = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (5)$$

A diferencia de los detectores basados en gradiente, los detectores basados en la Laplaciana son generalmente algoritmos, por lo que aumentan considerablemente la carga computacional. Además, las máscaras que se usan tienen por lo general dimensiones mayores que las usadas para el cálculo del gradiente, dando lugar a veces a que aparezcan bordes dobles. No todo son inconvenientes: el cruce por cero permite una mejor localización del borde, además dependen menos de la orientación del borde. Veamos los algoritmos principales:

- **LoG:** Este método de detección de bordes fue propuesto por Marr and Hildreth, que introdujeron el principio de detecciones mediante el método de cruces por cero. Este método consiste en encontrar las coordenadas de una imagen donde la segunda derivada toma el valor cero, mientras que los valores a su alrededor son altos.

El operador LoG también es sensible al ruido. Sin embargo, los efectos de dicho ruido pueden ser reducidos si se ignoran los cruces por cero producidos por pequeños cambios en la intensidad de la imagen. Además, este operador nos proporciona información sobre la dirección de los ejes, determinada mediante la dirección del cruce por cero.

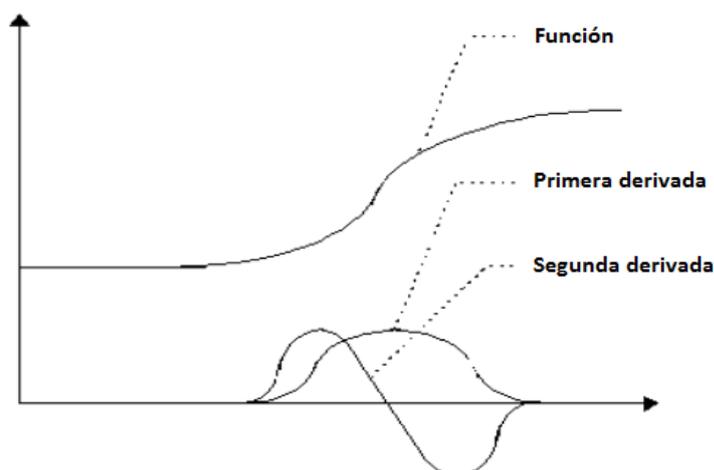


Ilustración 21: Detección de bordes por cruces por cero

La máscara que suele usar este operador es la siguiente:

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix} \quad (6)$$

- **Canny:** Este operador se basa en tres pasos: primero suaviza la imagen para eliminar ruido, posteriormente calcula el gradiente, y selecciona los máximos locales (de esta manera se evitan bordes anchos) y para finalizar, elimina los bordes creados por ruido.

La principal ventaja de este algoritmo es que es muy rápido y además es muy robusto contra el ruido, pero esa es su principal desventaja también, porque en ocasiones elimina bordes reales de la imagen.

A continuación, mostramos una comparación de las distintas imágenes de bordes obtenidas con los métodos que hemos descrito. Como se puede observar, los mejores resultados se obtienen con los dos últimos detectores.

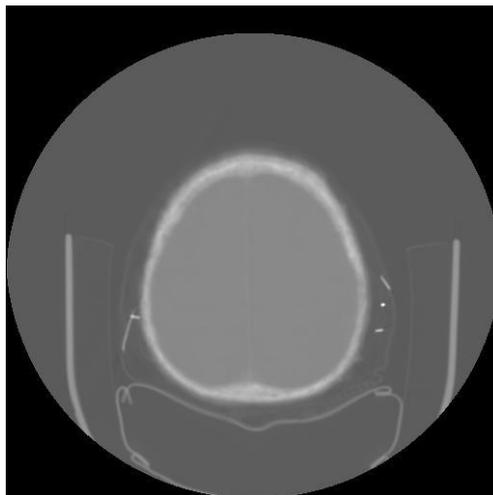


Ilustración 25: Imagen para comparar detectores de bordes

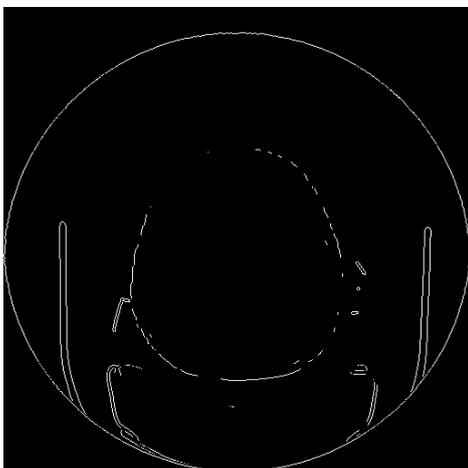


Ilustración 22: Detector de bordes Sobel

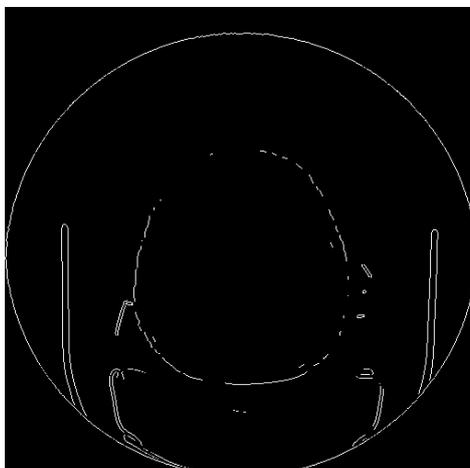


Ilustración 23: Detector de bordes Prewitt

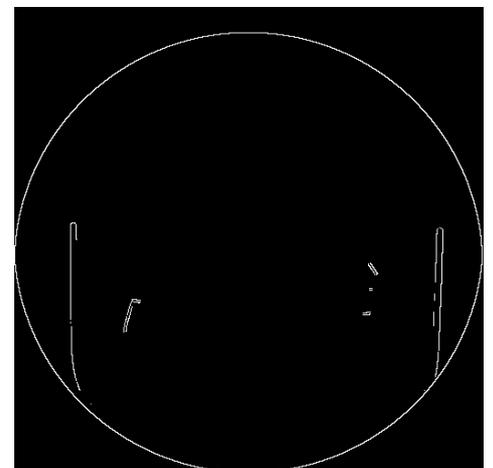


Ilustración 24: Detector de bordes Roberts

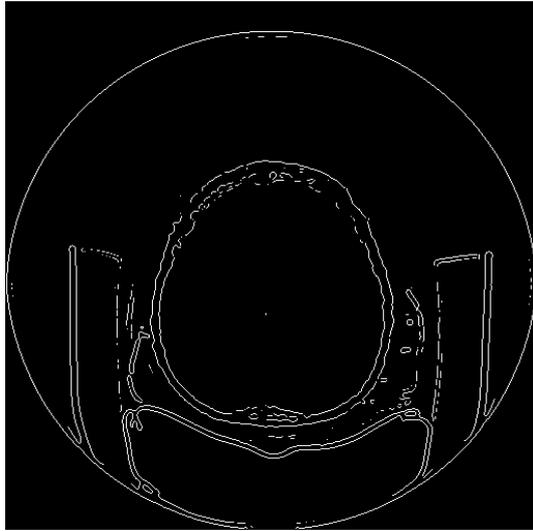


Ilustración 26: Detector de bordes LoG

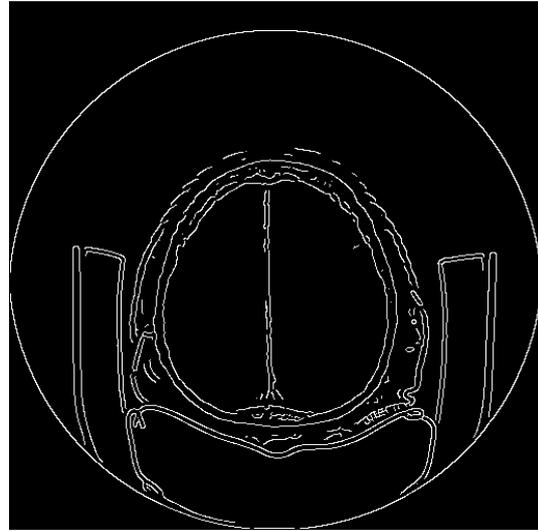


Ilustración 27: Detector de bordes Canny

## Capítulo 5: Técnicas de segmentación

Como ya se indicó en la introducción, en este trabajo, además de desarrollar un programa informático para la detección de huesos en imágenes de TAC, se ha realizado una comparación de las principales técnicas que existen en la actualidad para la segmentación en imágenes médicas, para así poder determinar si los resultados obtenidos con el programa desarrollado son buenos o no.

Antes de realizar la comparación vamos a explicar en qué consisten cada una de las técnicas.

### 5.1. Umbralización

Este método es el más básico y rápido y se ha decidido incluir porque en personas sanas y en determinadas zonas del cuerpo es capaz de dar resultados muy buenos.

La umbralización (en el ámbito de la segmentación de hueso) consiste en fijar un umbral a partir del cual el pixel evaluado se considera como hueso, ya que los huesos presentan generalmente valores altos en la escala de grises.

Un algoritmo muy conocido es el método de Otsu que usa el histograma de la imagen para calcular el valor umbral a partir del cual segmentar. Para hacerlo trata de encontrar un umbral que haga que la varianza en cada zona del histograma sea mínima, y que a la vez la varianza entre las dos zonas segmentadas sea máxima. Este método se suele aplicar a imágenes con histogramas bimodales (histogramas que tienen sus valores concentrados alrededor de valores separados entre sí), ya que el umbral suele estar en el valle entre los dos modos (picos del histograma).

El problema que presenta la segmentación mediante el método de Otsu en imágenes de TAC, es que el histograma de estas imágenes presenta más zonas que un histograma bimodal: la zona de fuera del FOV que tiene un peso importante en el histograma, la zona del FOV en la que no se halla el paciente y por último tenemos la zona que representa al paciente. Si sobre una imagen de TAC aplicamos este método directamente, nos separa la zona de dentro del FOV y la de fuera, y si procesamos la imagen, siempre nos va a umbralizar con valores cercanos a cero, por lo que el uso de este algoritmo en estas imágenes no nos produce los resultados deseados.

### 5.2. Level-sets

Los level-sets son una técnica numérica para el encontrar formas y figuras en imágenes. La principal ventaja que se les atribuye es que no usan la parametrización de la curva, por lo que se pueden adaptar a formas difícilmente parametrizables y además no dependen de la inicialización de la curva. Gracias a esa ventaja también pueden

dividirse, encontrar y separar varias figuras en una misma imagen, lo cual no es posible con la utilización de curvas paramétricas.

Los level-sets tratan de encontrar el nivel de mínima energía en un espacio de tres dimensiones. La curva  $C(x, y)$  en un principio es establecida por el usuario y representa la intersección del plano de nivel de energía cero con la función de tres dimensiones  $\phi$  que va evolucionando hasta alcanzar una zona estable de energía media mínima.

$$C(x, y) = \phi(x, y, z = 0) \tag{7}$$

La energía a minimizar se puede expresar en función de  $\phi$  como:

$$E_\phi = \int_{\Omega} L(x, y) \delta(\phi(x, y)) |\nabla \phi| dx dy \tag{8}$$

Donde  $L(x, y)$  es una función que toma valores bajos en los bordes y altos en las zonas homogéneas.

Para llegar al resultado final, se hace evolucionar la curva en la dirección de máximo descenso de energía:

$$\frac{\partial \phi}{\partial t} = \frac{-\partial E_\phi(\phi)}{\partial \phi} \tag{9}$$

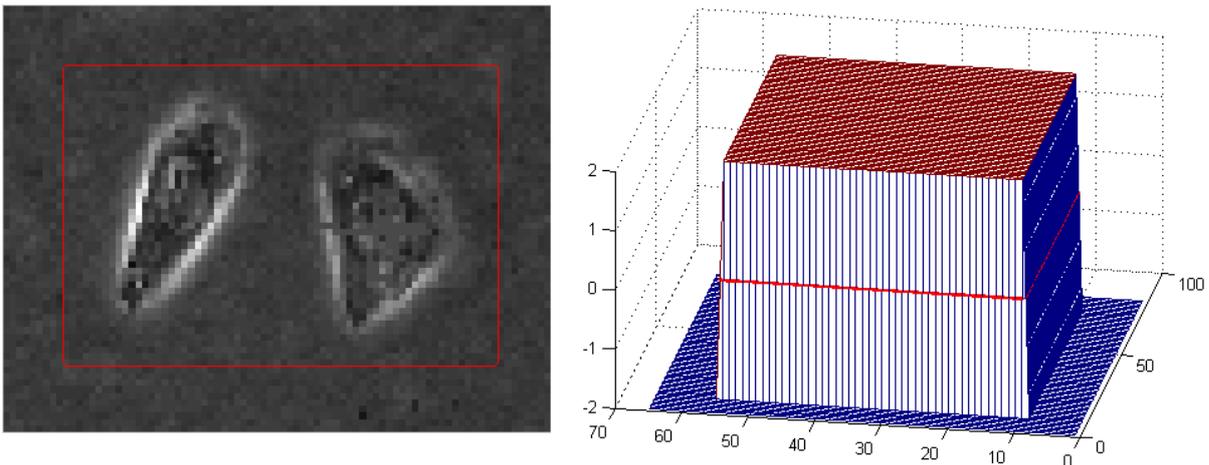


Ilustración 28: Ejemplo de inicialización de level-sets

Finalmente obtendríamos un resultado como el que se muestra en la siguiente ilustración, en el que podemos ver que la curva ha evolucionado hasta rodear completamente los dos objetos de la imagen.

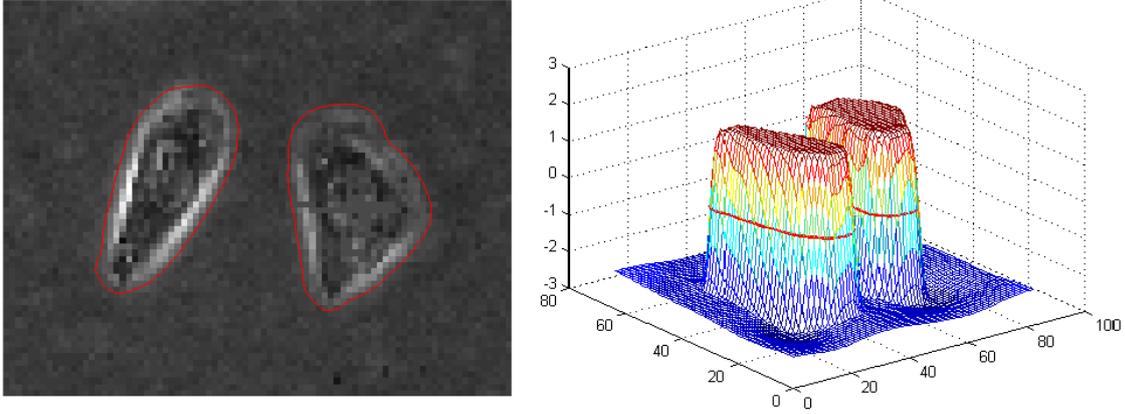


Ilustración 29: Resultado final de una segmentación usando level-sets

En este caso se ha trabajado con un algoritmo basado en level-sets que se denomina Distance Regularized Level Set Evolution (DRLSE) desarrollado por Chunming Li en [10]. El algoritmo de Chunming Li dispone de varios parámetros configurables, de los cuales solo vamos a usar los siguientes: el parámetro que hace variar el área ( $\alpha$ ) y el que hace variar el perímetro ( $\lambda$ ). Dependiendo del signo del parámetro  $\alpha$  la región en cuestión tiende a disminuir o a aumentar. En concreto, la función de energía que se va a minimizar en este caso es la siguiente:

$$\varepsilon_\varepsilon(\phi) = \mu \int_{\Omega} p(|\nabla\phi|)dx + \lambda \int_{\Omega} g\partial_\varepsilon(\phi)|\nabla\phi|dx + \alpha \int_{\Omega} gH_\varepsilon(-\phi)dx \quad (10)$$

Donde:

$$\partial_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon} \left[ 1 + \cos\left(\frac{\pi x}{\varepsilon}\right) \right], & |x| \leq \varepsilon \\ 0, & |x| > \varepsilon \end{cases} \quad (11)$$

$$H_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon} \left[ 1 + \frac{x}{\varepsilon} + \frac{1}{\pi} \sin\left(\frac{\pi x}{\varepsilon}\right) \right], & |x| \leq \varepsilon \\ 1, & |x| > \varepsilon \\ 0, & |x| < -\varepsilon \end{cases} \quad (12)$$

$$g \triangleq \frac{1}{1 + |\nabla G_\sigma * I|^2} \quad (13)$$

$$p(s) = \begin{cases} \frac{1}{2\pi} \sin(2\pi s), & \text{si } s \leq 1 \\ s - 1, & \text{si } s \geq 1 \end{cases} \quad (14)$$

En la ecuación (13),  $g$  es un indicador de borde,  $G_\sigma$  es un kernel gaussiano de desviación estándar  $\sigma$  e  $I$  es la imagen a tratar.

El parámetro  $\mu$  de la ecuación (9) es el coeficiente de regularización de distancia, que va a ser fijado a 0,2.

Por otro lado este algoritmo es bastante lento si la imagen es grande, porque en cada iteración tiene que calcular el gradiente de cuatro imágenes distintas del mismo tamaño que la imagen a segmentar y normalmente hacen falta al menos 500 iteraciones para conseguir un buen resultado. Además tampoco resolvería el problema de los huecos estrechos entre los huesos o cavidades dentro de los huesos, que es uno de los principales problemas a la hora de detectar hueso.

A pesar de estos inconvenientes y como veremos en las comparaciones, se consiguen buenos resultados si solamente se emplean en pequeñas zonas o recortes de la imagen.

### 5.3. Segmentación basada en Max-flow

Las técnicas de Max-flow consisten en crear una red con tantos nodos como píxeles, más un número de nodos que será igual a la cantidad de regiones que queremos separar en nuestra imagen. Cada nodo se conecta con sus vecinos y con los nodos destino y se asigna un coste a cada enlace. El algoritmo resuelve un problema de minimización de costes de tal manera que cada píxel queda unido a un único nodo destino y trata de maximizar el flujo entre el nodo *source* (fuente) y el nodo *sink* (sumidero) denotado como  $t$  en la Ilustración 30 .

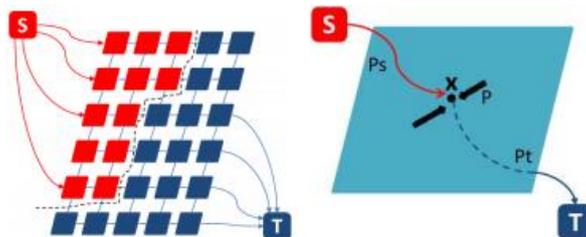


Ilustración 30: Inicialización de Max-flow. Figura extraída de [11]

En concreto, el código que se ha usado forma parte de una adaptación a Matlab que han desarrollado Andrew Delong y Anton Osokin usando los códigos de Max-flow en lenguaje C++ que implementan el algoritmo detallado en [11] [12].

En nuestro caso queremos obtener dos regiones: la región que va a representar al hueso puede ser caracterizada por el nodo  $s$  y la región que va a representar al resto de la imagen puede ser representada por el nodo  $t$ . Nos encontramos con que cada posición de la imagen tiene tres enlaces, el enlace a  $s$ , el enlace a  $t$  y el enlace propio de su posición que será  $p(x)$ , estos tres enlaces están delimitados por las capacidades:

$$p_s(x) \leq C_s(x), p_t(x) \leq C_t(x), |p(x)| \leq C(x) \forall x \in \Omega \quad (15)$$

Donde  $C_s$  es la capacidad de la fuente (hueso),  $C_t$  es la capacidad del sumidero (parte de la imagen que no es hueso) y  $C(x)$  es la capacidad del punto  $x$ . El problema consiste en maximizar el flujo total de la fuente hasta el sumidero.

$$\max_{p_s, p_t, p} \int_{\Omega} p_s dx \quad (16)$$

En el algoritmo de Yuan se propone que este problema es equivalente al siguiente problema de minimización:

$$\min_{u(x) \in [0,1]} \int_{\Omega} (u)C_s dx + \int_{\Omega} (1-u)C_t dx + \int_{\Omega} C(x)|\nabla u| dx \quad (17)$$

En realidad esta ecuación representa el modelo dual de la anterior. Si el problema de minimización está bien definido,  $C_s$  debería tomar valores bajos en el interior del hueso y valores altos fuera del hueso, y de la misma manera  $C_t$  debería tomar valores bajos en el tejido blando y valores altos en el tejido óseo.

## 5.4. Segmentación de estructuras óseas basada en métodos estadísticos y Max-flow

Este método ha sido desarrollado en [13]. En este método se empieza calculando el histograma local de cada pixel usando una vecindad  $21 \times 21$ . Se calcula la distancia (con prueba  $\chi^2$ ) de este histograma local a un histograma referencia de hueso. Tras finalizar este proceso para toda la imagen podremos formar una imagen distancia al histograma modelo. A esta imagen la podemos llamar *HDI* (Histogram Distance Image). Esta distancia será próxima a cero en las zonas con tejido óseo, y será alta en las zonas de tejido blando.

Para combinar la información estadística con la información de niveles de gris, creamos la imagen *CIT* (Cost Image Term), donde  $I_{norm}$  es la imagen umbralizada y normalizada entre 0 y 1:

$$CIT = \frac{((1 - I_{norm}) + HDI)}{2} \quad (18)$$

Se usa el termino  $(1 - I_{norm})$  en vez de  $I_{norm}$  porque nos interesan valores bajos en el tejido óseo, que son los que tienen el nivel de gris más alto.

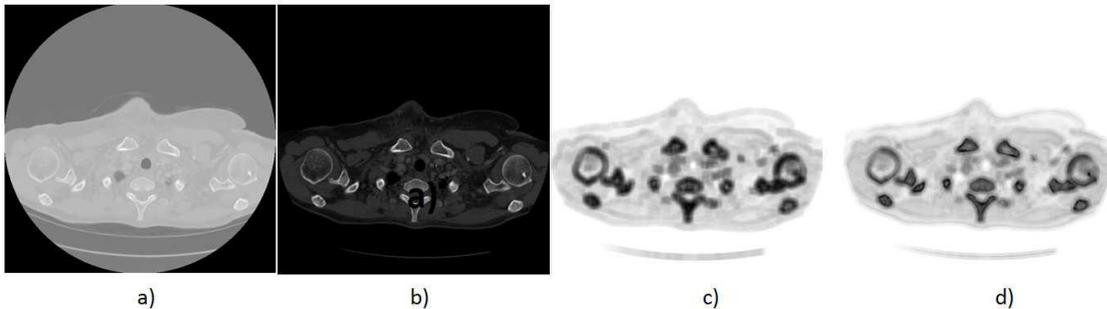


Ilustración 31: a) Imagen original, b) Imagen umbralizada, c) HDI: imagen distancia a histograma, d) CIT: imagen de costes

La imagen *CIT* es la imagen de costes que se le pasa al algoritmo Max-flow que hemos visto en el apartado anterior.

## 5.5. Método de segmentación por regiones

Para desarrollar el algoritmo cuyas características fueron enumeradas en el capítulo 1, tenemos que tener en cuenta ciertos factores de las imágenes y de las técnicas de procesamiento de imágenes.

En primer lugar, las imágenes con las que vamos a trabajar, van tener dimensiones de  $512 \times 512$  y van a ser del tipo `uint16`, por lo que para tratarlas y procesarlas se van a pasar a tipo `double`, y se van a escalar para que tengan valores entre cero y uno.

Como se puede ver en el apartado 4.3., los algoritmos de LoG y Canny dan mucho mejor resultado que los otros detectores de bordes explicados, por lo que serán los que se usen en nuestro algoritmo.

El algoritmo que se ha desarrollado es muy simple, y se puede dividir en tres bloques:

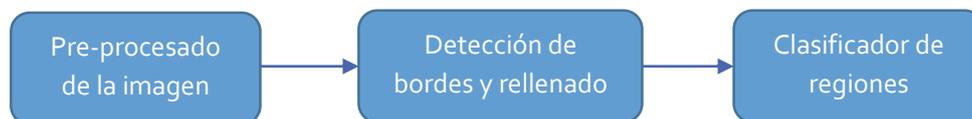


Ilustración 32: Diagrama de flujo del algoritmo desarrollado

A continuación describiremos con más detalle cada uno de los bloques del anterior diagrama. Para ello, usaremos una imagen de ejemplo para ir mostrando los distintos resultados que obtenemos en cada ejemplo.

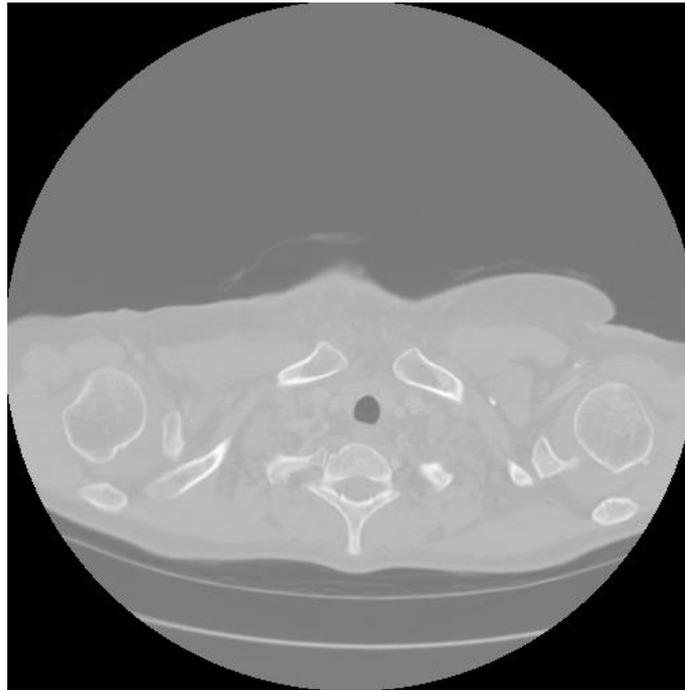


Ilustración 33: Imagen original sobre la que explicaremos el algoritmo desarrollado

### 5.5.1. Pre-procesado de la imagen

En primer lugar se normaliza la imagen para que tenga valores entre 0 y 1 y sea más cómodo de trabajar con ella. Para realizar esta tarea se desarrolló una sencilla función.

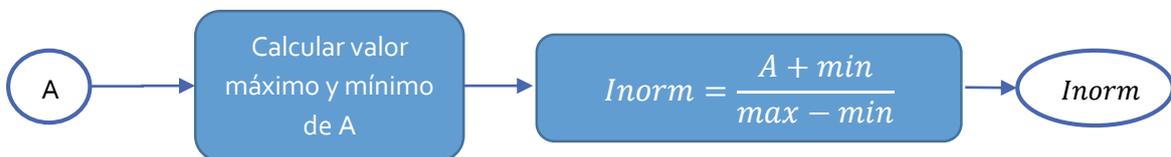


Ilustración 34: Diagrama de flujo de la función de ajuste.

Una vez que tenemos la imagen normalizada con todos los valores entre cero y uno pasamos a hacerle un ajuste de contraste. El ajuste de contraste se realiza de manera dinámica, es decir, se calculan unos valores mínimos y máximos entre los que se va a expandir el histograma de manera lineal. Por ejemplo en este caso se ha expandido el histograma que está entre los valores 0.7373 y 0.8549.

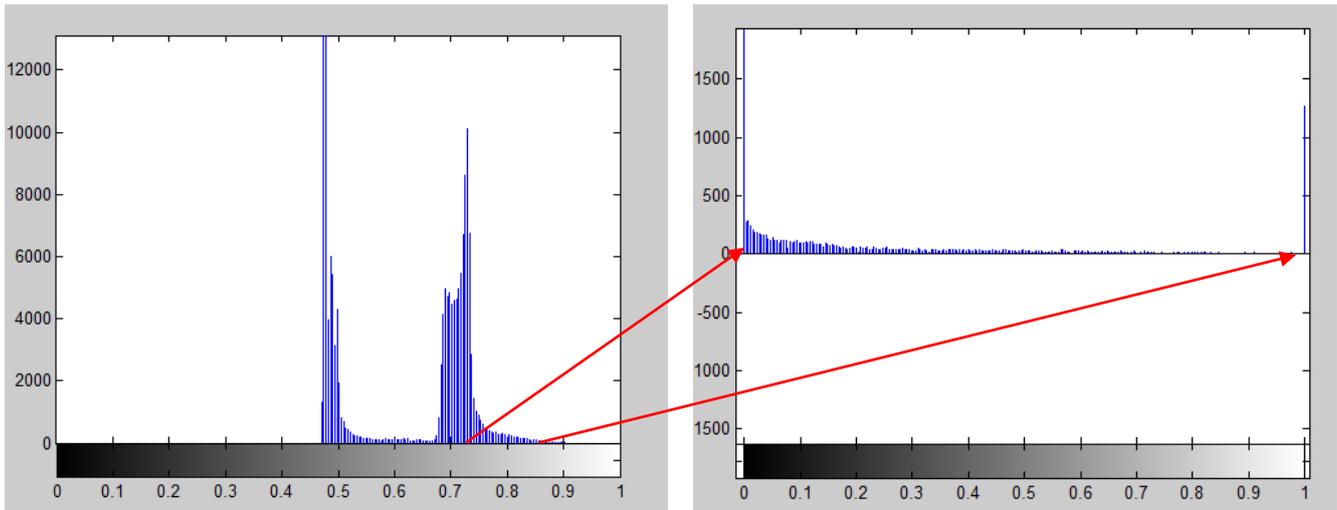


Ilustración 35: Histograma original e histograma expandido entre los valores 0.7373 y 0.8549

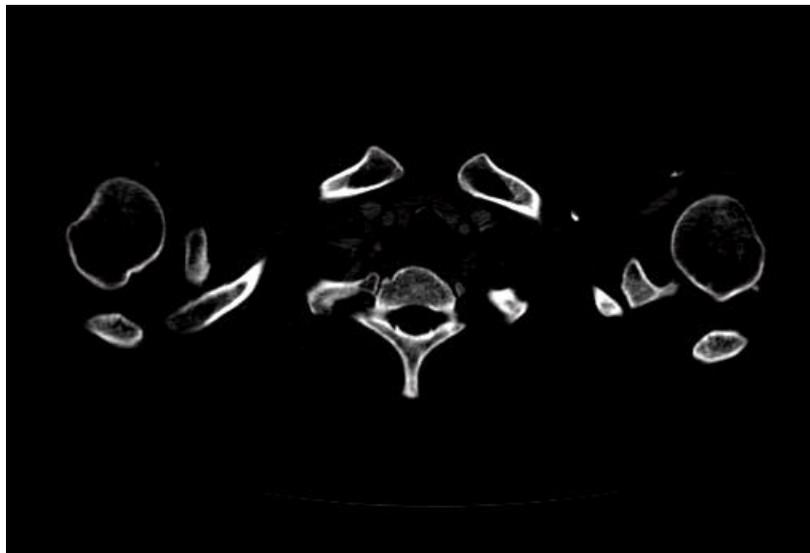


Ilustración 36: Ejemplo de uso de la función de ajuste de contraste automático en la Ilustración 33

Para calcular estos valores, el algoritmo usa la primera derivada. Puesto que la primera derivada, se puede interpretar como el incremento entre un punto y el anterior, el valor inferior a partir del cual expandiremos el histograma, será el último incremento superior a 200 píxeles mientras que el valor superior de la ventana lo marca el último valor de la derivada del histograma que sea superior a 30 píxeles. Los valores de 200 y 30 se han hallado experimentalmente y los resultados que obtenemos con ellos son buenos.

A continuación se muestra el diagrama de flujo que sigue la función de ajuste de contraste:

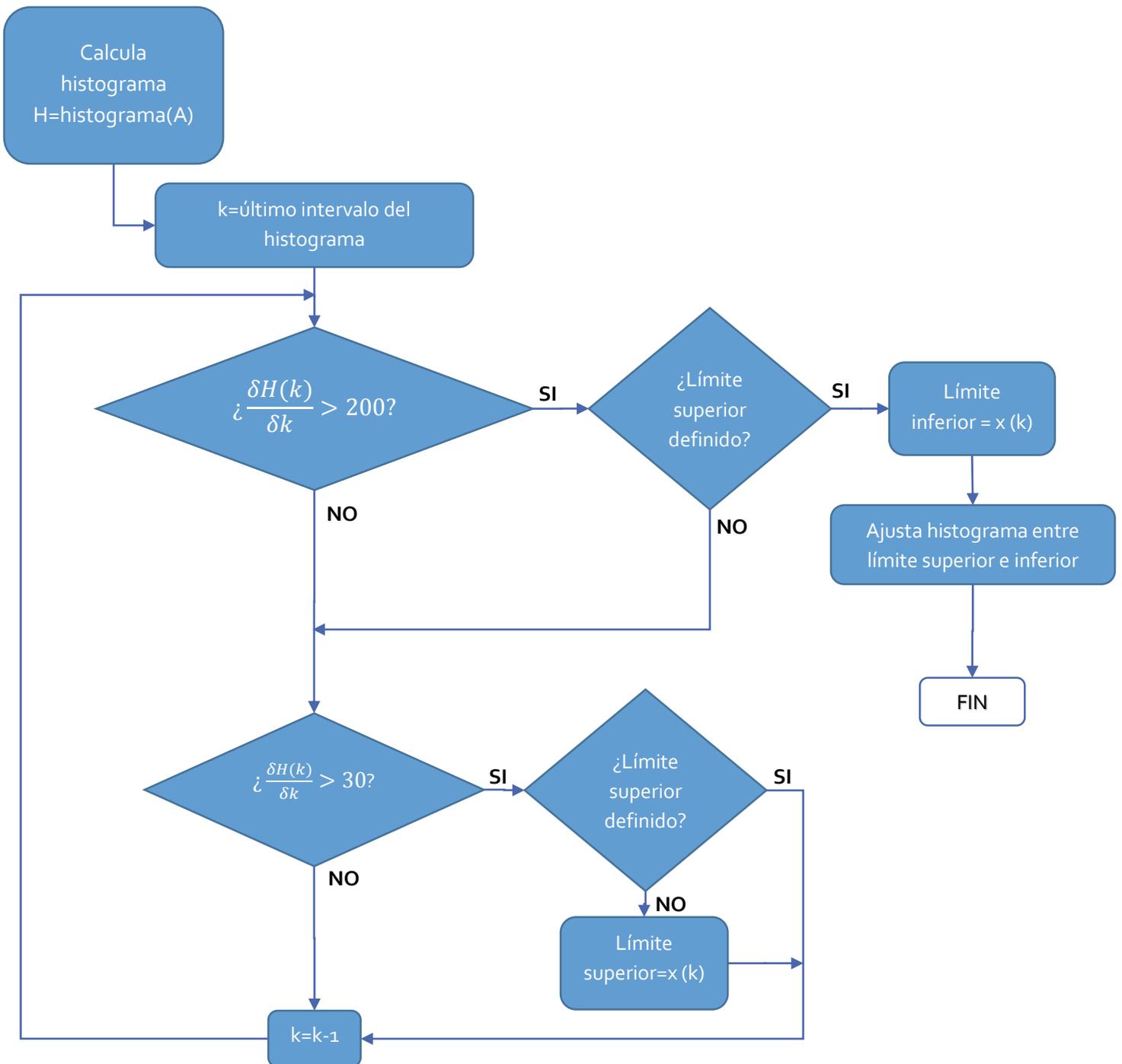


Ilustración 37: Diagrama de flujo de la función de ajuste de contraste

### 5.5.2. Detección de bordes y rellenado

Para la detección de bordes se han usado los métodos LoG y Canny. Se usan ambos porque es muy importante que se formen bordes cerrados, y se ha comprobado que al usar ambos métodos y combinarlos se consiguen mejores resultados.

Una vez que hemos detectado los bordes, hacemos una operación de dilatación con una máscara de 2x2, de esta manera cerramos algunos bordes que han quedado abiertos intentando deformar los contornos que hemos hallado lo menos posible.

Tras cerrar la mayor parte (o todos) los bordes inconexos, hacemos un relleno de todas las regiones mediante el algoritmo que explicamos en el apartado 4.2. y seguidamente le aplicamos una erosión con una máscara de tamaño  $3 \times 3$ .

Como ya sabemos una dilatación seguida de una erosión se trata, en realidad, de un cierre, pero no lo hemos nombrado como tal ya que entre ambos pasos se realiza el relleno de regiones. Lo hacemos de este modo porque en algunos casos la erosión vuelve a abrir algunos bordes que han sido cerrados por la dilatación, pero si hacemos el relleno entre esas dos operaciones sellamos los bordes, evitando así que la erosión los vuelva a abrir.



Ilustración 38: Ilustración 36 tras haber detectado los bordes. Se marcan los bordes inconexos



Ilustración 39:



Ilustración 38 tras haberle aplicado la dilatación. Los bordes inconexos se han cerrado



Ilustración 40:

tras rellenarla y aplicarle una erosión

La elección de una máscara de erosión de  $3 \times 3$  en lugar de usar la misma máscara usada para la dilatación, se ha debido a que proporcionaba mejores resultados. Vamos a poner un ejemplo en el que con la máscara de  $2 \times 2$  se marcaba una zona que no correspondía a hueso, mientras que con la máscara de  $3 \times 3$ , esto desaparece. Se puede observar en las imágenes a continuación.

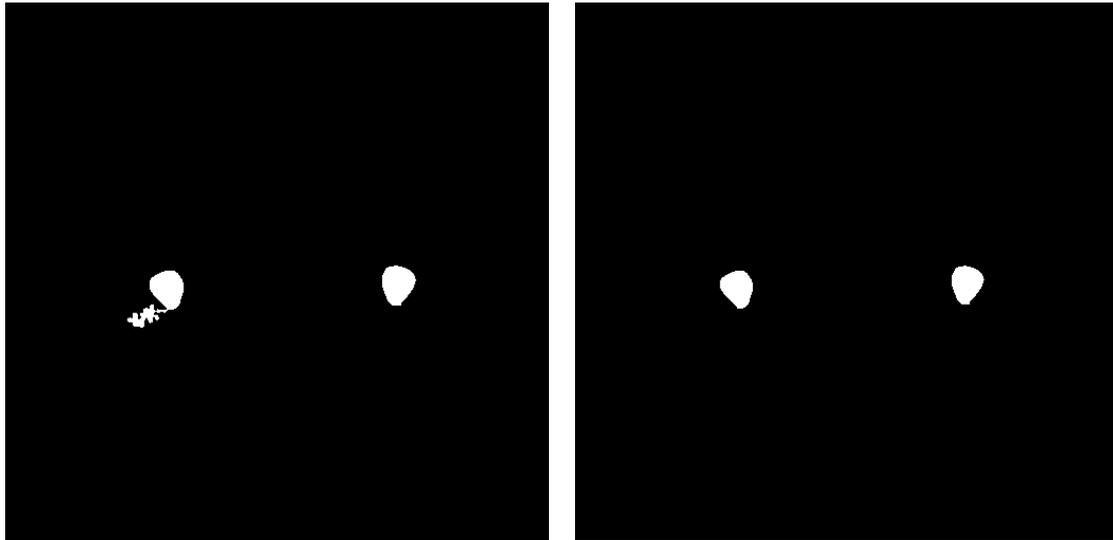


Ilustración 41: Imagen de la izquierda obtenida con máscara de erosión de tamaño 2x2. Imagen de la derecha obtenida con máscara de erosión 3x3

### 5.5.3. Clasificador de regiones

Como se puede observar en la Ilustración 40, hay muchas regiones rellenas que no corresponden a hueso, por lo que el siguiente paso consiste en limpiar esta imagen para deshacernos de las regiones que no corresponden a hueso. Para realizar esta tarea se van a realizar una serie de cálculos de cada región en particular, que serán los siguientes:

- **Área de la región:** a partir del cálculo del área de la región vamos a descartar las regiones cuya área sea menor de 140 píxeles, de esta manera eliminamos regiones que han sido originadas por pequeñas calcificaciones y ruido en general. Además esta medida nos va a servir para saber si nos encontramos en el caso de un TAC del cráneo, en cuyo caso debemos vaciar la región en cuestión, ya que si no nos marcaría como hueso todo el interior del cráneo. Experimentalmente, hemos comprobado que en los TACs de cráneos, el área correspondiente al corte axial del cráneo es siempre superior al 10% del área total de la imagen. Por lo tanto, para evitar que el algoritmo marque todo el cráneo como hueso, se comprobará si el área de la región que estamos estudiando es mayor que 26214 píxeles<sup>7</sup> en cuyo caso se aplicará una umbralización local en la región estudiada sobre la imagen con el contraste ajustado.
- **Nivel máximo de intensidad en el interior de la región:** esta medida es necesaria porque tras el ajuste de contraste se realzan partes de la imagen que corresponden a órganos u otros tejidos que debemos ignorar. Aunque existen

<sup>7</sup> Este número surge de calcular el 10% del área de toda la imagen, es decir, como nuestras imágenes tienen un tamaño de 512x512 píxeles, hemos hecho:  $512 \times 512 \times 0.1 = 26214$  píxeles.

zonas óseas que tienen niveles de Hounsfield por debajo de estos tejidos, estas zonas óseas suelen estar conectadas con otras zonas de la misma región que si tienen niveles de Hounsfield lo suficientemente altos como para poder usarlos para clasificar.

- **Longitud de los ejes de la elipse inscrita en la región:** con esta medida podemos eliminar artefactos del exterior del cuerpo que aparecen en ocasiones con altos niveles de Hounsfield. Este tipo de artefactos suelen encontrarse en las imágenes de TAC de cráneo.

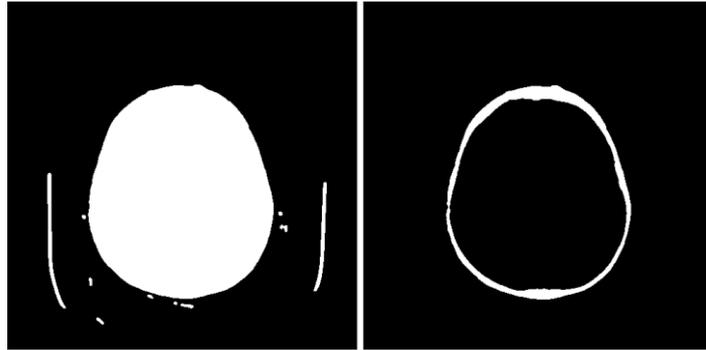


Ilustración 42: Imagen de TAC del cráneo antes y después del clasificador de regiones

## Capítulo 6: Resultados

Para realizar la comparación de los cuatro métodos descritos se han usado 20 cortes de secciones de imágenes de TAC de distintas partes del cuerpo y de pacientes distintos. Con el fin de obtener los resultados más objetivos posibles, inicialmente se ha realizado una segmentación manual de los huesos en estas 20 imágenes. Estas imágenes segmentadas a mano (en las que suponemos que todos los píxeles marcados como hueso realmente lo son), las llamaremos a partir de ahora imágenes *ground truth* y nos servirán para poder hacer una comparación entre los resultados obtenidos mediante los distintos métodos.

La comparación de estos resultados se hará en base a una serie de medidas que definiremos pero necesitamos conocer en primer lugar otros parámetros:

- **TP (True Positive)**: es una medida de los píxeles que han sido marcados como hueso y realmente lo son.
- **FP (False Positive)**: es una medida de los píxeles marcados como hueso que realmente no lo son.
- **TN (True Negative)**: es una medida de aquellos píxeles que han sido marcados como negativos (es decir, que no son hueso) y realmente no lo son.
- **FN (False Negative)**: es una medida de los píxeles que son huesos, pero que han sido marcados como negativos.

Ahora que conocemos estos parámetros se pueden definir las siguientes medidas, que serán las que usemos para calcular los resultados:

- **PPV (Positive Predicted Value)**: Esta medida responde a la pregunta: ¿de los píxeles marcados como hueso, cuantos realmente lo son? Es decir da un valor entre cero y uno de la tasa de píxeles que se han marcado como hueso y realmente lo son.

$$PPV = \frac{TP}{TP + FP} \quad (19)$$

- **Sensitivity**: Esta medida responde a la pregunta: ¿De todos los píxeles que son hueso, cuantos han sido detectados? Es decir da un valor entre cero y uno que representa la relación entre los píxeles que son hueso y los píxeles que se han marcado como tal.

$$Sensitivity = \frac{TP}{TP + FN} \quad (20)$$

- **Specificity:** Esta medida responde a la pregunta: ¿de todos los píxeles que son no son hueso, cuantos han sido detectados? Es decir da un valor entre cero y uno que representa la relación entre los píxeles que no son hueso y los píxeles que se han marcado como que no son hueso.

$$Specificity = \frac{TN}{TN + FP} \quad (21)$$

- **Sørensen–Dice coefficient:** Este coeficiente nos da la similitud entre dos muestras estadísticas, en nuestro caso nos da la similitud entre la imagen *ground-truth* y la calculada por los distintos algoritmos. Su fórmula es:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (22)$$

Donde *A* representa la imagen resultado, y *B* la imagen *ground-truth*.

Es conveniente indicar que para la obtención de todos estos resultados se ha usado un ordenador portátil Toshiba Satellite Pro, con procesador Intel Pentium Dual Core de 2GHz, y 3GB de memoria RAM.

## 6.1. Resultados obtenidos mediante umbralización

Para la obtención de estos resultados, se ha hallado un umbral óptimo y ese es el que se ha usado en la umbralización. El hecho de hacer esto y no usar el método de Otsu, por ejemplo, se debe a que los resultados obtenidos de esta manera son mejores.

Para hallar este umbral se ha hecho un barrido de valores posibles para el umbral (entre 900 y 1300), comparando los resultados obtenidos con cada umbral. La imagen que mejores resultados presentaba era la imagen con un valor de umbral de 1200.

A continuación, mostramos una tabla con los valores obtenidos para las medidas anteriormente definidas.

Tabla 1: Resultados obtenidos mediante umbralización

Imagen	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
IM1	0,7631	0,5977	0,9900	0,6704	0,0005
IM2	0,7693	0,6481	0,9910	0,7036	0,0010
IM3	0,9934	0,0253	1,0000	0,0493	0,0008
IM4	1,0000	0,5740	1,0000	0,7293	0,0008
IM5	0,6075	0,9985	0,9829	0,7554	0,0008
IM6	0,6257	1,0000	0,9838	0,7697	0,0011
IM7	0,8360	0,9949	0,9926	0,9085	0,0010
IM8	0,7059	1,0000	0,9867	0,8276	0,0068
IM9	0,7564	1,0000	0,9883	0,8613	0,0008
IM10	0,6913	0,8914	0,9976	0,7787	0,0005

IM11	0,9464	0,6703	0,9976	0,7847	0,0005
IM12	0,9702	0,7223	0,9987	0,8281	0,0005
IM13	0,9579	0,8328	0,9985	0,8910	0,0011
IM14	0,5696	0,8527	0,9762	0,6830	0,0013
IM15	0,8121	0,9914	0,9849	0,8929	0,0011
IM16	0,7086	0,6367	0,9970	0,6707	0,0005
IM17	0,7850	0,6167	0,9973	0,6908	0,0005
IM18	0,8192	0,5685	0,9986	0,6712	0,0010
IM19	0,9316	0,7006	0,9987	0,7998	0,0011
IM20	0,9514	0,7523	0,9981	0,8402	0,0008
<b>PROMEDIO</b>	<b>0,8100</b>	<b>0,7537</b>	<b>0,9929</b>	<b>0,7403</b>	<b>0,0011</b>

A continuación expongo algunas de las imágenes obtenidas.

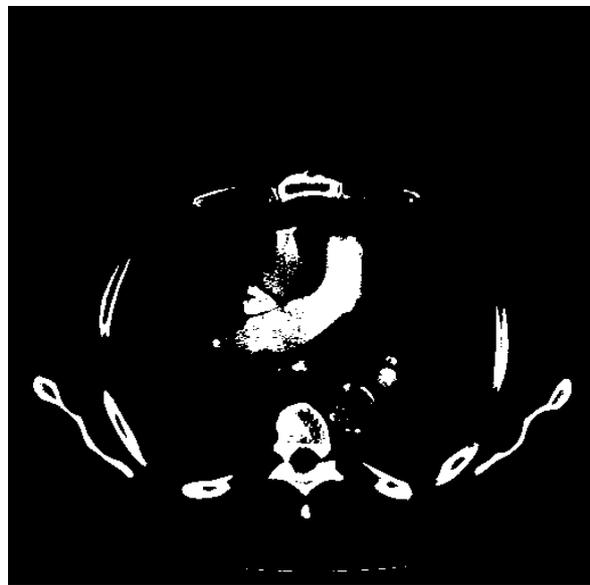
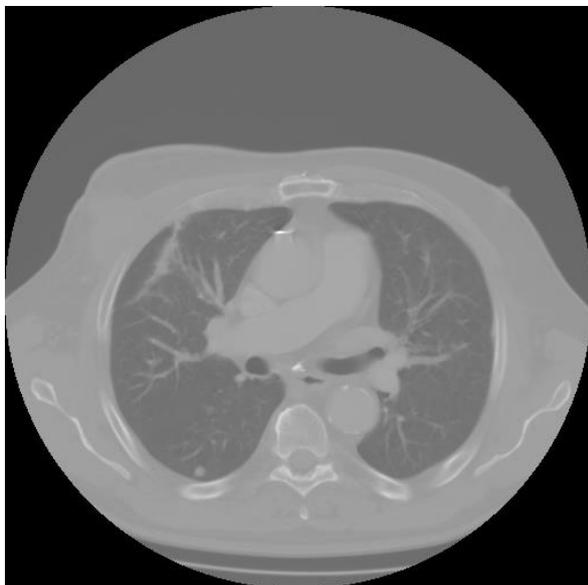


Ilustración 43: Imagen IM14 segmentada mediante umbralización

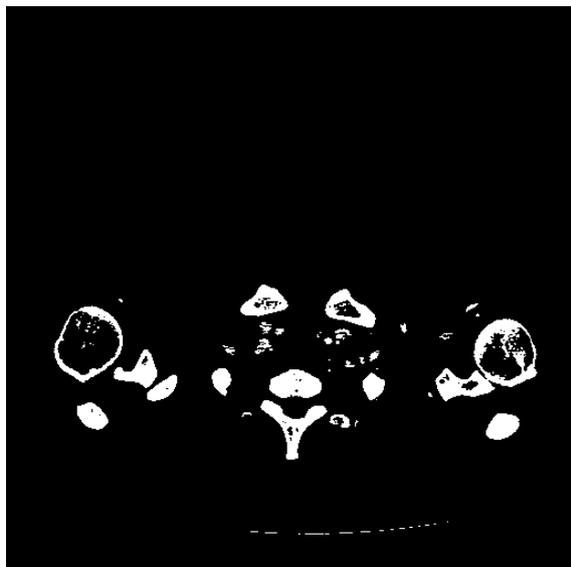
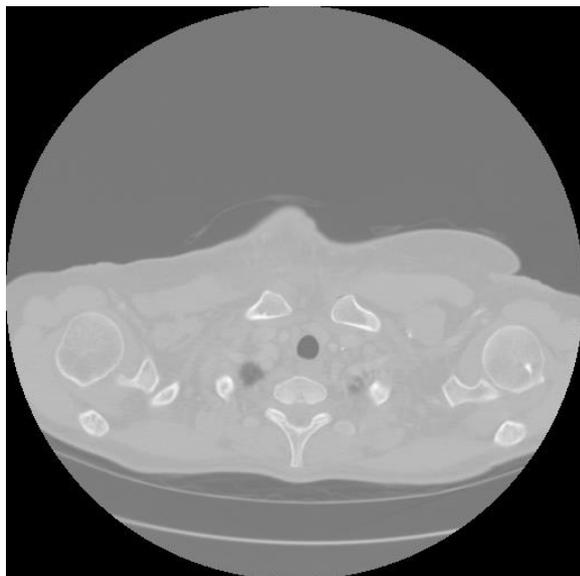


Ilustración 44: Imagen IM2 segmentada mediante umbralización

## 6.2. Resultados obtenidos con el método de Level-Sets

Como ya adelantábamos en la explicación de esta técnica, los Level-Sets van a tener muchos problemas a la hora de segmentar imágenes de  $512 \times 512$  píxeles, tamaño de las imágenes con las que estamos trabajando porque son imágenes bastante grandes, ralentizando mucho el proceso. Debido a esto, solo se ha aplicado a dos imágenes, que vemos a continuación:

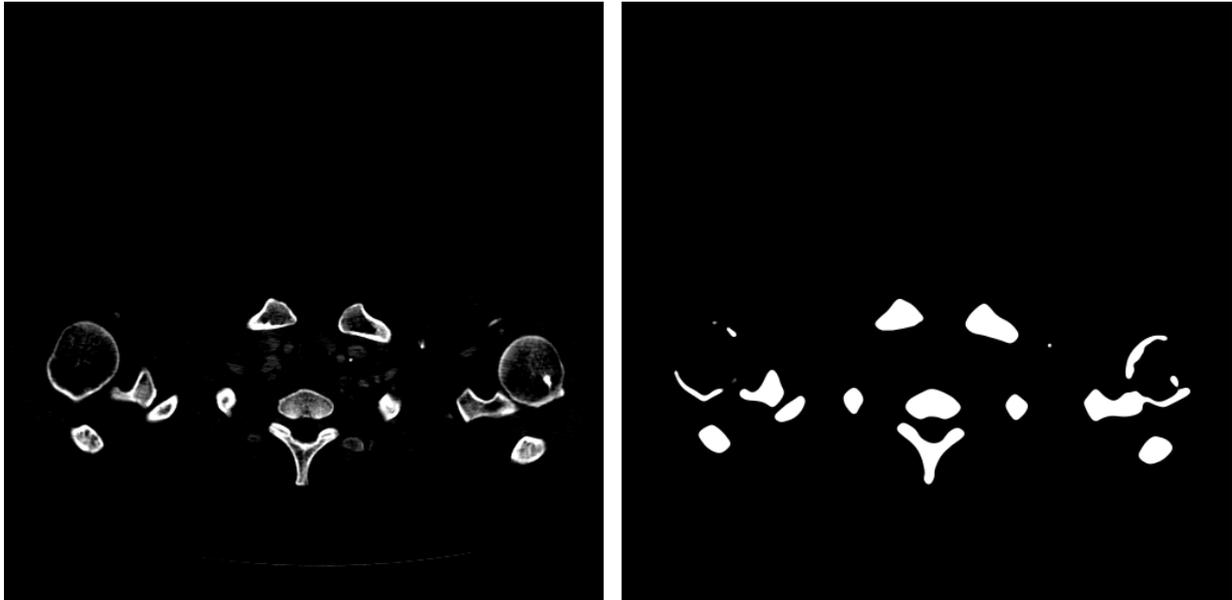


Ilustración 45: Ejemplo de aplicar Level-Sets sobre la imagen pre-procesada IM2

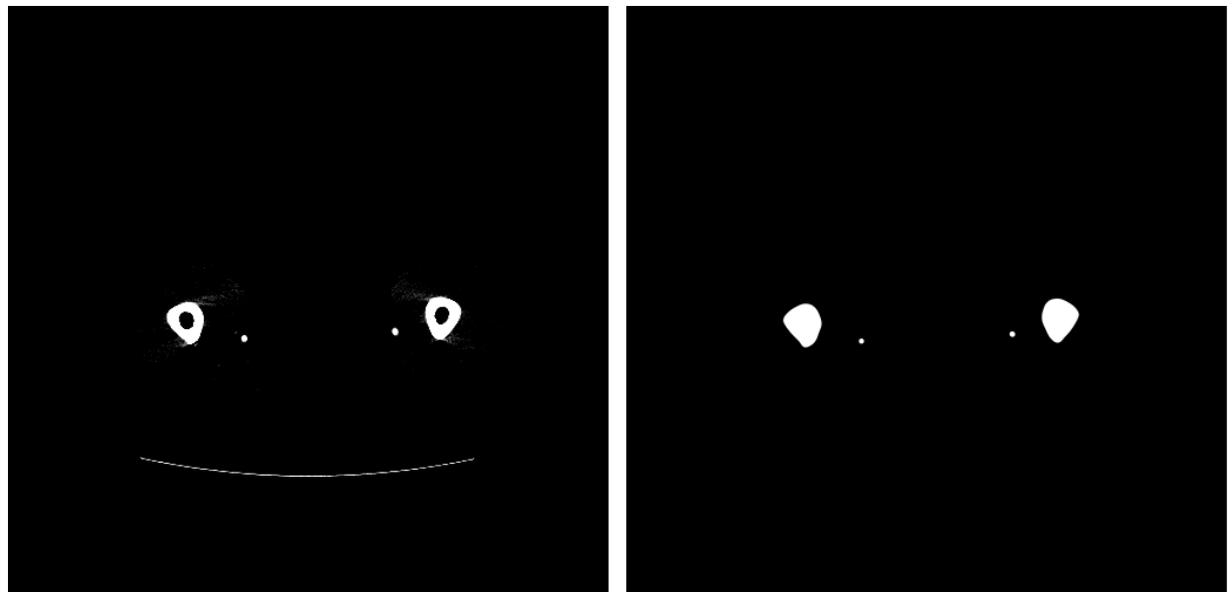


Ilustración 46: Ejemplo de aplicar Level-Sets sobre la imagen pre-procesada IM4

Para ambas imágenes se han usado un total de 600 iteraciones y los siguientes parámetros:  $\lambda=0.5$  y  $\alpha=8$ .

Los resultados que se han obtenido son los siguientes:

Tabla 2: Resultados obtenidos mediante Level-sets

Imagen	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
IM1	0,864	0,598	0,996	0,707	440
IM2	0,836	0,998	0,999	0,9011	372
<b>PROMEDIO</b>	<b>0.8635</b>	<b>0.7980</b>	<b>0.9930</b>	<b>0.804</b>	<b>406</b>

Si en vez de usar la imagen entera usamos recortes de ella, el método es mucho más rápido y además podemos comparar el resultado de aplicar distintos parámetros de una manera más directa.

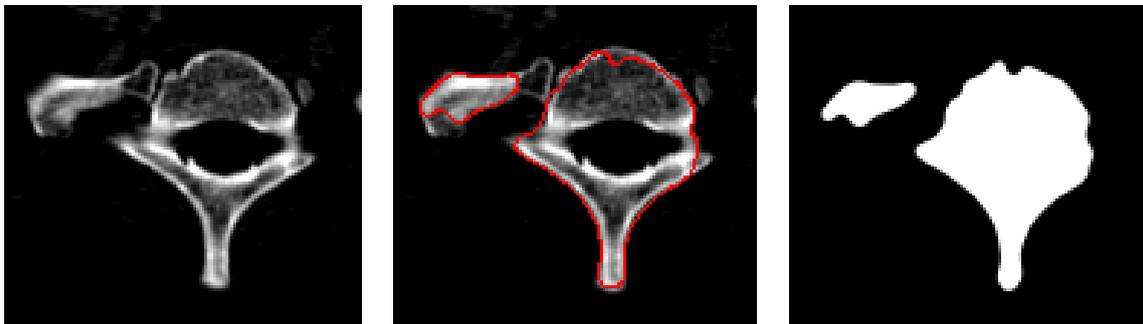


Ilustración 47: Level-sets con parámetros  $\lambda=0.5$  y  $\alpha=8$  (600 iteraciones)

Como vemos el parámetro  $\alpha=8$  es demasiado agresivo ya que se pasa el borde del hueso. Veamos cómo evoluciona la curva si modificamos este parámetro.

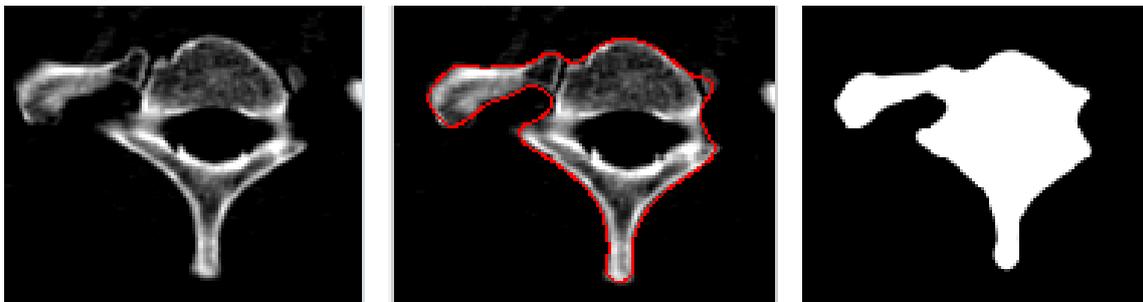


Ilustración 48: Level-sets con parámetros  $\lambda=0.5$  y  $\alpha=1$  (600 iteraciones)

Como podemos ver, hemos conseguido un buen resultado pero vamos a comprobar cómo afectaría a este resultado aumentar el parámetro  $\lambda$ , que como sabemos es que el regula la longitud del perímetro de la región segmentada.

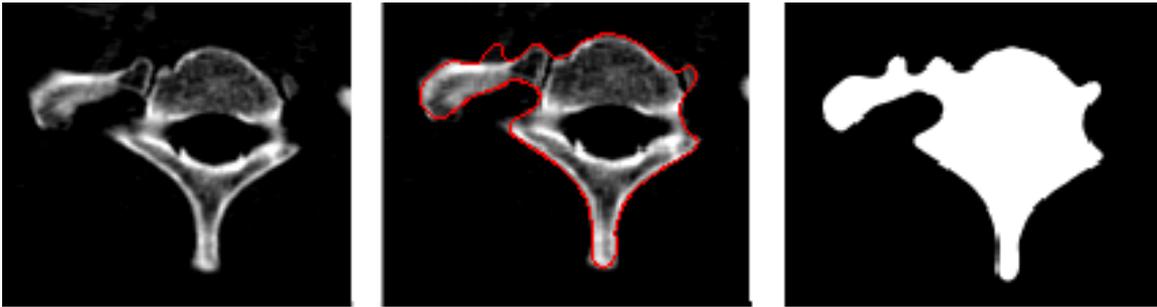


Ilustración 49: Level-sets con parámetros  $\lambda=8$  y  $\alpha=1$  (600 iteraciones)

Al aumentar el parámetro  $\lambda$  la curva se queda enganchada en cualquier ruido, empeorando los resultados obtenidos.

Por último, como prueba, podemos hacer una umbralización tras aplicar Level-Sets. Esta prueba se va a hacer con los parámetros que obtienen mejor resultado (Ilustración 48).



Ilustración 50: Umbralización más Level-sets (umbral=0)

### 6.3. Resultados obtenidos con el método de Max-Flow

Durante el desarrollo de nuestro algoritmo se pensó que usar estas técnicas sería de gran ayuda, ya que unos de los principales problemas que encontramos, como ya se ha indicado, es el de la descalcificación de los huesos, y se pensó que con esta técnica se podían conseguir mejores resultado por tener en cuenta los pixeles vecinos a la hora de asignar una región a cada pixel.

El problema, como vamos a ver a continuación, es que si solamente se usa la información de nivel de gris para asignar los costes a los enlaces, el algoritmo no es capaz de diferenciar los tejidos óseos descalcificados de las zonas que no son hueso.

Como sabemos, hay que inicializar los centroides de los Max-flow. Para elegir los valores que mejores resultados nos ofrecían, se han realizado una serie de pruebas para poder determinarlo. Estas pruebas se han realizado tanto en la imagen original como en la imagen con el contraste modificado para resaltar los huesos. A continuación mostramos las pruebas realizadas en la imagen original:

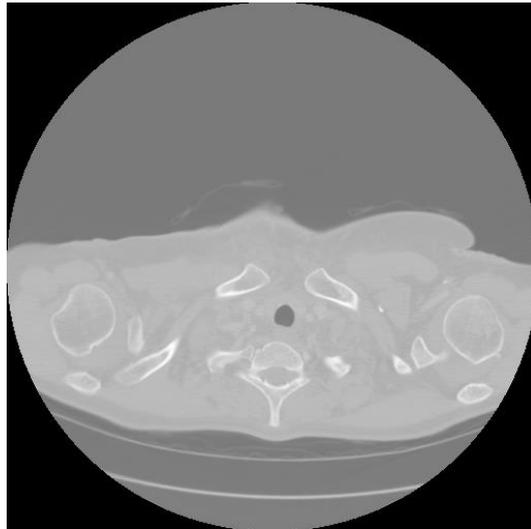


Ilustración 51: Imagen original que usamos para encontrar los centroides óptimos



Ilustración 52: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.765 (hueso) y 0.55 (fuera del hueso). Tiempo de ejecución: 3.628579 s

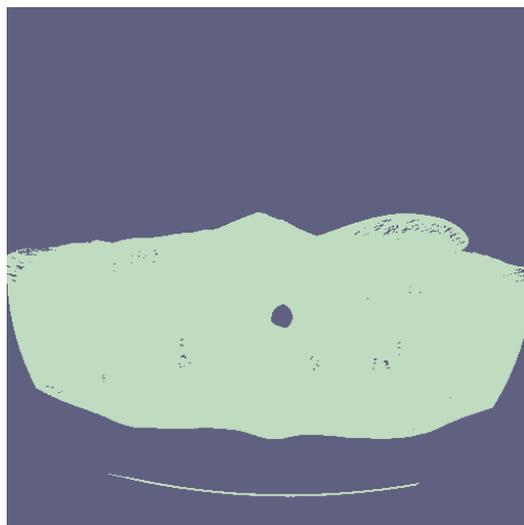


Ilustración 53: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.4 (fuera del hueso). Tiempo de ejecución: 3.551244 s

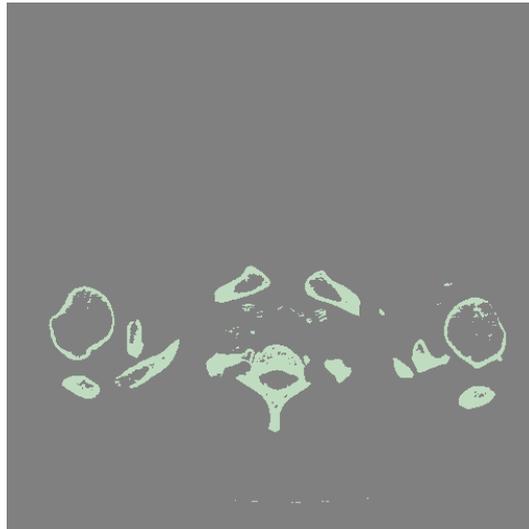


Ilustración 54: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.43 (fuera del hueso). Tiempo de ejecución: 4.184355 s



Ilustración 55: Segmentación con Max-flow a partir de la imagen original y con centroides en 0.823 (hueso) y 0.51 (fuera del hueso). Tiempo de ejecución: 3.486082 s

Si en vez de usar la imagen original, usamos la imagen con el contraste ajustado, vemos que obtenemos resultados semejantes, o incluso peores, ya que en este caso se ha perdido información. Lo vemos a continuación:



Ilustración 56: Imagen ajustada

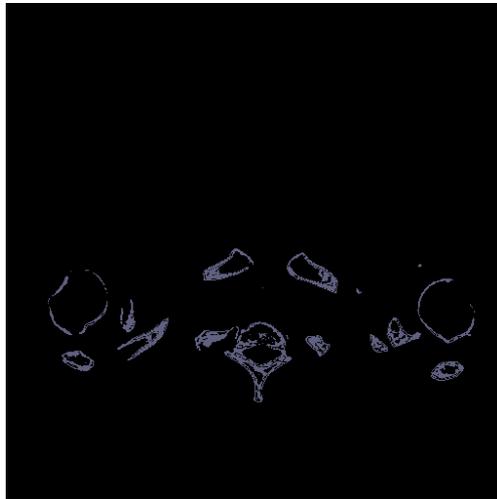


Ilustración 57: Segmentación con Max-flow a partir de la imagen ajustada y con centroides en 0.392 (hueso) y 0 (fuera del hueso). Tiempo de ejecución: 3.662611 s



Ilustración 58: Segmentación con Max-flow a partir de la imagen ajustada y con centroides en 0.784 (hueso) y 0 (fuera del hueso). Tiempo de ejecución: 3.621492 s

Una vez realizadas todas estas pruebas se ha optado por usar los parámetros que se han usado en la Ilustración 54, con los cuales se han conseguido los siguientes resultados:

Tabla 3: Resultados obtenidos mediante Max-flow

Imagen	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
IM1	0,2238	0,8671	0,8373	0,3558	3,0974
IM2	0,2134	0,9430	0,8398	0,3481	2,9716
IM3	0,9720	0,0410	1,0000	0,0786	2,9218
IM4	0,9979	0,6169	1,0000	0,7624	2,9389
IM5	0,3587	0,9988	0,9527	0,5278	2,9466
IM6	0,3407	1,0000	0,9476	0,5083	2,9494
IM7	0,5612	1,0000	0,9702	0,7189	2,9501
IM8	0,4455	1,0000	0,9602	0,6164	2,9241
IM9	0,5304	1,0000	0,9678	0,6932	2,9638
IM10	0,0571	0,9372	0,9054	0,1077	2,9367
IM11	0,2983	0,9424	0,8579	0,4532	2,9514
IM12	0,2845	0,9520	0,8588	0,4381	2,9559
IM13	0,1673	0,9906	0,8029	0,2863	2,9444
IM14	0,1974	0,9792	0,8531	0,3285	2,9792
IM15	0,2521	0,9582	0,8124	0,3992	2,9482
IM16	0,0870	0,9949	0,8813	0,1601	2,9427
IM17	0,1268	0,9833	0,8914	0,2247	2,9385
IM18	0,1074	0,9853	0,9099	0,1937	2,9314
IM19	0,2322	0,9836	0,9201	0,3757	2,9440
IM20	0,3224	0,9509	0,9025	0,4815	2,9369
<b>PROMEDIO</b>	<b>0,3388</b>	<b>0,9062</b>	<b>0,9036</b>	<b>0,4029</b>	<b>2,9537</b>

Es interesante resaltar que umbralizando se conseguían mejores resultados.

Estos malos resultados se deben a que usar en Max-flow como imagen de costes la imagen de intensidad no es lo más apropiado, sin embargo en el siguiente punto vamos a ver como creando una imagen de costes apropiada se consiguen unos resultados mucho mejores.

## 6.4. Resultados obtenidos con el método de segmentación de estructuras óseas basada en métodos estadísticos y Max-flow

Este algoritmo ha sido desarrollado con la finalidad de segmentar hueso en imágenes de TAC, que es precisamente el tema de estudio en este trabajo. Como se ha visto en el apartado 5.4., éste algoritmo usa la técnica de segmentación Max-flow pero en vez de

utilizar como imagen de costes la imagen a tratar como hemos hecho en el apartado anterior, se ha construido una imagen de distancias de histograma combinada con la imagen de niveles de intensidad. Esto es de gran utilidad ya que el histograma tiene información de textura, que es muy importante a la hora de diferenciar entre tejido óseo y otras zonas.

Gracias a esto consigue los mejores resultados vistos hasta ahora, como podemos ver a continuación:

Tabla 4: Resultados obtenidos mediante métodos estadísticos y Max-flow

Imagen	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
IM1	0,787	0,991	0,985	0,877	896
IM2	0,793	0,997	0,988	0,883	897
IM3	0,876	0,998	0,997	0,933	897
IM4	0,962	0,983	1,000	0,973	913
IM5	0,689	0,993	0,988	0,814	882
IM6	0,692	1,000	0,988	0,818	867
IM7	0,944	0,992	0,998	0,967	899
IM8	0,781	1,000	0,991	0,877	907
IM9	0,846	0,999	0,993	0,916	933
IM10	0,986	0,945	1,000	0,965	866
IM11	0,925	0,855	0,996	0,889	886
IM12	0,937	0,984	0,996	0,960	911
IM13	0,892	0,994	0,995	0,940	885
IM14	0,767	0,963	0,989	0,853	910
IM15	0,818	0,815	0,985	0,817	1026
IM16	0,885	0,995	0,999	0,937	1055
IM17	0,901	0,996	0,998	0,946	969
IM18	0,884	0,998	0,999	0,938	918
IM19	0,882	0,985	0,997	0,930	915
IM20	0,942	0,996	0,997	0,968	909
<b>PROMEDIO</b>	<b>0,859453</b>	<b>0,974022</b>	<b>0,993926</b>	<b>0,91012</b>	<b>917</b>

Si comparamos el tiempo de ejecución de este código con el anterior vemos que de los 917 segundos que tarda de media solamente unos tres corresponden a la parte de maximización de flujo (Max-flow). El resto del tiempo se emplea en calcular los histogramas locales de cada pixel.

Como es inmediato de ver, el principal problema de este algoritmo es su gran carga computacional, teniendo un tiempo medio de cómputo por imagen de 15 minutos y 17 segundos. Si estamos usando un equipo modesto, como es el caso en este proyecto,

resulta inviable tratar una pila de imágenes lo suficientemente grande que nos permitan formar un modelo 3D.

## 6.5. Resultados obtenidos mediante el método de segmentación por regiones

Debido a que este método se ha explicado con gran detalle en el apartado 5.5., en este apartado nos limitaremos a exponer los resultados obtenidos.

Tabla 5: Resultados obtenidos mediante el método de segmentación por regiones

Imagen	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
IM1	0,882	0,975	0,993	0,926	2,024
IM2	0,926	0,993	0,996	0,958	0,849
IM3	0,881	0,995	0,997	0,935	0,911
IM4	0,926	0,996	1,000	0,960	0,537
IM5	0,802	0,939	0,994	0,865	0,811
IM6	0,948	0,947	0,999	0,948	0,784
IM7	1,000	0,688	1,000	0,815	0,809
IM8	0,999	0,875	1,000	0,933	0,739
IM9	1,000	0,772	1,000	0,871	0,828
IM10	0,998	0,970	1,000	0,984	0,711
IM11	0,960	0,703	0,998	0,811	0,730
IM12	0,956	0,901	0,998	0,928	0,727
IM13	0,974	0,806	0,999	0,882	0,885
IM14	0,984	0,722	1,000	0,833	0,696
IM15	0,764	0,829	0,988	0,795	0,773
IM16	0,904	0,972	0,999	0,936	0,702
IM17	0,915	0,984	0,999	0,948	0,725
IM18	0,952	0,994	0,999	0,972	0,674
IM19	0,919	0,989	0,998	0,953	0,713
IM20	0,957	0,958	0,998	0,957	0,737
<b>PROMEDIO</b>	<b>0,93231</b>	<b>0,90038</b>	<b>0,99763</b>	<b>0,91057</b>	<b>0,81830</b>

A continuación, se exponen las imágenes resultados obtenidos mediante el método de segmentación basado en métodos estadísticos y Max-flow y mediante el método de segmentación por regiones.

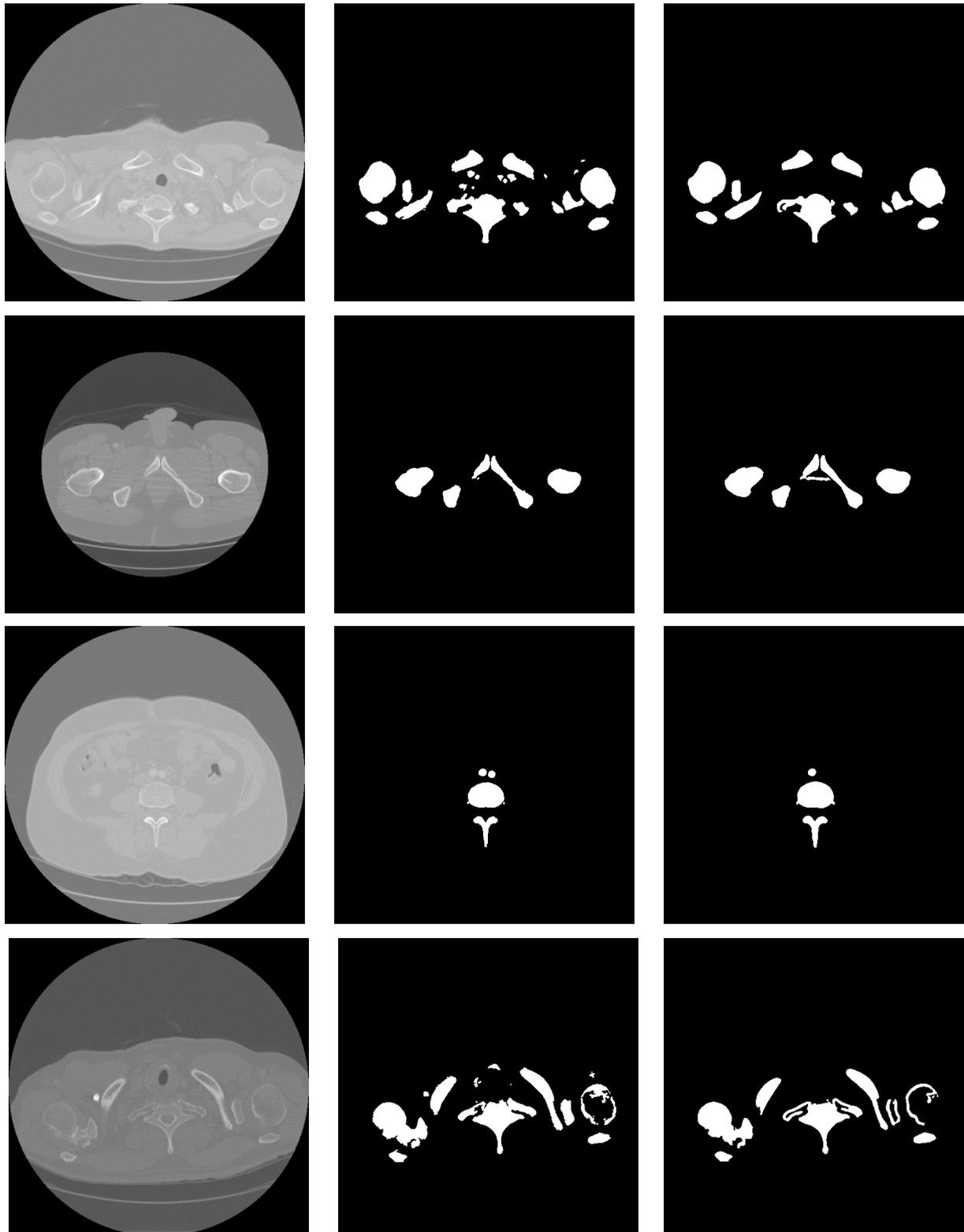


Ilustración 59: Comparación de técnicas de segmentación. De izquierda a derecha. Imagen original, segmentación por distancia a histograma + Max-flow, y segmentación mediante regiones

## Capítulo 7: Comparación de resultados y conclusiones

A continuación se añade una tabla de las medidas promedio obtenidas con cada uno de los distintos métodos para que el lector pueda compararlos cómodamente.

Tabla 6: Comparación de resultados obtenidos con los distintos métodos

Método	PPV	Sensitivity	Specificity	DiceCoef	Tiempo (s)
<b>Umbralización</b>	0,8100	0,7537	0,9929	0,7403	0,0011
<b>DRLSE*</b>	0.8635	0.7980	0.9930	0.804	406
<b>Max-flow</b>	0,3388	0,9062	0,9036	0,4029	2,9537
<b>Distancia a histograma + Max-flow</b>	0,859453	0,974022	0,993926	0,91012	917
<b>Segmentación por regiones</b>	0,93231	0,90038	0,99763	0,91057	0,81830

\*Sólo se han calculado con dos imágenes.

En esta tabla vemos que mediante el método de segmentación por regiones conseguimos resultados semejantes a los que se consiguen usando el método de distancia a histograma y también podemos apreciar la diferencia entre ambos métodos en la medida *sensitivity*. Ésta diferencia se debe a que el método de segmentación por regiones necesita bordes pronunciados para poder realizar correctamente el llenado de las regiones, y en los casos en los que los bordes son muy débiles, suele dejar el hueso vacío, este efecto se puede apreciar bien en las imágenes del Anexo A. A pesar de esto, tiene a favor que el método de segmentación por regiones es 1120 veces más rápido.

El método de segmentación por regiones desarrollado en este trabajo es capaz de tratar y segmentar una gran cantidad de imágenes en poco tiempo, usando un ordenador portátil de usuario, llegando a tratar hasta 526 imágenes en 4 minutos y 50 segundos, por lo que se considera que usando equipamiento especializado podría ser capaz de tratar las imágenes en tiempo real y dar un modelo 3D in situ.

A pesar de los buenos resultados obtenidos, se considera que queda mucho por avanzar en esta línea de trabajo, y se proponen como líneas futuras la incorporación de algoritmos y técnicas de visión por computador que permitan usar los slides contiguos para determinar si una región es hueso o no, o para detectar cambios bruscos en la frontera de la región tratada y poder realizar un mejor procesado de la región.

Queda mucho por investigar en relación con las imágenes médicas. Gracias a ellas la medicina no para de avanzar, facilitando el diagnóstico en todo tipo de campos. Gracias

a algoritmos y técnicas como las que se han estudiado en este proyecto cada día resulta más fácil la labor de planificación quirúrgica, en gran medida gracias a los modelos 3D que permiten al personal médico ver con detalle las estructuras con las que se van a encontrar, y se espera que trabajos como este aporten un pequeño grano de arena a esta labor.

## Capítulo 8: Bibliografía y referencias

---

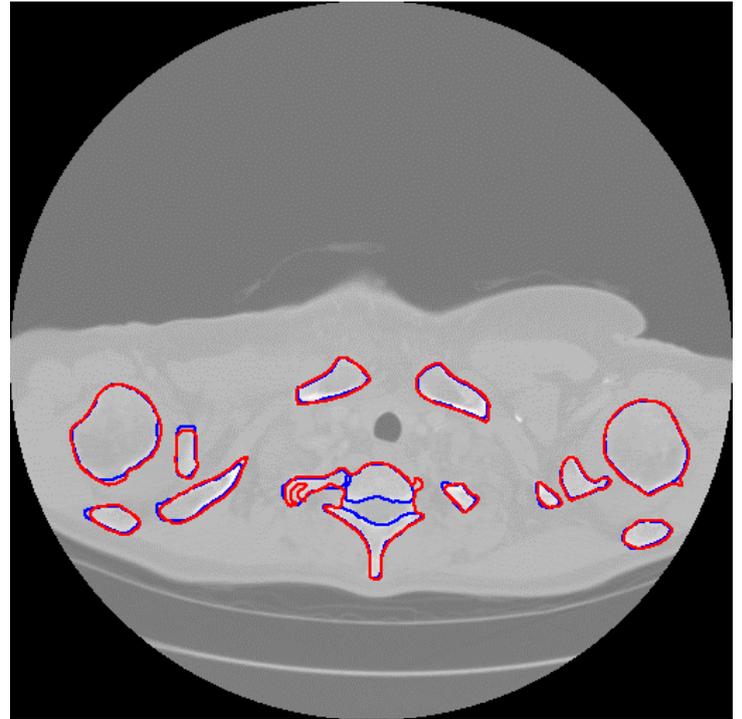
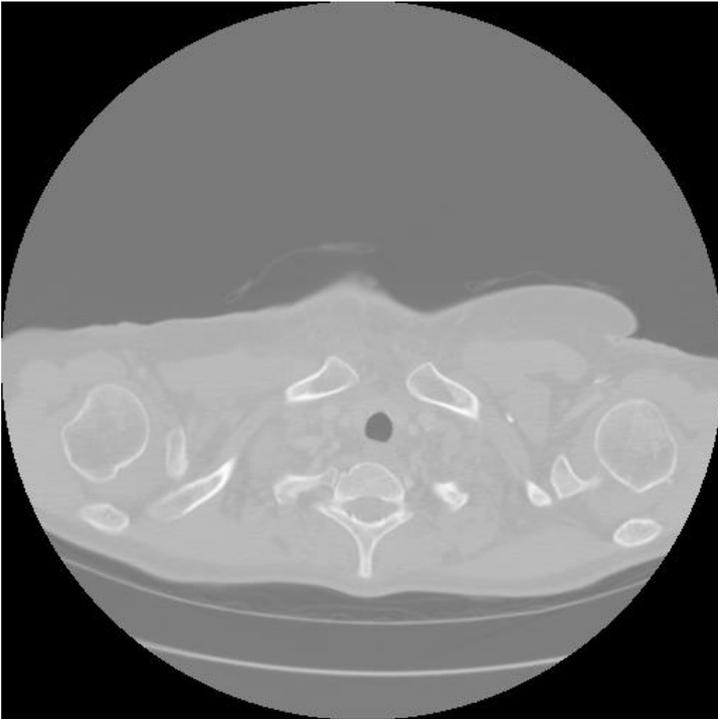
- [1] D. A. Christensen, *Ultrasonic bioinstrumentation*, New York : John Wiley & Sons, 1998.
- [2] D. G. Ortiz y M. P. A. director, «Resonancia magnética nuclear: descripción general y adecuación de un local para la realización de resonancias magnéticas,» Sevilla, 2012.
- [3] P. A. G. Ruiz, «Introducción a la resonancia magnética nuclear,» Diego Marín, Murcia, 2006.
- [4] R. Chandra, «Nuclear medicine physics : the basics,» Lippincott Williams & Wilkins, Philadelphia, 2012.
- [5] B. A. C. S. Rangaraj M. Rangayyan, «Color image processing with biomedical applications,» SPIE Press, Bellingham, Washington , 2011.
- [6] K. N. P. Ratislav Lukac, «Color image processing : methods and applications,» Taylor & Francis, Boca Raton [Florida], 2007.
- [7] «Digital Image and Communications,» ACR-NEMA Standards Publications, Washington, D.C., 1985.
- [8] N. P. / I. I. 12052, «The DICOM Standard 2014a,» Rosslyn, VA, USA, 2014.
- [9] R. E. W. S. L. E. Rafael C. Gonzalez, *Digital Image processing using MATLAB*, New Delhi : McGraw Hill Education, 2010 .
- [10] C. X. Chunming Li, «Distance Regularized Level Set Evolution and Its,» *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 19, pp. 3243-3254, 2010.
- [11] E. B. X.-C. T. a. Y. B. T. r. C. J. Yuan, «A study on continuous max-flow and,» *UCLA*, pp. 10-61, 2010.
- [12] E. B. a. X. T. J. Yuan, «A study on continuous max-flow and min-cut,» de *CVPR*, USA, San Francisco, 2010.
- [13] C. S. C. S.-M. B. A. J. A. Pérez Carrasco, «Statistical-based Segmentation of Bone,» de *ICIAR*, Vilamoura, Algarve, Portugal, 2014.
- [14] C. S. Begoña Acha, «Apuntes de Tratamiento Digital de Imágenes Medicas,» Sevilla, 2014.
- [15] I. G. Fenoll, «Aportaciones a la Segmentación y Caracterización de Imágenes Médicas 3D,» sevilla, 2010.
- [16] C. S. mejías, «"Preprocesado y Segmentación de Imágenes Tomograficas Axiales Computerizadas",» Sevilla, 2007.

- [17] J. J. C. Plaza, «“SEGMENTACIÓN 3D DE IMÁGENES MÉDICAS DE TOMOGRAFÍA AXIAL COMPUTERIZADA (TAC) POR CRECIMIENTO DE REGIONES MULTITOLERANCIA”,» Sevilla, 2007.

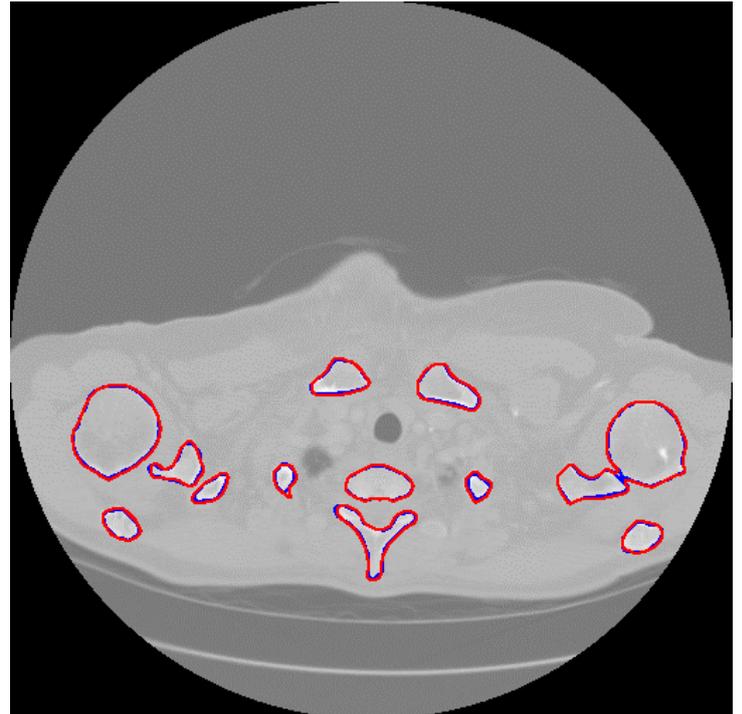
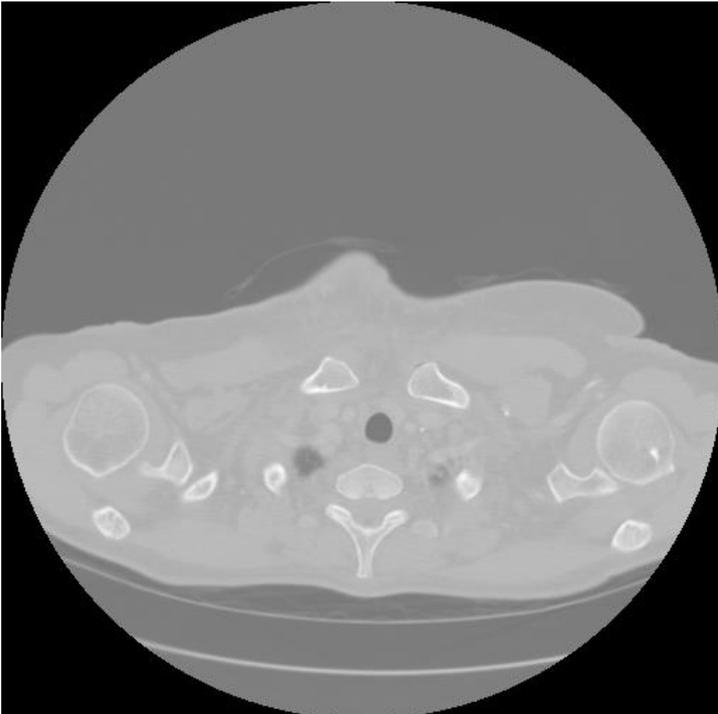
## Anexo A

En este Anexo se incluyen todas las imágenes obtenidas tras aplicar el **método de segmentación basado en regiones**. En rojo se encuentran marcados los huesos detectados por este método, mientras que en azul se encuentran los huesos reales (obtenidos a partir de las imágenes *ground-truth*).

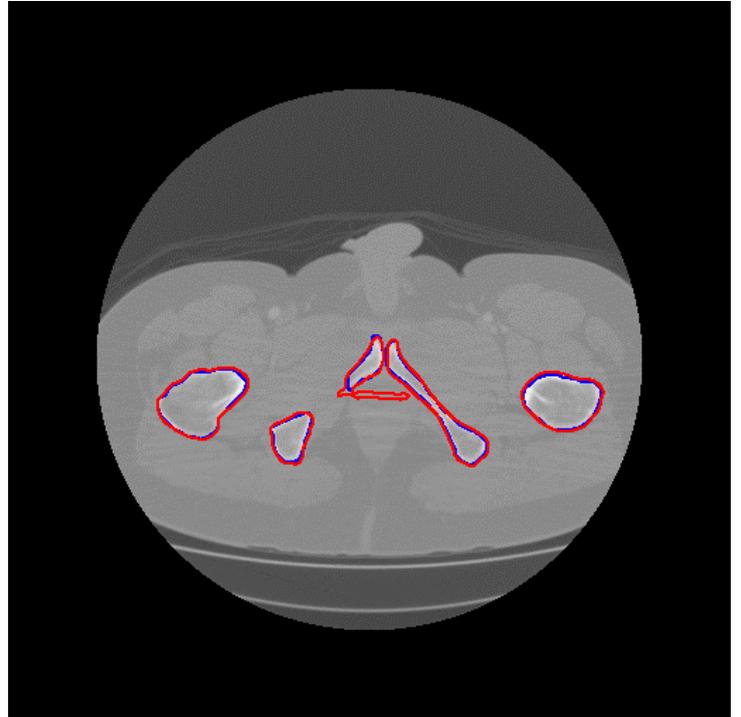
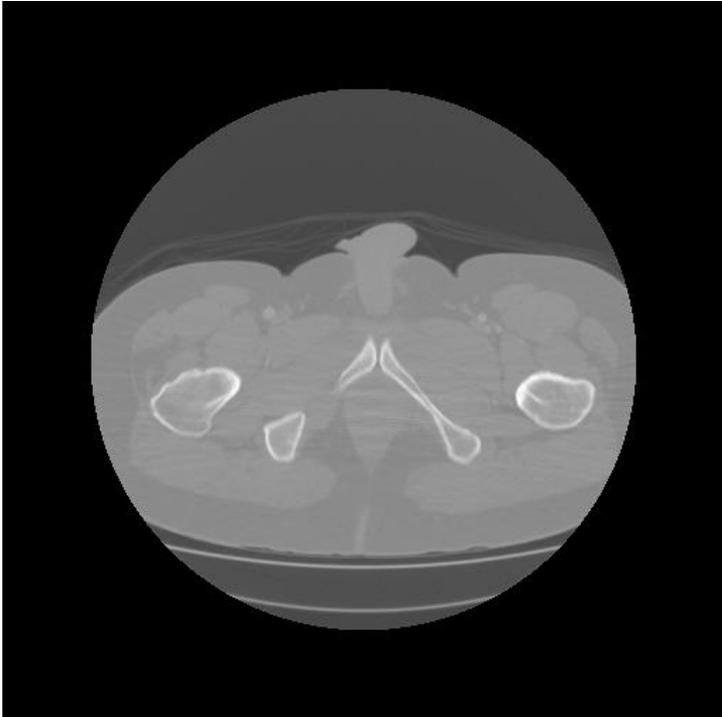
IM1



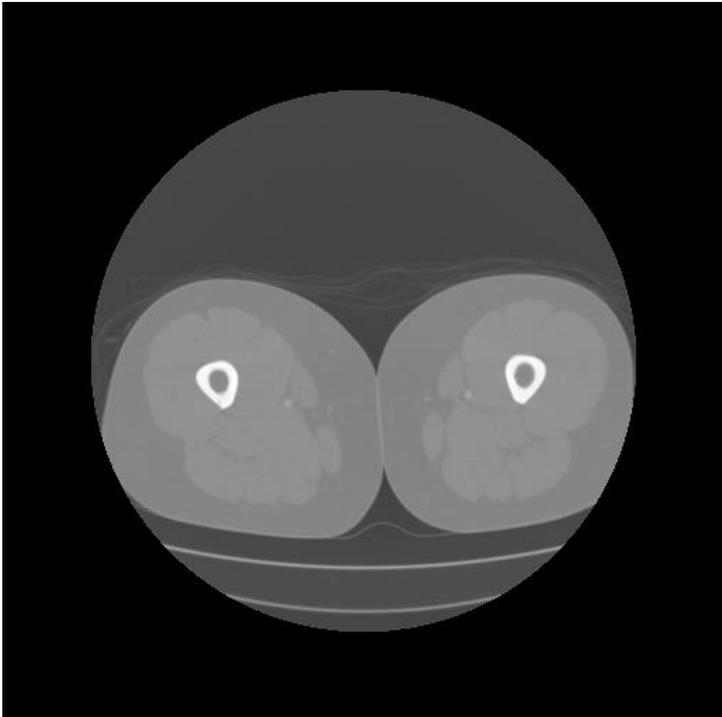
IM2



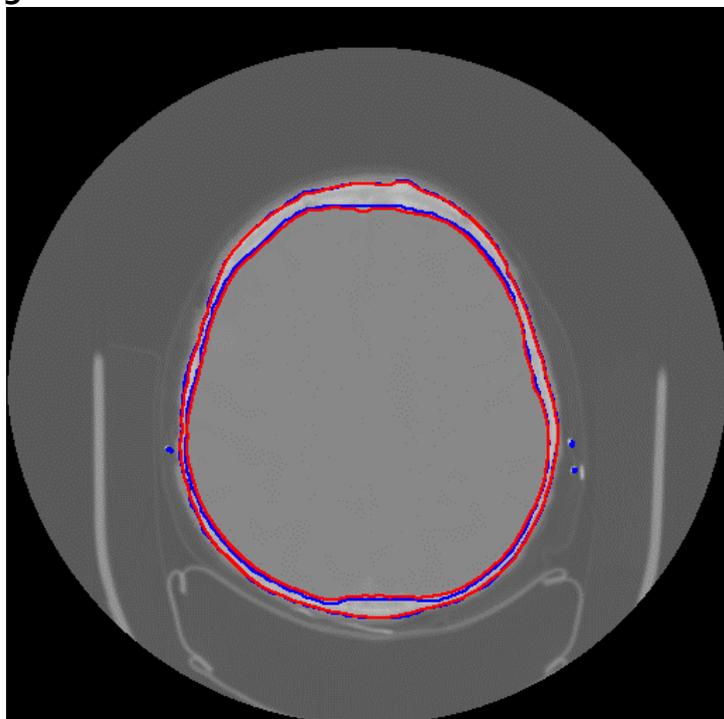
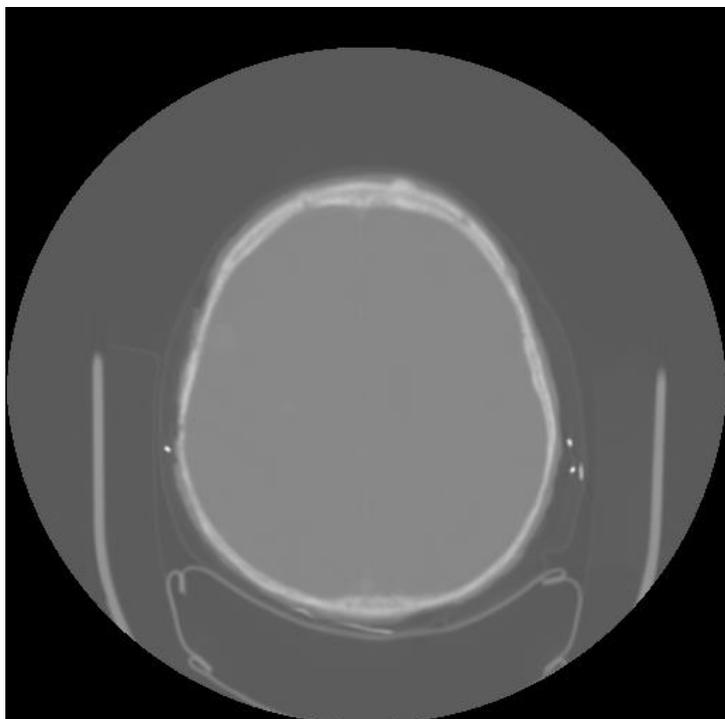
IM3



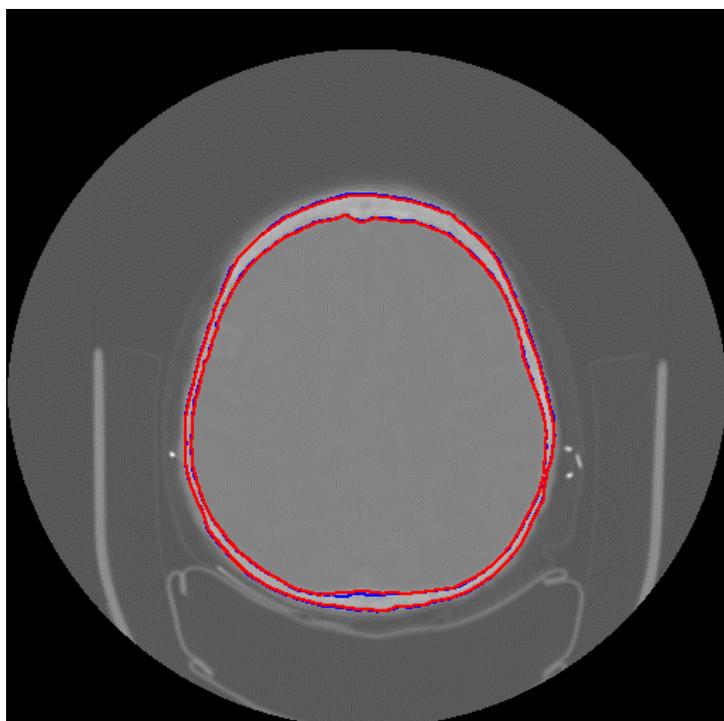
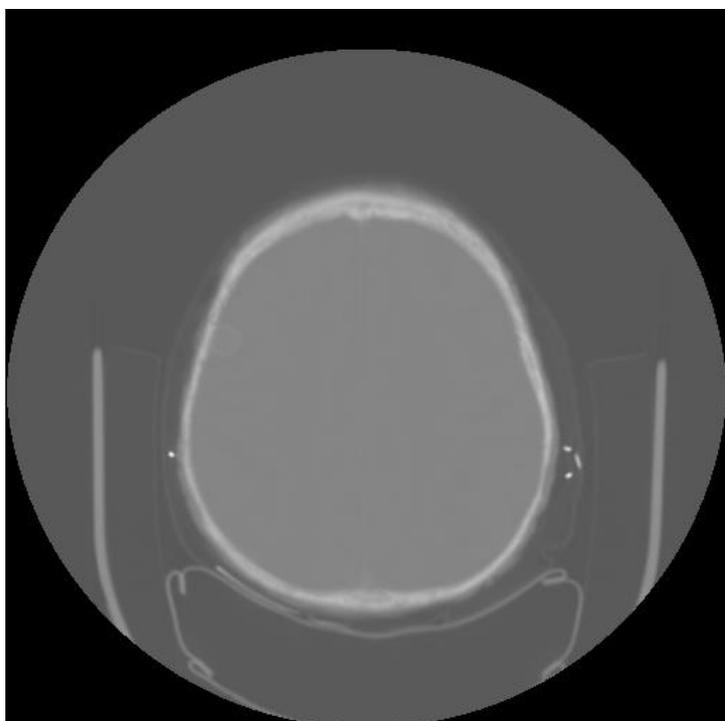
IM4



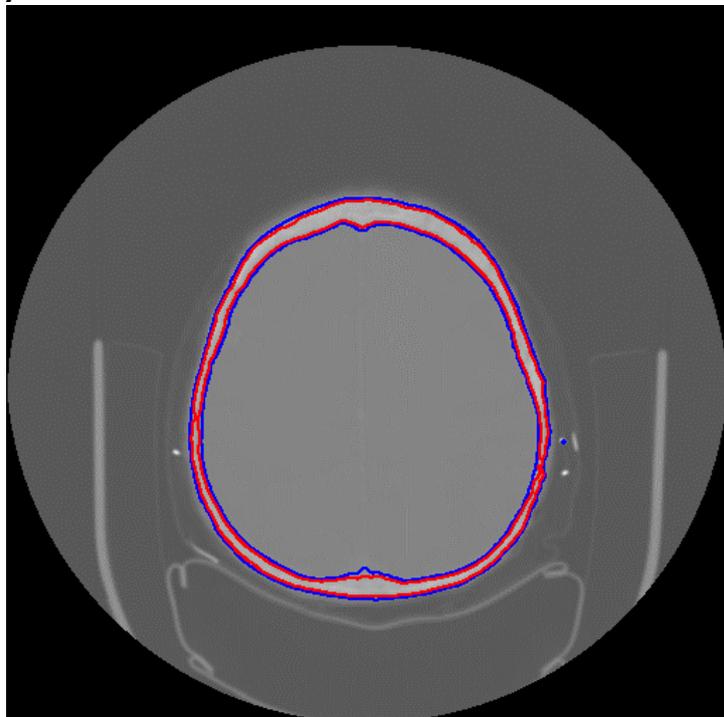
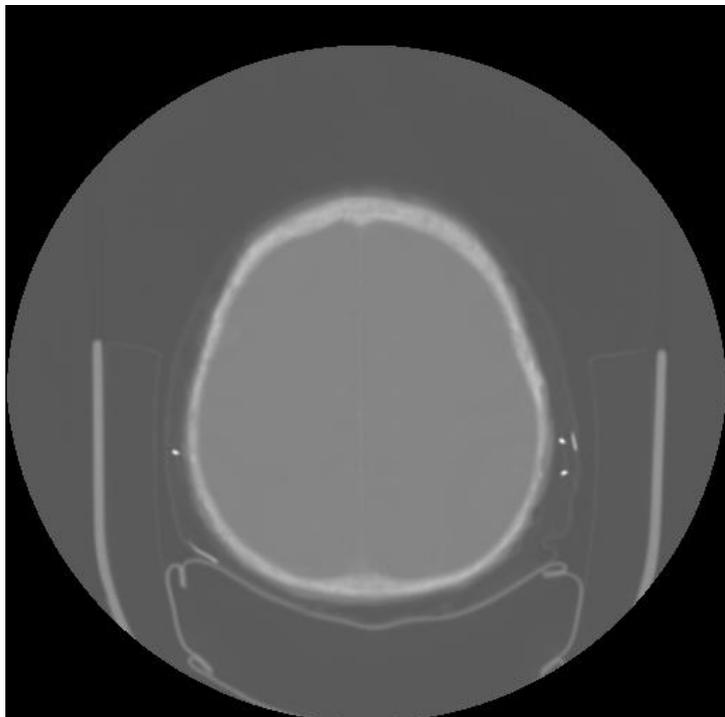
IM5



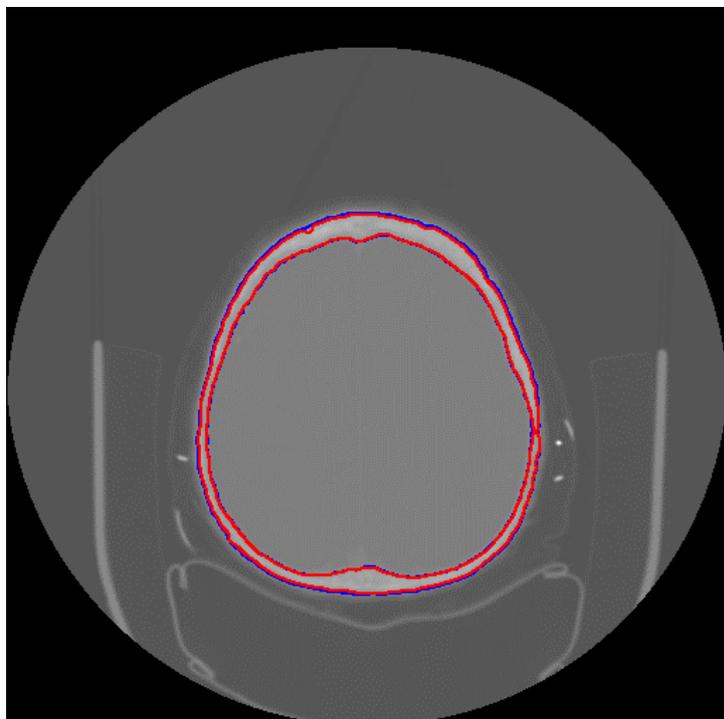
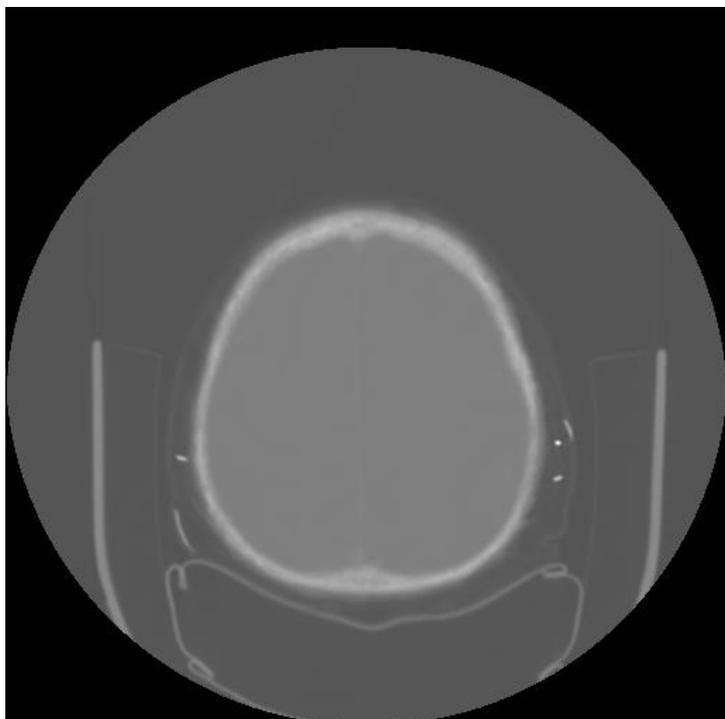
IM6



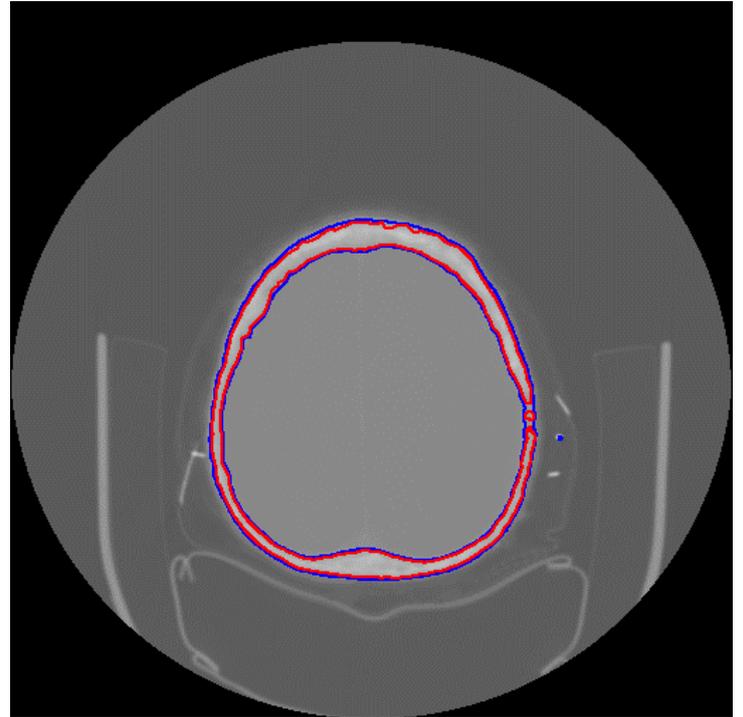
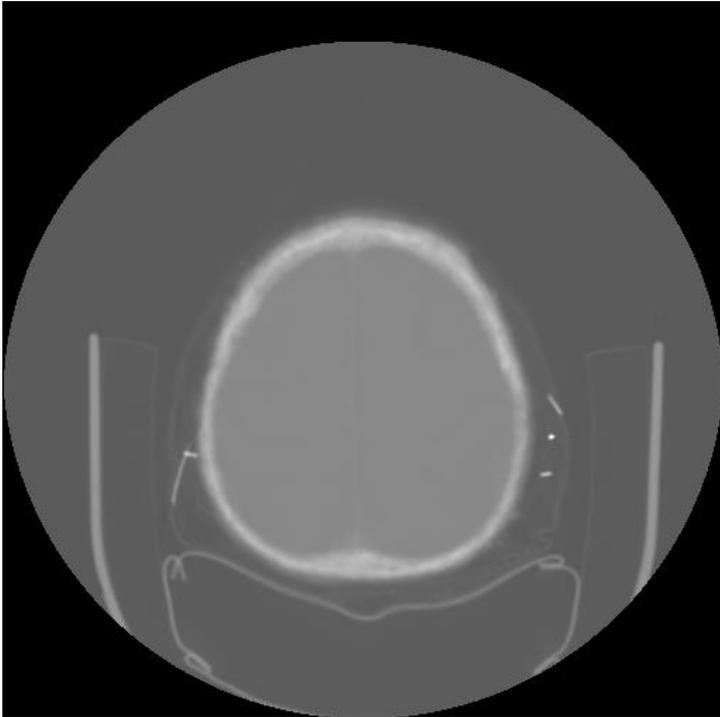
IM7



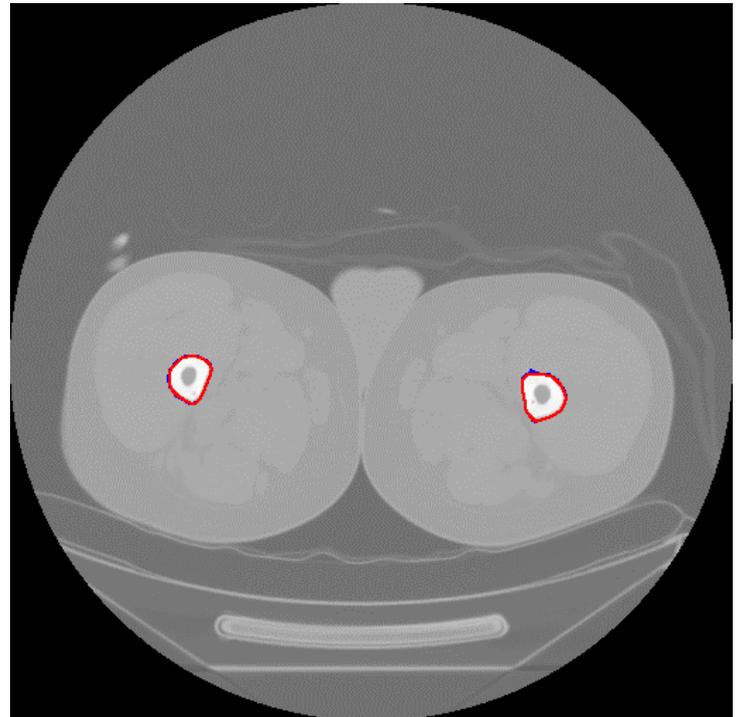
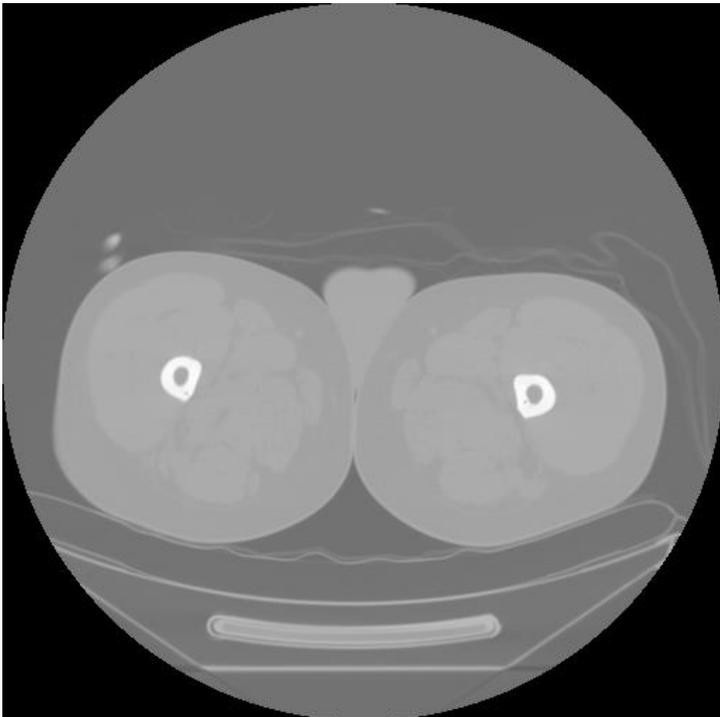
IM8



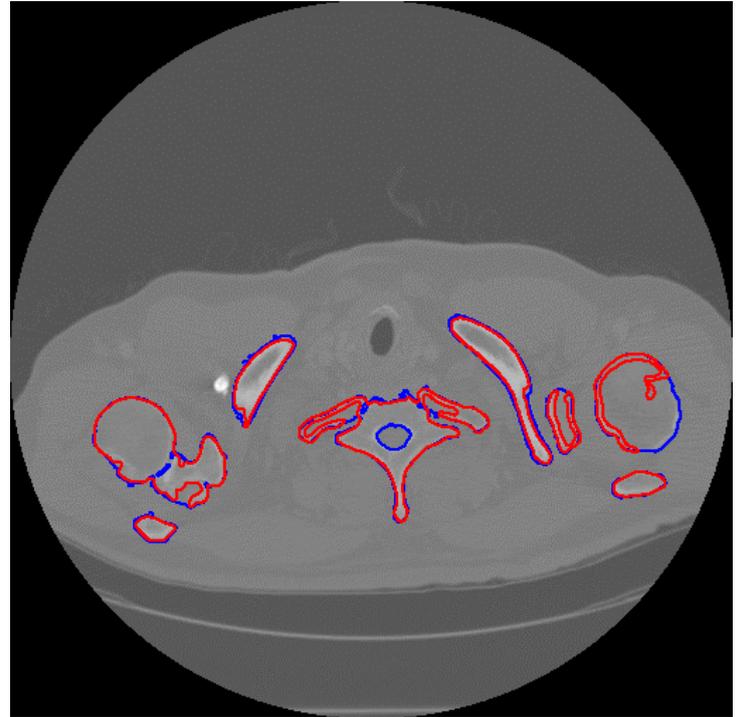
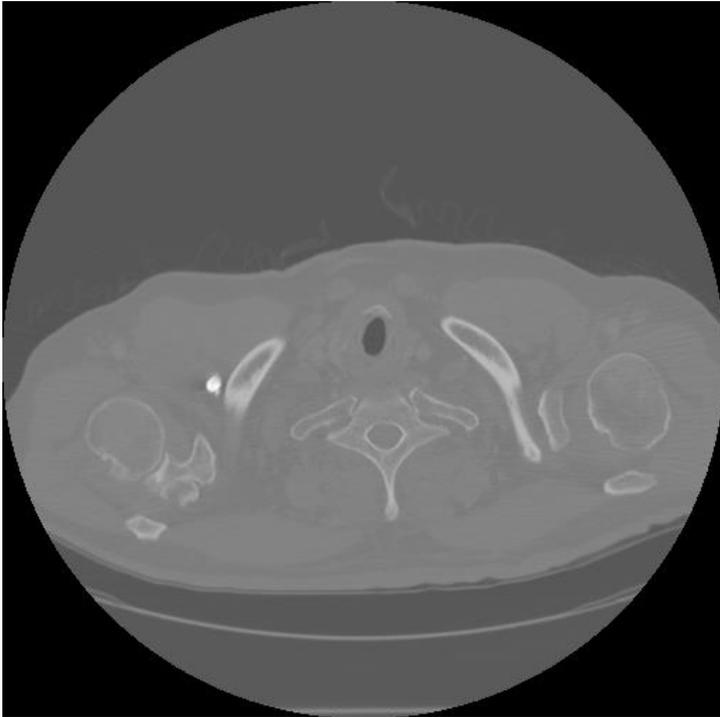
IM9



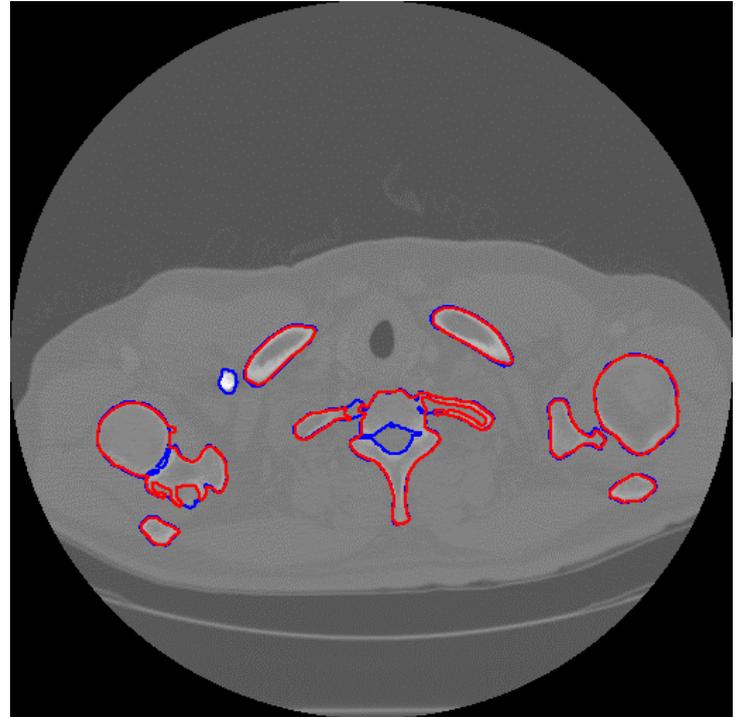
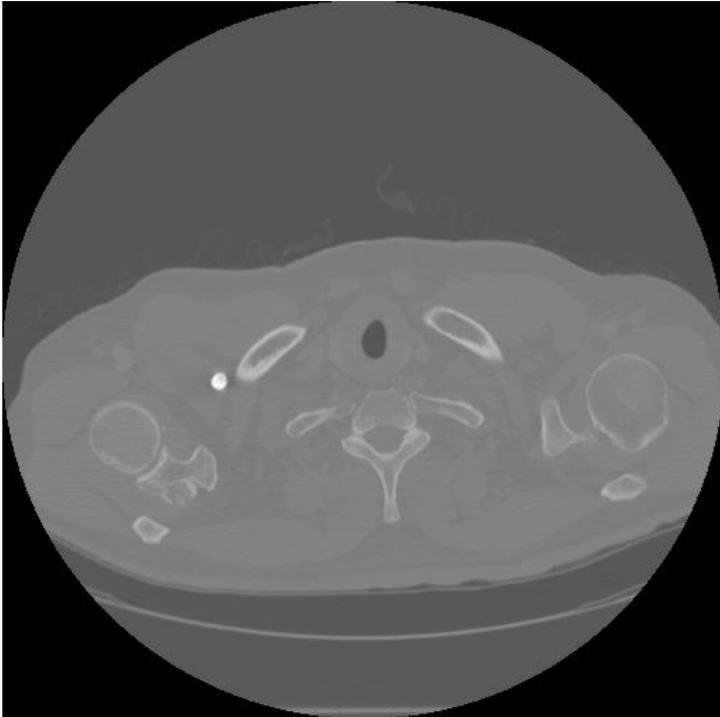
IM10



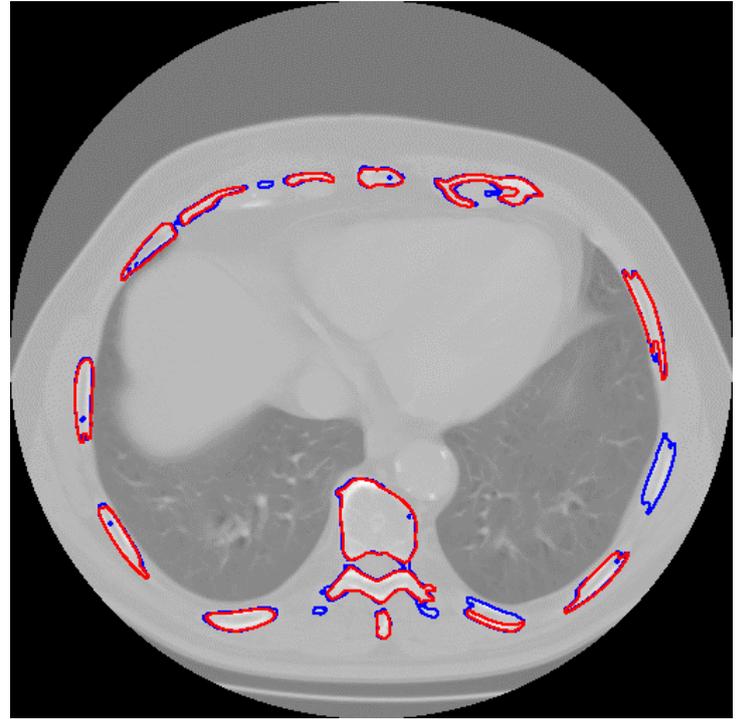
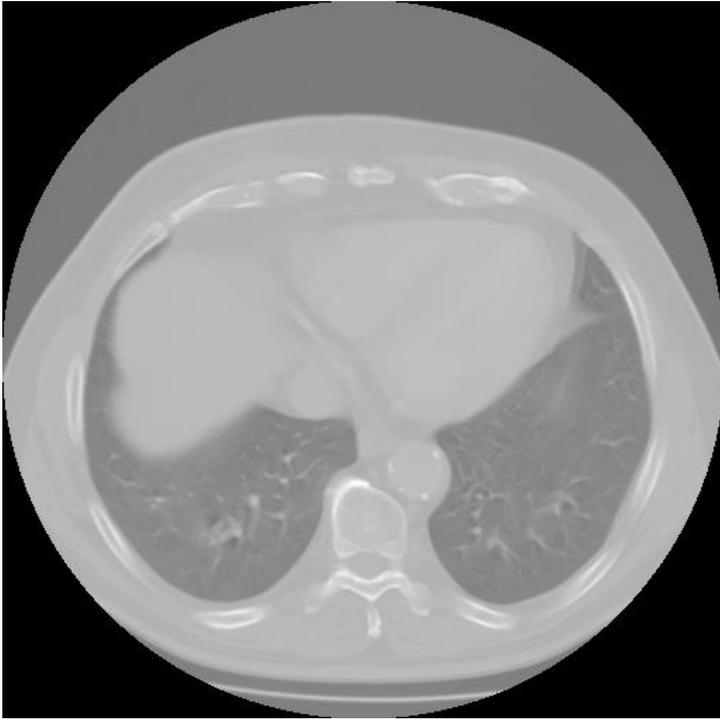
IM11



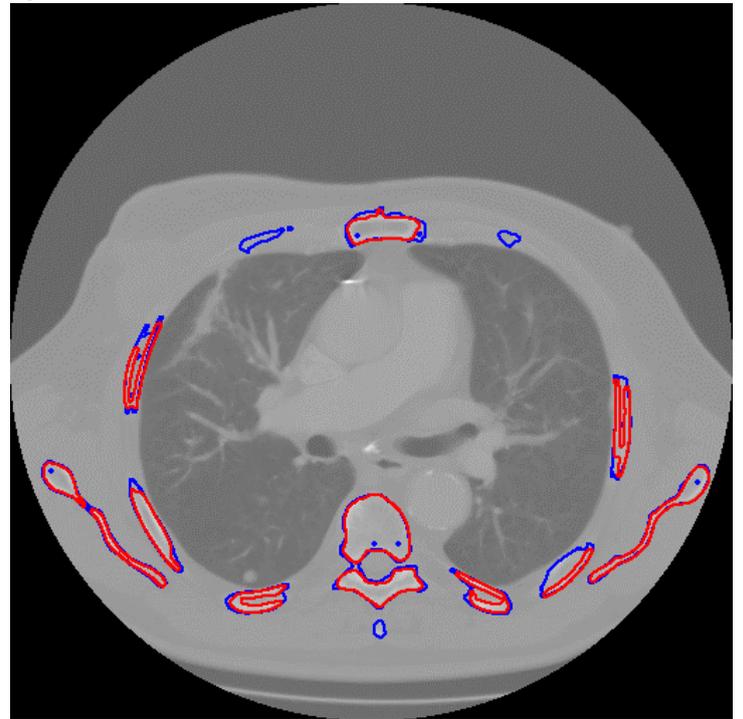
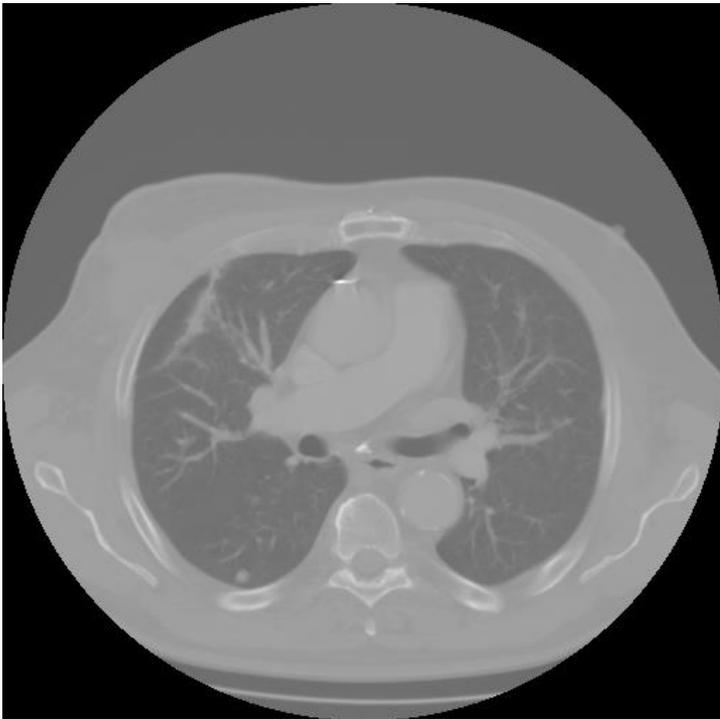
IM12



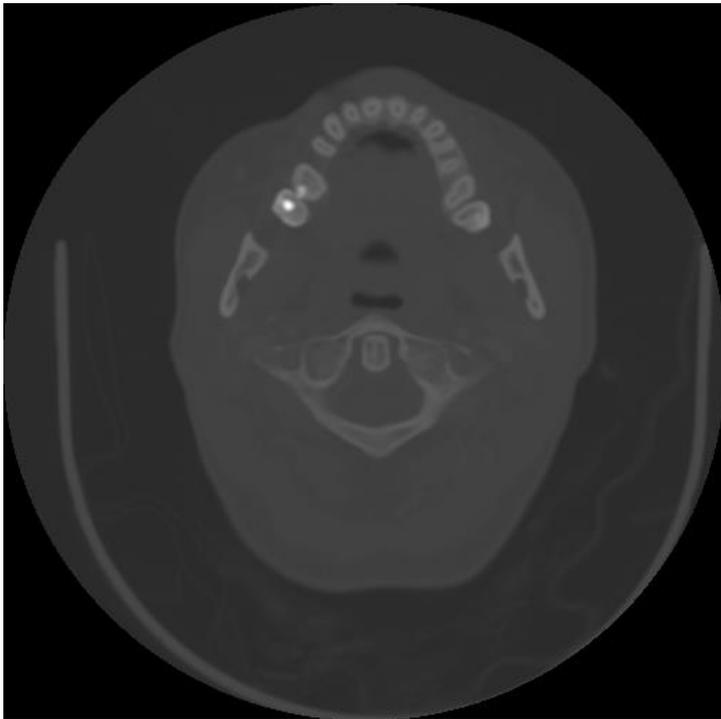
IM13



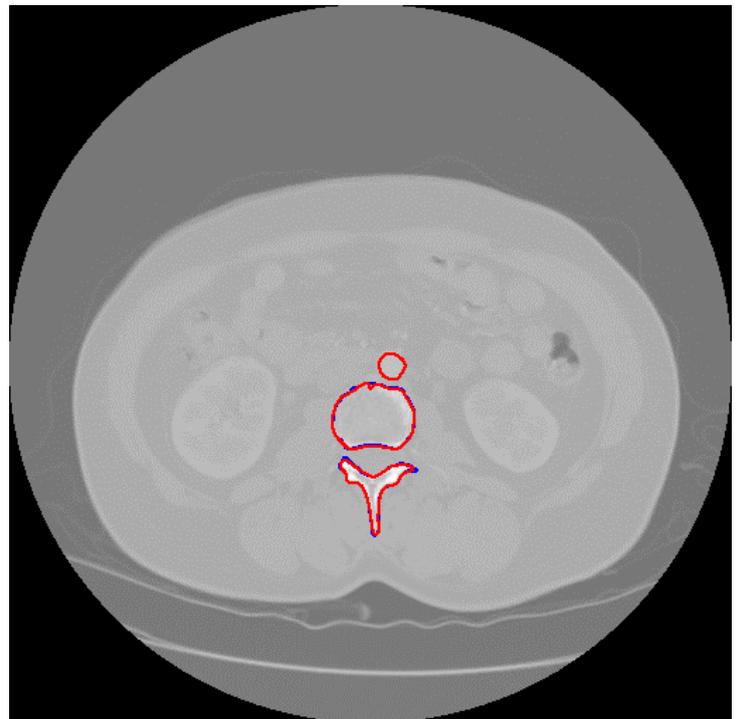
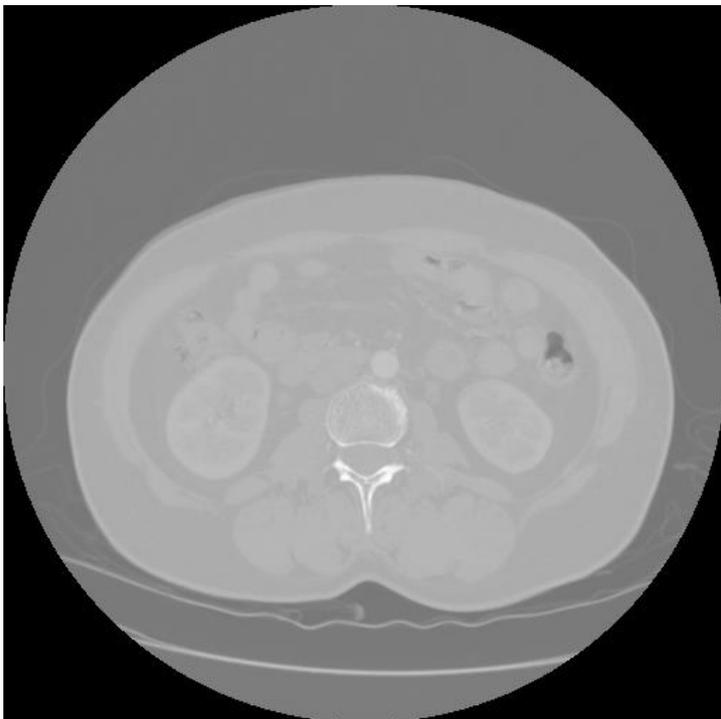
IM14



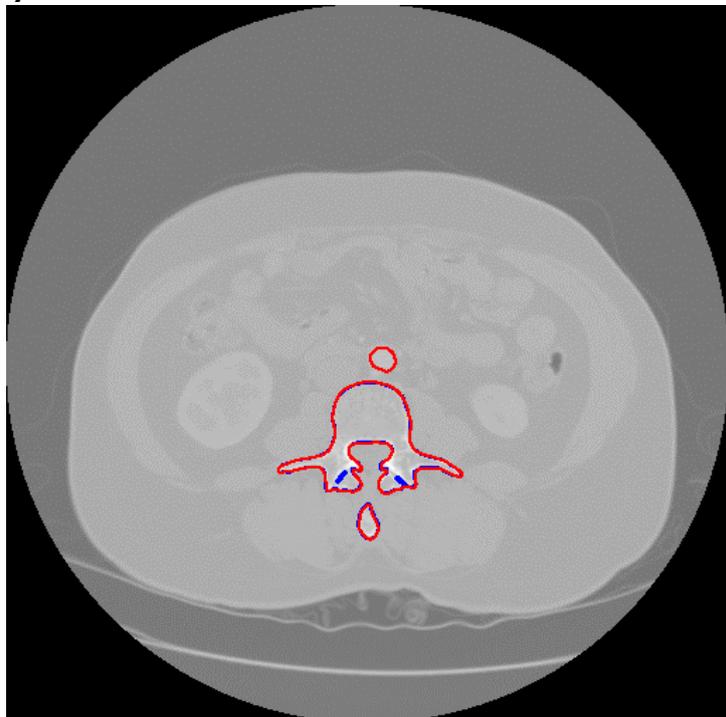
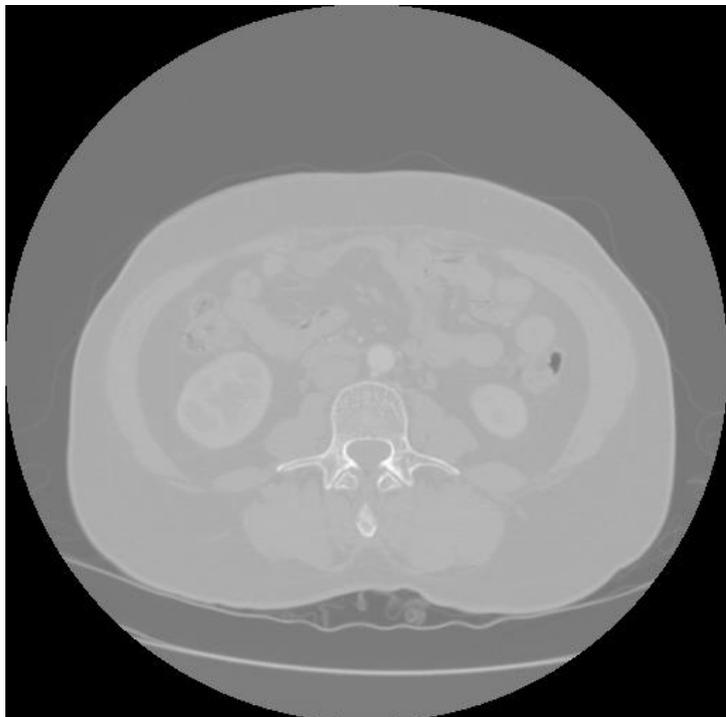
IM15



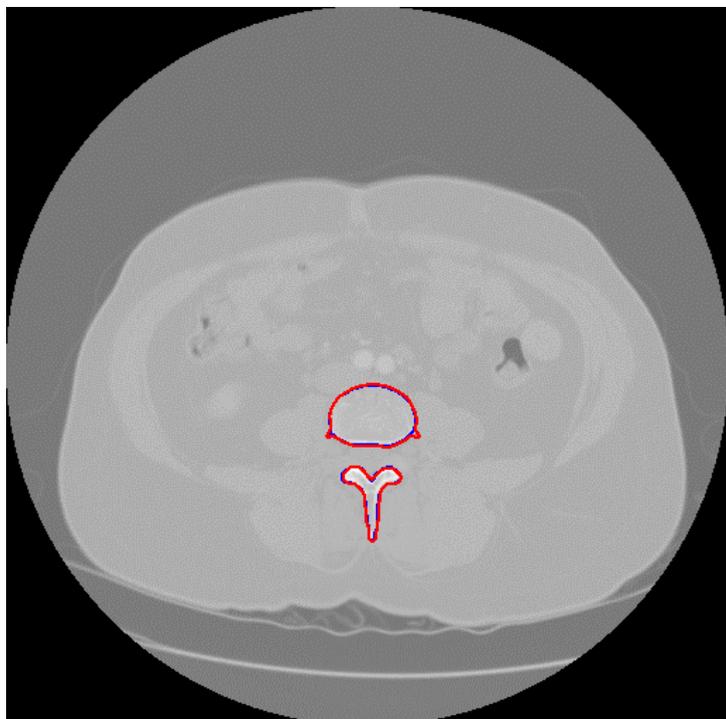
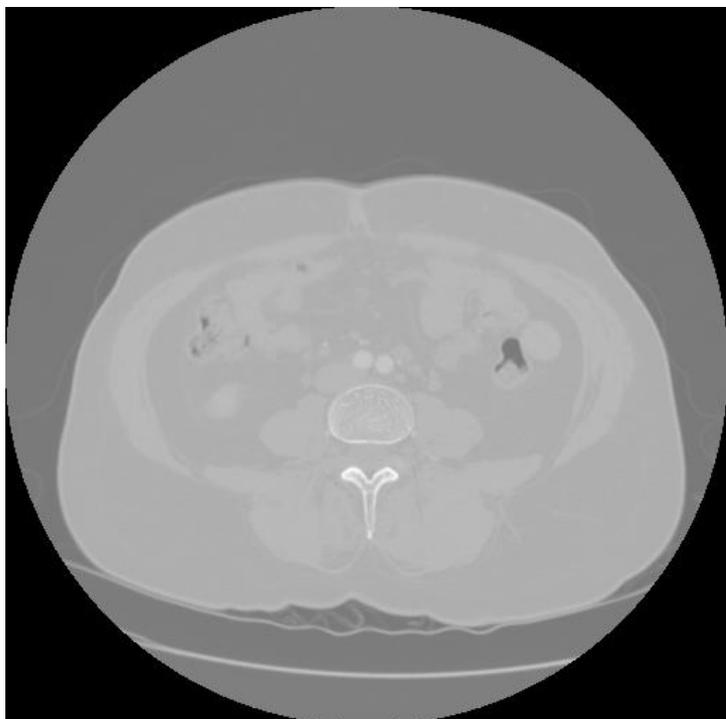
IM16



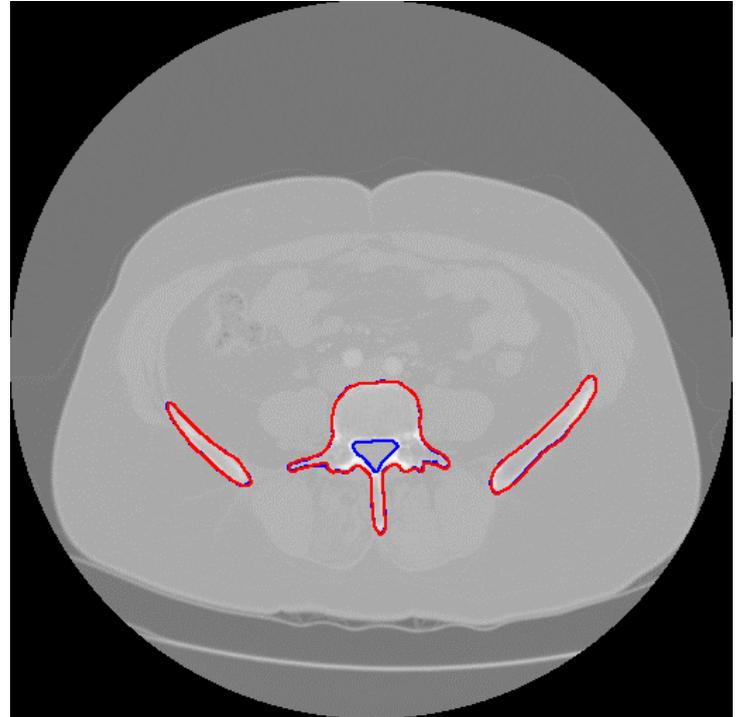
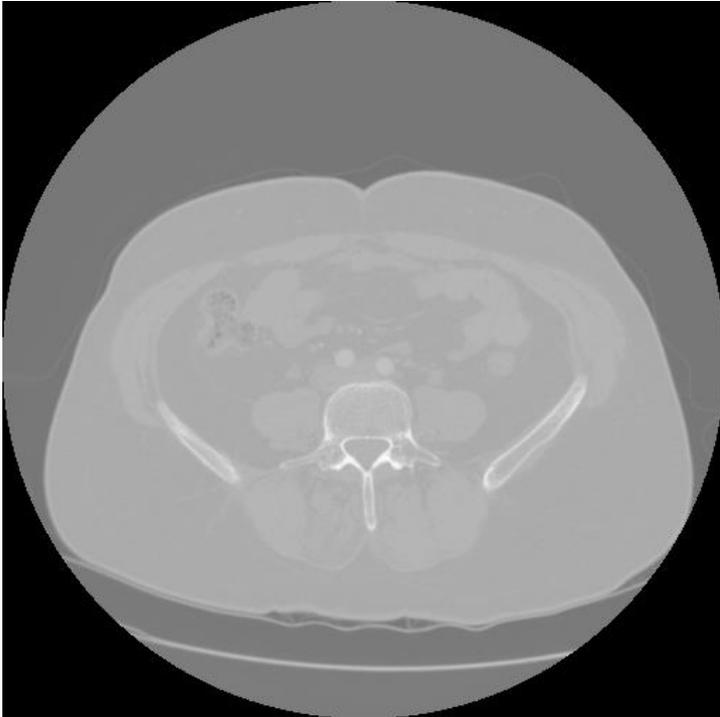
IM17



IM18



IM19



IM20

