

Nuevos métodos para el análisis  
automático del volumen de estructuras  
cerebrales a partir de imágenes de  
resonancia magnética nuclear



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

Tesis doctoral

Autor

Jose Enrique Romero Gómez

Director

Jose Vicente Manjón Herrera

6 de Marzo de 2018, Valencia (España)

Es ahora en la madurez cuando comprendo  
cuanto me querías.

Para mi abuela Ana.

## Agradecimientos

La realización de esta tesis ha sido posible gracias al compromiso y esfuerzo de muchas personas hacia las cuales siento la necesidad de dedicar unas palabras de agradecimiento.

Empezaré agradeciendo el apoyo y trabajo incansable de mi director de tesis, Jose Vicente Manjón Herrera que ha sido para mí un referente, no solo a nivel profesional, que durante todos estos años ha inculcado en mí el gusto por el buen hacer y la atención a los detalles siendo la luz que ha guiado este proyecto a buen puerto.

Quiero mencionar a mis compañeros de trabajo, que han vivido junto a mí las penas y alegrías del desarrollo de mi tesis apoyándome de forma desinteresada. Son Juanmi, Javi, Elies, Carlos, Adrian, Miguel y Salva.

Quiero mencionar también a los que fueron mis tutelados ya que enseñando también se aprende y de ellos he aprendido mucho: Fuensanta, Celia, Silvia, Jose Manuel y Jose Alberto.

No puedo dejar de agradecer el gran apoyo de mis amigos Carlos, Antonio y Óscar los cuales han sido un refugio para mí en los momentos difíciles y los primeros en alegrarse con mis logros.

Quiero también mencionar a unas personas que me dieron todo su cariño y a las que después de años sigo considerando mi familia: Jose, Charo y Carol.

Por último, quiero agradecer toda una vida de dedicación aquellos sin los cuales no sería quien soy. Ellos son mis tíos Paco y Loli, mi prima Eva y por supuesto mis padres Jose Enrique y Ana María gracias a los cuales he aprendido a afrontar la vida con decisión.

# Índice

Glosario .....	8
Capítulo 1. Introducción y estado del arte .....	11
1.1. Prólogo.....	11
1.2. El cerebro humano .....	12
1.3. Resonancia Magnética Nuclear.....	17
1.4. Segmentación .....	20
1.4.1. Definición .....	20
1.4.2. Antecedentes históricos .....	21
1.4.3. Estado del arte.....	25
1.4.4. El método Nonlocal label fusion .....	29
Capítulo 2: Hipótesis y aportaciones.....	34
2.1. Hipótesis .....	34
2.2. Aportaciones .....	34
2.3. Organización .....	35
Capítulo 3. Segmentación de hemisferios cerebrales (NABS) .....	36
3.1. Segmentación del cerebro.....	36
3.2. Materiales y métodos .....	37
3.2.1. Creación de la librería .....	37
3.2.2. Block-wise Nonlocal label fusion aplicado a múltiples clases .....	39
3.2.3. Reducción del área de proceso usando información a priori .....	40
3.3. Experimentos y resultados .....	42
3.3.1. Tamaño del área de búsqueda .....	42
3.3.2. Tamaño de la librería .....	42
3.3.3. Implementación Block-wise.....	43
3.3.4. Comparativa con el método Adaptive Disconnection .....	44
3.3.5. Estimación de la asimetría .....	46
3.3.6. Tolerancia a lesiones y atrofia.....	47
3.3.7. Estimación de la asimetría en pacientes con demencia .....	48
3.3.8. Coste computacional .....	51
3.4. Conclusión y discusión .....	51
Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES) .....	52
4.1. Segmentación del cerebelo .....	52
4.2. Material y métodos .....	54

4.2.1. Datos .....	54
4.2.2. Preprocesado .....	55
4.2.3. Creación de la librería .....	57
4.2.3. Optimized PatchMatch label fusión .....	57
4.2.4. Regularización de etiquetas .....	60
4.2.5. Extensión de la librería.....	61
4.3. Experimentos y resultados .....	61
4.3.1. Parámetros del método.....	61
4.3.2. Construcción de la librería .....	61
4.3.3. Multi-escala adaptativo .....	62
4.3.4. Regularización de etiquetas.....	63
4.3.5. Extensión de la librería.....	64
4.3.6. Comparativa de métodos .....	65
4.3.7. Tiempo de ejecución .....	66
4.4. Conclusión y discusión .....	67
Capítulo 5: Segmentación de subregiones del hipocampo (HIPS) .....	69
5.1. Segmentación del hipocampo .....	69
5.2. Materiales y métodos .....	72
5.2.1. Datos .....	72
5.2.2. Preprocesado .....	73
5.2.3. Construcción de la librería .....	74
5.2.4. Multiespectral Optimized PatchMatch (MOPAL).....	74
5.2.5. Multiscale label fusion .....	75
5.2.6. Corrector de errores sistemáticos .....	75
5.3. Experimentos y resultados .....	77
5.3.1. Parámetros de MOPAL.....	78
5.3.2. Corrector de errores sistemáticos .....	79
5.3.3. Resolución estándar frente a alta resolución.....	81
5.3.4. Comparativa de métodos .....	84
5.4. Conclusión y discusión .....	86
Capítulo 6: Conclusiones generales y líneas de futuro.....	88
Publicaciones realizadas .....	91
Revistas .....	91
Conferencias .....	91
Bibliografía.....	92

## Resumen

En la actualidad, la medicina y en especial el área de imagen médica, ha sido uno de los campos de la ciencia que más se ha beneficiado de las tecnologías de la información. En esta tesis nos centramos en las imágenes de resonancia magnética cerebral y los métodos de segmentación automática del volumen cerebral.

En el primer capítulo describimos los antecedentes de este trabajo mediante una breve introducción de los principios físicos de la resonancia magnética y una revisión del estado del arte en relación con las técnicas de segmentación y su evolución a lo largo de las últimas décadas.

En los capítulos tres, cuatro y cinco presentamos tres métodos de segmentación automática de diferentes partes del cerebro que mejoran el estado del arte en términos de calidad del resultado y velocidad. Finalmente, en el capítulo seis se comentan las conclusiones generales y se proponen líneas de desarrollo futuras.

## Resum

En l'actualitat, la medicina i en especial l'àrea d'imatge mèdica, ha estat un dels camps de la ciència que més s'ha beneficiat de les tecnologies de la informació. En aquesta tesi ens centrem en les imatges de ressonància magnètica cerebral i el desenvolupament de mètodes de segmentació automàtica del volum cerebral.

En el primer capítol descrivim els antecedents d'aquest treball mitjançant una breu introducció dels principis físics de la ressonància magnètica i una revisió de l'estat de l'art en relació amb les tècniques de segmentació i la seva evolució al llarg de les últimes dècades.

En els capítols tres, quatre, i cinc presentem tres mètodes de segmentació automàtica per a diferents parts del cervell que milloren l'estat de l'art en quant a qualitat dels seus resultats i velocitat. Finalment, en el capítol sis es comenten les conclusions generals i es proposen línies de desenvolupament i es proposen línies de desenvolupament futures.

## **Abstract**

Nowadays, medicine and specially the medical image area, is one of the science fields that has benefited more from the information technologies. In this thesis we focus on cerebral magnetic resonance imaging and automatic segmentation methods of the brain volume.

In the first chapter we describe a background for this work by a brief introduction of the physic fundamentals of magnetic resonance and a revision of the state-of-the-art regarding to segmentation techniques and its evolution along the last decades.

In the chapters three, four and five we present three methods for automatically segment different parts of the brain that leverage the state-of-the-art in terms of result quality and computation time. Finally, in chapter six we comment the general conclusions and propose lines for future works.

# Glosario

## A

- AD: Enfermedad de Alzheimer (Alzheimer's Disease).
- ANNF: Mapa aproximado de vecino más próximo (Approximated Nearest Neighbor Field).
- ANTs: Software de registro de imagen (Advanced Normalization Tools).
- AUC: Dada una gráfica, valor del área bajo la curva (Area Under the Curve).

## B

- Boosting: Estrategia de entrenamiento de redes neuronales.
- B-Spline: Método de interpolación basado en curvas diferenciables.

## C

- CDR: Grado de demencia (Clinical Dementia Rating).
- CPU: Unidad central de proceso (Central Process Unit).
- CSF: Líquido cefalorraquídeo (del inglés CerebroSpinal Fluid).

## D

- Dataset: Conjunto de imágenes.
- DP: Densidad protónica.

## E

- EM: Estimador de máxima verosimilitud (Expectation-Maximization).
- Espín: Término tomado del inglés que significa rotación.

## G

- GPU: Unidad de procesamiento gráfico (Graphical Process Unit).

## H

- HC: Hipocampo.
- HR: Alta resolución (High Resolution).

## I

- IRM: Imagen de resonancia magnética nuclear.

## K

- Kernel: Matriz de valores utilizada como base de convolución.

## L

- LASR: Método de superresolución de imágenes (Local Adaptive SuperResolution).



- LCR: Líquido cefalorraquídeo.
- Lóbulo: porción de un órgano que se considera como una unidad diferenciada por sus características anatómicas o funcionales.
- LOO: Validación cruzada excluyendo el sujeto bajo evaluación (Leave One Out).

## M

- MAD: Promedio del valor absoluto de la substracción (Mean Absolute Difference).
- MAGeT: Método de segmentación de múltiples estructuras (Multiple Automatically Generated Templates).
- MALF: Fusión de etiquetas basada en múltiples atlas (Multi Atlas Label Fusion).
- MAS: Segmentación basada en múltiples atlas (Multi Atlas Segmentation).
- MNI: Instituto neurológico de Montreal (Montreal Neurological Institute).
- MSN: Métrica multi-contraste de diferencia entre parches (Multispectral Semi Norm).

## N

- N4: Software de corrección de inhomogeneidad en imágenes de resonancia magnética cerebral.
- NICE: Método de segmentación del volumen intracraneal (Nonlocal Intracranial Cavity Extraction).

## O

- OPAL: Método de segmentación basado en parches (Optimized PAtchmatch Label fusion).

## P

- Perceptrón: Red neuronal con una topología multicapa básica.
- PET: Tomografía por emisión de positrones (Positron Emission Tomography).

## R

- RASCAL: Método de segmentación del cerebelo (Rapid Automatic Segmentation of the human cerebellum And its Lobules).
- ReLU: Unidad lineal rectificadora (Rectified Linear Unit).
- RF: Radiofrecuencia.
- RMN: Resonancia Magnética Nuclear.
- ROC: Curva de rendimiento de un clasificador (Receiver-Operating Characteristic).
- ROI: Región de interés (Region Of Interest).

## S

- SANLM: Filtro adaptativo a ruido espacialmente variable para imágenes (Spatially Adaptive Non-Local Means).

- SB: Sustancia Blanca.
- SG: Sustancia Gris.
- Sobrecompleto: Enfoque de segmentación basada en parches en la que estos solapan entre sí.
- SPECT: Tomografía computarizada de emisión monofotónica (Single Photon Emission Computed Tomography).
- SPM: Software de procesado de imagen cerebral (Statistical Parametrical Mapping).
- SSD: Suma de diferencias al cuadrado (Sum of Squared Differences).
- SUIT: Software de segmentación del cerebelo (Spatially Unblased Template).
- SurfPatch: Método de segmentación automática de hipocampo.

## T

- T1: Constante de relajación longitudinal.
- T2: Constante de relajación transversal.
- TAC: Tomografía Axial Computarizada.

## V

- Vóxel: Unidad de la que se compone una imagen tridimensional. Es el equivalente en tres dimensiones a un pixel.

## Capítulo 1. Introducción y estado del arte

### 1.1. Prólogo

Resulta difícil pensar en algún campo de la ciencia que no se haya visto revolucionado por las tecnologías de la información. La medicina ha sido probablemente una de las áreas más beneficiadas por la incorporación de técnicas que permiten obtener información valiosísima que ha hecho posible diagnósticos y tratamientos que unas pocas décadas atrás eran impensables. En concreto, las técnicas de imagen médica han cambiado radicalmente la forma en la que vemos el interior del cuerpo humano y la información que podemos obtener de él.

La aparición de las primeras técnicas de imagen se remontan al siglo XIX con el descubrimiento de los rayos X por parte de Wilhelm C. Roentgen gracias a los cuales se hizo posible contemplar el interior del cuerpo sin necesidad realizar arriesgados procedimientos.

En la actualidad, las técnicas de imagen han evolucionado y existen diversas modalidades como la Resonancia Magnética Nuclear (RMN), la ecografía, la Tomografía Axial Computarizada (TAC) o las técnicas de medicina nuclear (como PET o SPECT).

Una de las técnicas más empleadas en el diagnóstico médico es la RMN que junto con técnicas de procesado de imagen permiten la obtención de información cuantitativa muy útil tanto en ámbitos de investigación como en la práctica clínica. Dichas técnicas de procesado permiten realizar un análisis objetivo, reproducible y automático de las imágenes. Sin ellas, la información que podemos obtener de las imágenes se limita a un análisis cualitativo (es decir, a la inspección visual por parte de un experto). Sin embargo, usando técnicas de procesado podemos estudiar parámetros físicos concretos como la densidad, el volumen, etc.

Una de las técnicas de procesado en las que más se ha trabajado en los últimos años es la segmentación de imágenes médicas. En especial, la segmentación de imágenes de RMN cerebral, donde la medición de los volúmenes asociados a los tejidos que componen el cerebro, a las estructuras anatómicas en las que se divide e incluso la medición de regiones anómalas como lesiones contribuyen a un seguimiento más preciso y un diagnóstico más temprano de enfermedades como el Alzheimer ([Dickerson and Sperling, 2005](#); [Jack et al., 2005, 2000](#)) o la esquizofrenia ([Okugawa et al., 2002](#)) entre muchas otras.

En esta tesis, se presentan tres métodos de segmentación de estructuras cerebrales que permiten la estimación automática y precisa de los volúmenes de dichas estructuras a partir de imágenes de RMN cerebral.

## 1.2. El cerebro humano

El cerebro, del latín *cerebrum* cuyo significado arcaico es “lo que se lleva en la cabeza” es un término general que hace referencia al carácter central del sistema nervioso más complejo del reino animal. En la actualidad, al contenido de cráneo se le denomina encéfalo aun que comúnmente se sigue llamando cerebro a todo el contenido craneal. La figura 1.2.1 muestra una ilustración del encéfalo humano.



Figura 1.2.1: Ilustración del cerebro humano.

El cerebro es el órgano de mayor tamaño del sistema nervioso central y desempeña una gran cantidad de funciones de regulación y cognitivas como la memoria, las emociones, el razonamiento y el lenguaje. Se encuentra alojado en el interior del cráneo bañado en un líquido llamado líquido cefalorraquídeo ([Parent and Carpenter, 1996](#)).

La unidad funcional más pequeña del cerebro son las neuronas (figura 1.2.2). Células que interpretan señales eléctricas (impulsos nerviosos) y químicas (neurotransmisores y hormonas) y generan una respuesta eléctrica o química. Estas células se interconectan entre sí formando complejas redes que dan origen a las funciones antes mencionadas. Dichas conexiones se realizan mediante prolongaciones del cuerpo de la neurona llamadas dendritas y axón. Las dendritas reciben señales que tras ser integradas pueden desencadenar una respuesta que se propaga a otras neuronas a través del axón. El axón, de mayor longitud que las dendritas, actúa como cable conductor haciendo llegar el impulso desde el cuerpo de la neurona (o soma) que lo origina a las dendritas de las neuronas que lo reciben ([Cajal, 2012](#)).

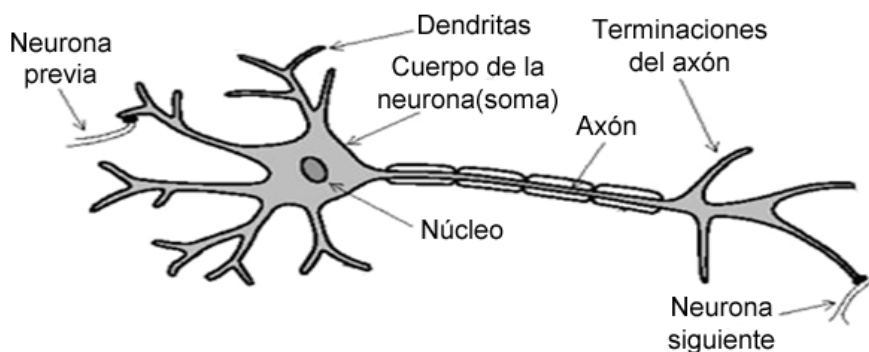


Figura 1.2.2: Ilustración de una neurona.

Las neuronas forman parte de los dos tejidos constituyentes del cerebro que son la sustancia gris (SG) formada fundamentalmente por los cuerpos de las neuronas y la sustancia blanca (SB) formada mayoritariamente por los axones de las neuronas. Completando el espacio intracraneal se encuentra el líquido cefalorraquídeo (LCR) el cual baña al encéfalo y la médula espinal. La figura 1.2.3 muestra la distribución espacial de estos tres tejidos. A partir de la descripción anterior de la neurona podemos decir que la sustancia gris corresponde a las regiones donde se procesa la información mientras que la sustancia blanca es responsable de interconectar los diferentes centros de sustancia gris. El líquido cefalorraquídeo baña al encéfalo y circula a través de él. Proporciona protección contra traumatismos, participa en la nutrición de sus tejidos y actúa como vía para la eliminación de subproductos del metabolismo entre otras funciones (Sakka et al., 2011).

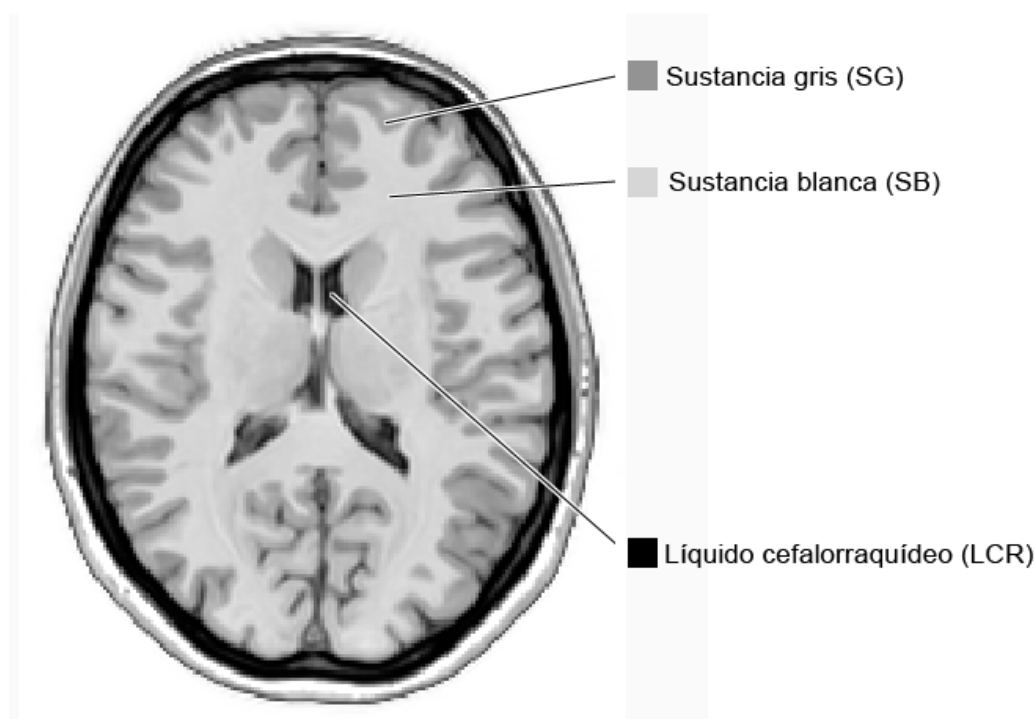


Figura 1.2.3: Imagen de RMN cerebral señalando los tres tejidos fundamentales que forman el contenido intracraneal.

Además de la clasificación de los ya mencionados tejidos, el encéfalo se divide en diferentes áreas anatómicas y funcionales siendo la división más general la de sus macroestructuras que son el cerebro, dividido en dos hemisferios cerebrales, izquierdo y derecho (que representan la mayor parte del volumen intracraneal) separados por la fisura interhemisférica e interconectados mediante una estructura de SB llamada cuerpo calloso, el cerebelo, dividido en los hemisferios cerebelares izquierdo y derecho y por último el tronco cerebral a partir del cual se forma la médula espinal. En el capítulo 3 se abordará con más detalle la estructura del encéfalo.

En la superficie de los hemisferios cerebrales se encuentra la corteza cerebral. Una formación de SG que se subdivide en los lóbulos frontal, parietal, temporal y occipital (siendo estos bilaterales). Estos lóbulos reciben el nombre de los huesos que los recubren (figura 1.2.4).

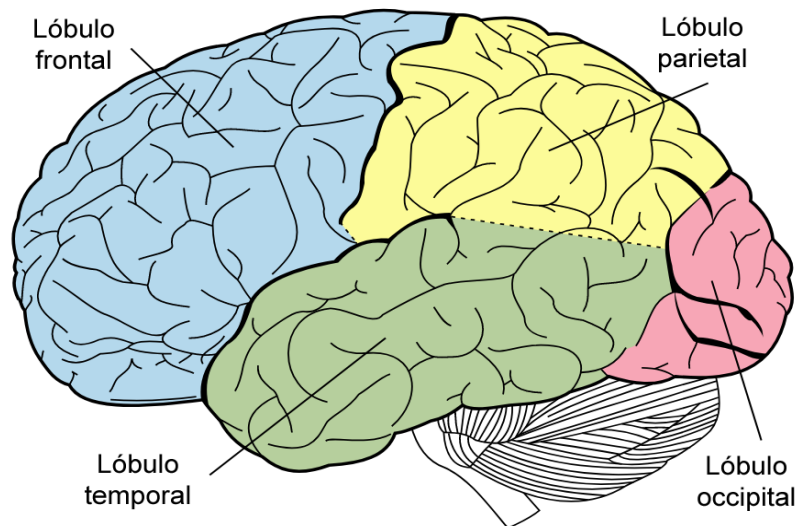


Figura 1.2.4: Ilustración de los lóbulos cerebrales.

Profundizando en los hemisferios cerebrales se encuentran una serie de estructuras denominadas ganglios basales y el tálamo. Se trata de núcleos de SG que se encuentran en la base del cerebro entre las vías nerviosas, sobre el tronco del encéfalo. Los ganglios basales relacionados mayormente con la función motora interconectan la corteza motora con el tronco encefálico mientras que el tálamo (formado por 80 núcleos neuronales) se ocupa de filtrar los estímulos que llegan al cerebro (excepto los olfativos) decidiendo cuales llegan a la corteza y cuales se descartan. Los ganglios basales son el caudado, el globo pálido, el putamen y la amígdala (figura 1.2.5). Entre ellos se encuentran 4 cavidades por las que circula el LCR llamadas ventrículos (Herrero et al., 2002).

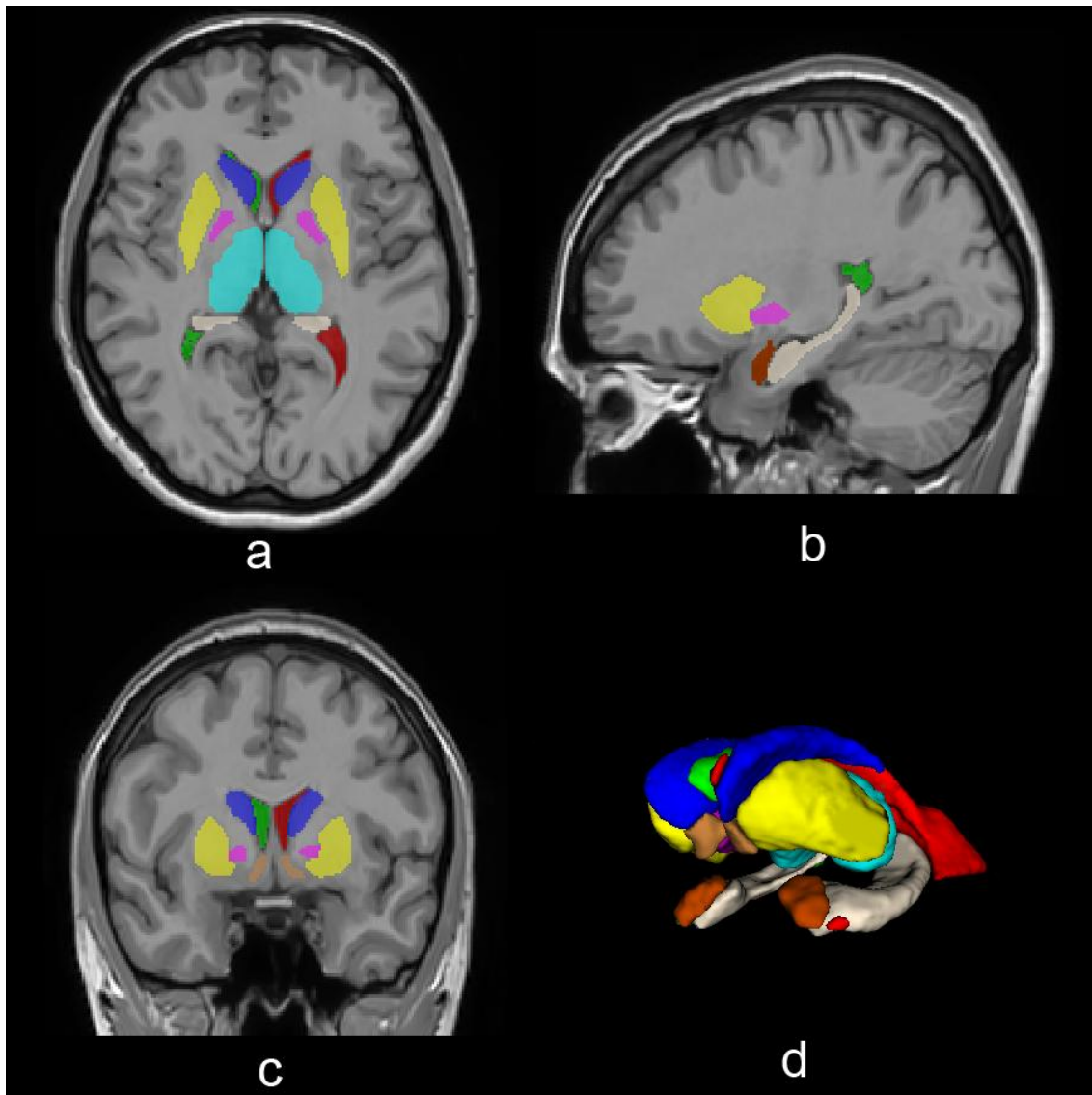


Figura 1.2.5: Ejemplo de segmentación manual de estructuras subcorticales: Hipocampo (Blanco), Tálamo (Cian), Putamen (Amarillo), Globo pálido (Rosa), Caudado (Azul), Núcleo acumbens (Marrón), Amígdala (Naranja), Ventrículo izquierdo (Verde) y Ventrículo derecho (Rojo). Se muestran los cortes axial (a), sagital (b), coronal (c) y representación tridimensional (d). Fuente: Plataforma volBrain.upv.es.

Próximo a los ganglios basales, tras el lóbulo temporal, se encuentra una estructura bilateral llamada hipocampo que es responsable de la gestión de la memoria. Se trata de una formación de gran complejidad anatómica formada por capas que pertenecen tanto al sistema límbico como a la corteza cerebral. Su anatomía se divide en segmentos que van desde el *dentate gyrus* en su parte más superior, continua con el *cornu amonis* (dividido en subsegmentos desde el CA1 a CA4) y termina en el *subículum* en su parte más inferior. Además posee capas diferenciadas llamadas *strata* que contienen neuronas especializadas (Duvernoy et al., 2013). En el capítulo 5 se abordará en mayor detalle la anatomía del hipocampo. En la figura 1.2.6 puede verse una imagen del hipocampo y sus partes constituyentes.

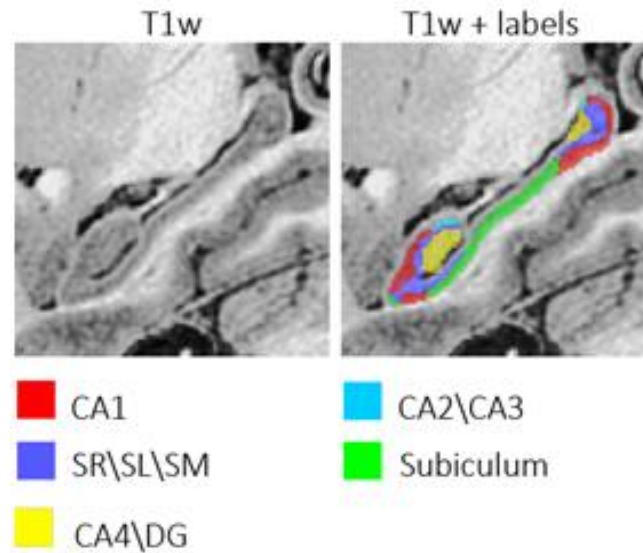


Figura 1.2.6: Ejemplo de segmentación del hipocampo en 5 subregiones.

Finalmente, bajo el cerebro se encuentra el cerebelo, con una estructura similar al cerebro pero de menor tamaño. Este, se subdivide en lóbulos que a su vez se dividen en lobulillos. En el capítulo 4 se describirá en detalle la anatomía del cerebelo y sus lóbulos. Su función principal es integrar las señales motoras y sensitivas aunque también se le atribuye un papel importante en procesos cognitivos como la atención y las emociones (Manto et al., 2013). La figura 1.2.7 muestra una ilustración de un corte axial del cerebelo.

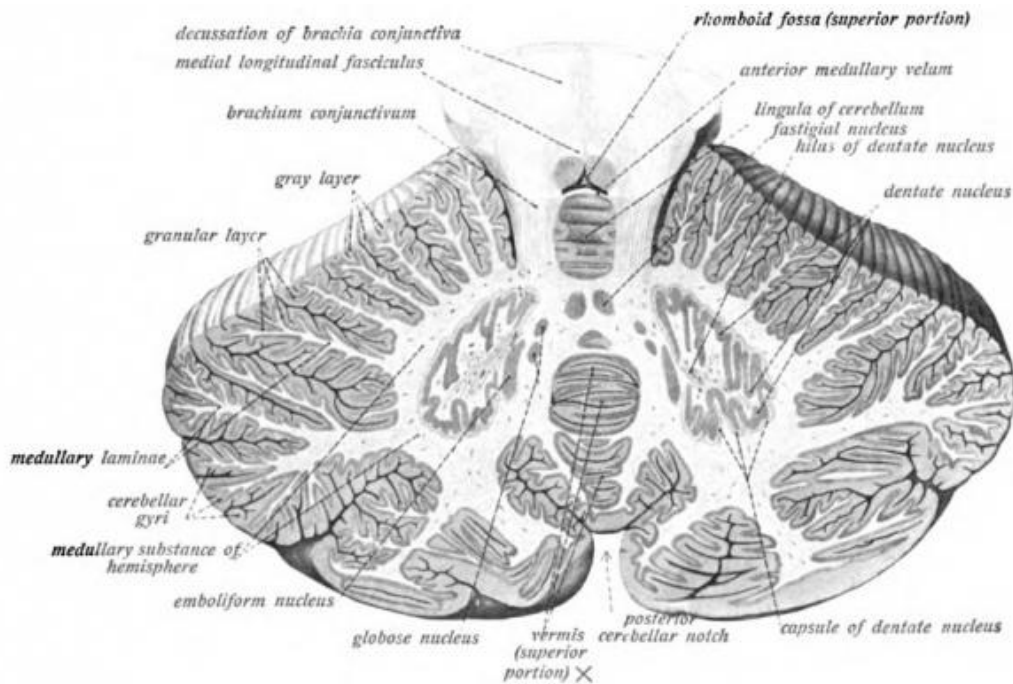


Figura 1.2.7: Ilustración de un corte del cerebelo. Fuente: Sobotta's Textbook and Atlas of Human Anatomy 1908.



### 1.3. Resonancia Magnética Nuclear

La resonancia es un fenómeno físico por el cual determinados sistemas solo admiten energía si esta se proporciona a una frecuencia específica. Un ejemplo clásico de este proceso es un columpio en el cual el balanceo de las piernas incrementa la velocidad del columpio solo si se hace con la frecuencia adecuada.

En 1946, Felix Block en Stanford y Edward Prucell en Harvard descubrieron independientemente el fenómeno de resonancia magnética nuclear que les valió a ambos el premio Nobel del física en 1952.

A grandes rasgos, el fenómeno de la RMN consiste en que determinados núcleos atómicos (como por ejemplo, el del Hidrógeno) cuando se someten a la acción de un campo magnético intenso pueden absorber energía selectivamente en forma de ondas de radio-frecuencia (RF) de una frecuencia específica (fenómeno de resonancia). Dicha energía es devuelta al retornar al estado de equilibrio, generando una señal eléctrica inducida en una antena o bobina de recepción que analizada y procesada proporciona las imágenes de RMN (IRM). En la figura 1.3.1 se muestra un esquema de dicho proceso.

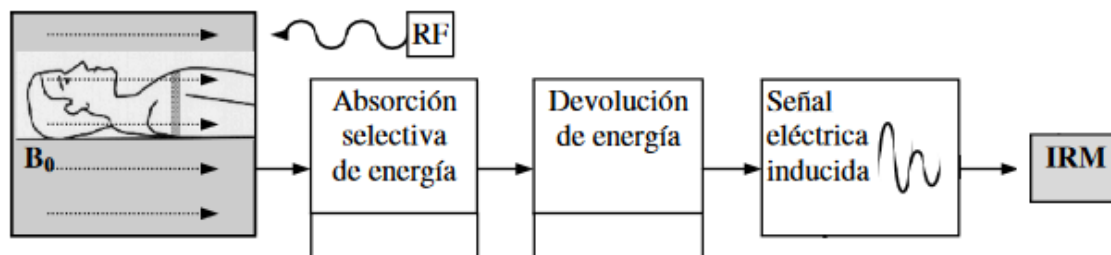


Figura 1.3.1: Esquema del proceso de obtención de imágenes de RMN.

En concreto, en IRM, se utiliza una propiedad denominada espín (rotación de protones o electrones sobre su propio eje). Esta rotación supone una carga en movimiento y por tanto producen un momento magnético en un plano perpendicular al plano de rotación (figura 1.3.2).

Cuando el número de protones presentes en el núcleo de un átomo es par la interacción entre espines hace que el espín total sea nulo y por tanto su momento magnético resultante también será nulo. Por el contrario, cuando el número de protones es impar como en el caso del hidrógeno (un protón), el núcleo posee un momento magnético neto distinto de cero.

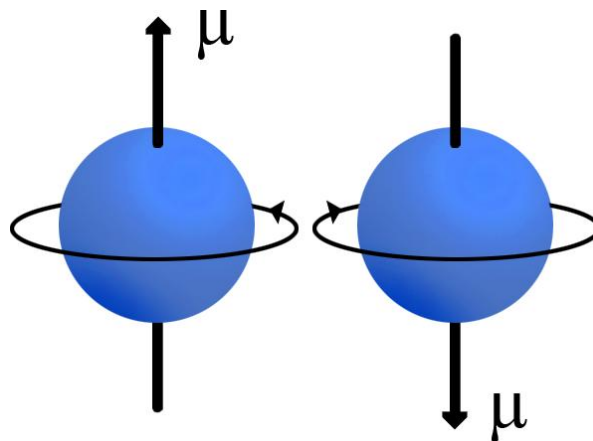


Figura 1.3.2: Ejemplo de dos protones y sus momentos magnéticos.

A nivel macroscópico el campo magnético resultante es prácticamente nulo debido a la orientación aleatoria de los momentos magnéticos. Sin embargo, en presencia de un campo magnético externo  $B$  los momentos se alinean con él resultando en un campo neto no nulo en la dirección del campo aplicado.

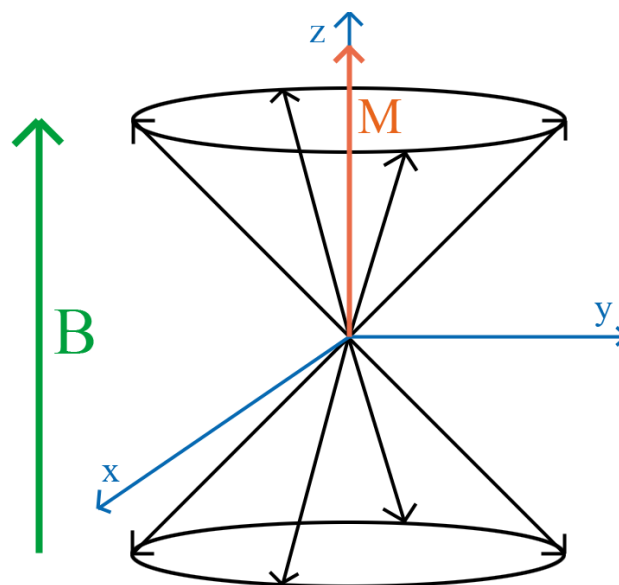


Figura 1.3.3: Vector magnetización  $M$  resultante de la suma de todos los momentos magnéticos presentes en un vóxel.

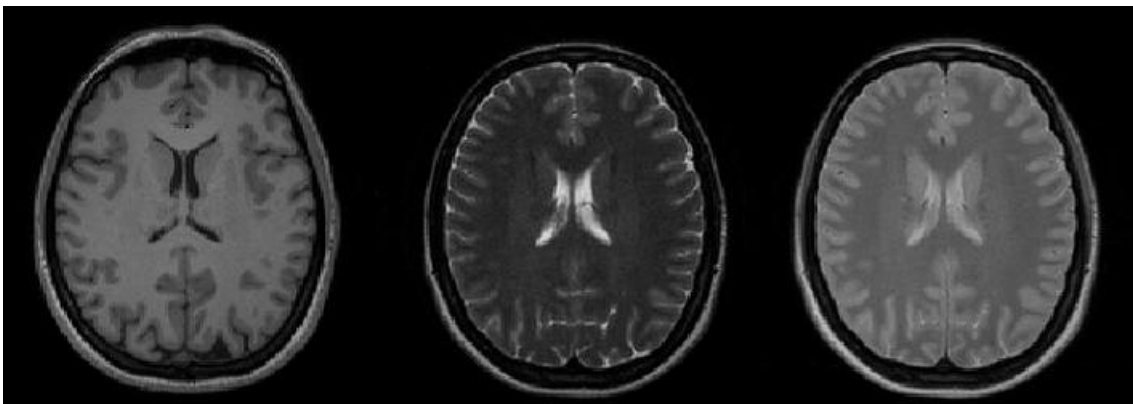
Si consideramos una pequeña porción de tejido a estudiar, el promedio de todos sus momentos magnéticos da como resultado un vector que se conoce como vector magnetización  $M$ , paralelo al campo  $B$  (figura 1.3.3) y directamente proporcional al número de átomos de hidrógeno en la región de estudio. A esta región se la llama vóxel (término que resulta de la unión de los términos en inglés VOLUME y pIXEL).

Cuando excitamos estos núcleos de H con pulsos de RF a su frecuencia natural de rotación (frecuencia de Larmor (Bargmann et al., 1959)) algunos átomos en orientación paralela (la del campo aplicado) pasarán a orientación antiparalela, ganando energía, la cual devolverán al volver a su orientación inicial emitiendo energía en forma de onda de RF. Este proceso se conoce como relajación y es dependiente de las

## Capítulo 1. Introducción y estado del arte

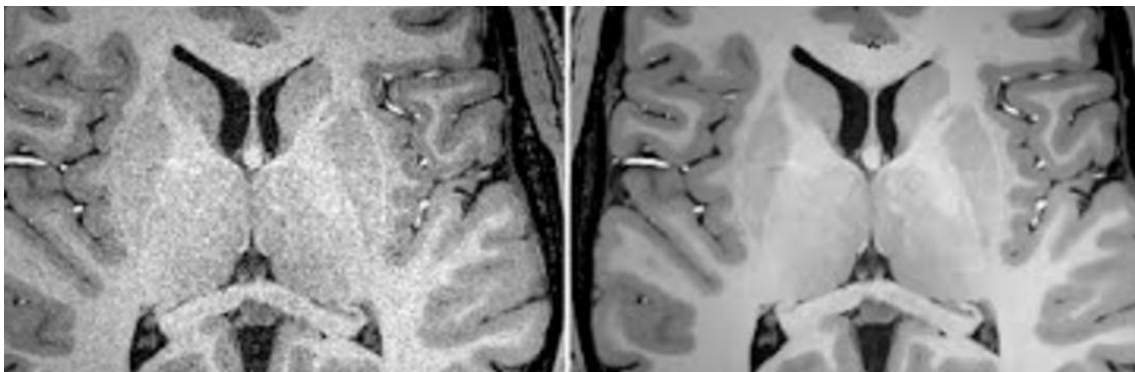
condiciones que rodean al átomo como su movilidad y la interacción entre procesos de relajación. Durante la excitación también se produce un fenómeno de coherencia de fase que hace que el vector magnetización adquiera una componente transversal no nula que es la detectada en las antenas de recepción. Dependiendo de la duración y momento en que aplicamos estos pulsos de RF potenciamos distintos fenómenos físicos en el proceso de relajación como son por ejemplo, las constantes de relajación longitudinal y transversal ( $T_1$  y  $T_2$ ) o la densidad de protones (DP).

Estos parámetros son intrínsecos de los tejidos y determinan el contraste de la imagen y para estudiarlos se excita el tejido con diferentes secuencias de pulsos de RF diseñadas para potenciar el efecto de uno de ellos. Por ejemplo, imágenes con potenciación del efecto DP contienen información relativa a la densidad de protones que hay en el tejido así como su movilidad. En la figura 1.3.4 se puede ver un ejemplo de una imagen del mismo sujeto potenciada en  $T_1$ ,  $T_2$  y DP.



*Figura 1.3.4: De izquierda a derecha: ejemplo de imágenes potenciadas en  $T_1$ ,  $T_2$  y DP. Las imágenes pertenecen al dataset IXI.*

La adquisición de imágenes de RMN no está exenta de problemas de adquisición ya que suele estar afectada por varios tipos de ruido específicos de esta técnica. Los dos tipos de ruido que más afectan al análisis cuantitativo de las imágenes son el ruido aleatorio (figura 1.3.5) y la inhomogeneidad de señal (figura 1.3.6).



*Figura 1.3.5: Ejemplo de imagen ruidosa y su versión filtrada.*

Actualmente, existen técnicas para reducir el efecto del ruido aleatorio como la aplicación de filtros adaptativos específicos (figura 1.3.5).

Otro problema propio de las imágenes de RMN es la presencia de inhomogeneidades que consiste en variaciones de intensidad de señal de baja frecuencia que hace que un mismo tejido presente patrones de brillo espacialmente variables como puede verse a la izquierda en la figura 1.3.6. Esto se debe principalmente a imperfecciones en la bobinas de excitación/detección entre otros factores (Sled et al., 1998). En la actualidad existen varios métodos automáticos que permiten reducir este ruido de forma efectiva (Sled et al., 1998).

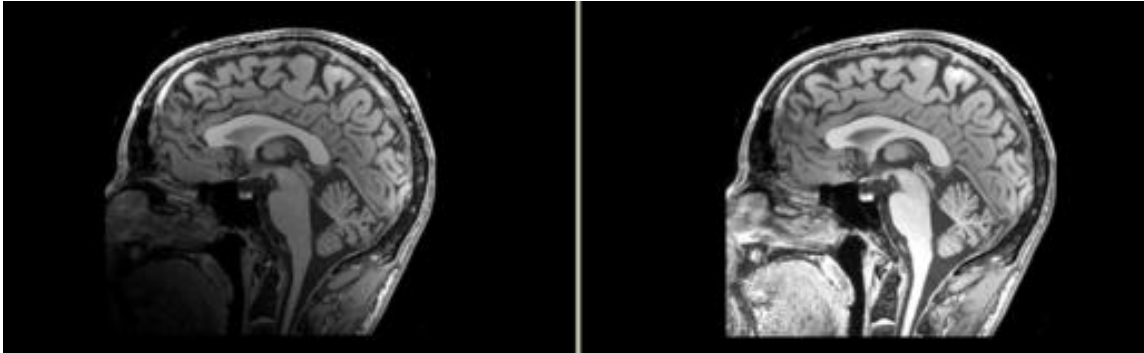


Figura 1.3.6 De izquierda a derecha. Imagen afectada por inhomogeneidades y su versión corregida respectivamente.

## 1.4. Segmentación

En esta tesis se presentan varios métodos de segmentación de imágenes de RMN cerebral por lo que repasaremos a continuación distintos aspectos del estado de la técnica.

### 1.4.1. Definición

Según Duda y Hart la segmentación es “la extracción de un objeto de interés de una imagen arbitraria” (Duda and Hart, 1973). Para Gonzalez y Woods es “la subdivisión de una imagen en sus partes constituyentes” (Gonzalez and Woods, 2007).

A menudo, interesa conocer el volumen, o en general las propiedades, de las partes que componen un órgano (en nuestro caso el cerebro). Mientras la evaluación visual de las imágenes permite hacer un análisis cualitativo, la segmentación de imágenes (figura 1.4.1) permite realizar un análisis cuantitativo contribuyendo a un diagnóstico más objetivo de diferentes patologías. Además, permite realizar estudios de normalidad y seguimiento de la evolución de pacientes. Esto, por ejemplo, es clave en el estudio de enfermedades neurodegenerativas, entre otras, que en la actualidad son uno de los grandes problemas en una sociedad cada vez más envejecida.

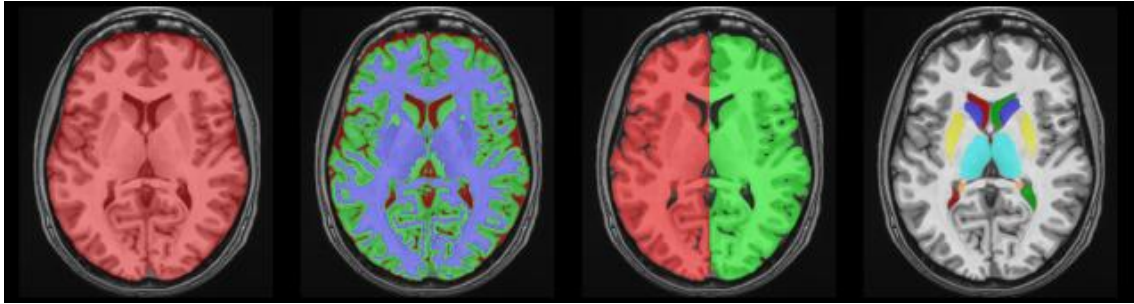


Figura 1.4.1: Ejemplos de segmentación de diferentes estructuras cerebrales. De izquierda a derecha: cavidad intracraneal, tejidos cerebrales, hemisferios cerebrales y estructuras sub-corticales.

### 1.4.2. Antecedentes históricos

La segmentación en imagen médica se puede realizar de tres formas: manual, semiautomática y automática.

#### *Segmentación manual*

La segmentación manual por parte de expertos tiene la ventaja de producir segmentaciones con bastante calidad. No obstante, es una tarea tediosa que requiere de tiempo así como de una inversión económica importante. Además, la segmentación manual depende de expertos y su nivel de entrenamiento con lo que produce una variabilidad entre expertos, es decir, es poco reproducible (Pruessner et al., 2000). Por estas razones, no es adecuada para realizar estudios longitudinales o estudios sobre la población en los que se deben procesar grandes cantidades de datos de forma consistente. Este hecho se agrava con el avance de las tecnologías de adquisición que logran imágenes de alta resolución que permiten estudiar estructuras cada vez más pequeñas y complejas lo cual ha hecho que la fiabilidad de la segmentación manual decaiga y el coste temporal se dispare.

Para resolver este problema, muchos métodos de segmentación semiautomática y automática han surgido a lo largo de los años.

#### *Segmentación semi-automática*

Una forma de reducir el coste temporal de la segmentación manual es utilizar técnicas semiautomáticas. Estas técnicas siguen necesitando la supervisión de un operador humano pero automatizan algunas partes del proceso lo que no solo reduce el tiempo de segmentación sino que generalmente produce segmentaciones más consistentes y reproducibles.

Por ejemplo, las técnicas de umbralización han sido ampliamente utilizadas para ayudar en el proceso de segmentación. La umbralización es una técnica sencilla y muy utilizada que consiste en la determinación de unos umbrales que delimitan el rango de intensidades que discriminan los distintos objetos de la imagen. La selección de estos

umbrales puede hacerse de forma manual o mediante técnicas automáticas como el método de Otsu (Otsu, 1979).

Otra forma popular de seleccionar de forma semiautomática un conjunto de píxeles es el crecimiento de regiones. El crecimiento de regiones consiste en la selección de un punto semilla a partir del cual se añaden nuevos puntos que cumplen una serie de condiciones tales como estar contacto con un punto previamente incluido y pertenecer a un determinado rango de intensidades (Adams and Bischof, 1994). Esta técnica presenta como ventaja principal respecto a la umbralización la posibilidad de seleccionar regiones con un rango similar de intensidad pero una localización y conectividad distinta. Se puede ver un ejemplo de esta técnica en la figura 1.4.2.

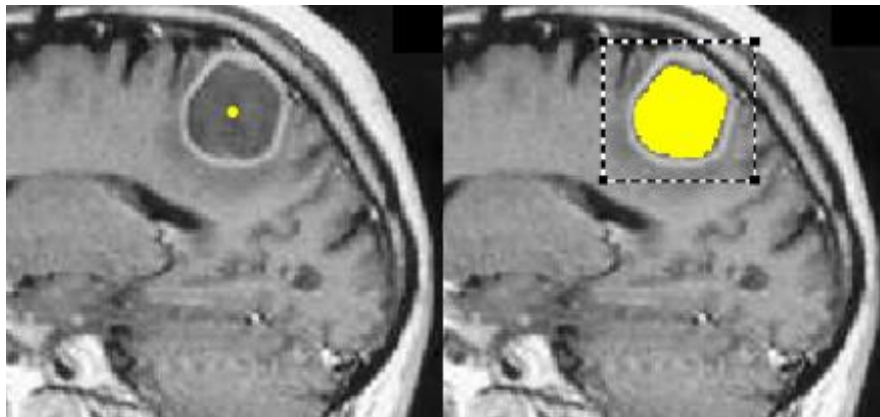


Figura 1.4.2. Ejemplo de crecimiento de regiones. A la izquierda el punto semilla seleccionado y la derecha el área seleccionada tras el proceso de crecimiento de regiones.

### Segmentación automática

Existen infinidad de métodos de segmentación automática que pueden clasificarse a grandes rasgos en dos categorías: supervisados y no supervisados. Los métodos supervisados usan pares dato-clase (ejemplos previamente etiquetados) para realizar un proceso de aprendizaje tras el cual se aplican a nuevos datos con el fin de obtener las clases a la que pertenecen. Por otro lado, los métodos no supervisados (*clustering*) usan la propia distribución de los datos para hallar la clasificación más coherente.

En segmentación de tejidos cerebrales los métodos no supervisados han sido ampliamente utilizados (probablemente debido a la falta de datos etiquetados). A continuación repasaremos algunos de los más usados.

Las técnicas de *clustering* son un conjunto de clasificadores no supervisados que se basan en la propia estructura de los datos para realizar la clasificación. Esta clasificación es un proceso iterativo que tiene por objetivo satisfacer unos criterios de homogeneidad intracase. Por tanto, las técnicas de *clustering* tienen como finalidad determinar las agrupaciones naturales de un grupo de datos sin etiquetar. Para ello utilizan criterios como la distancia entre muestras. Algunos ejemplos de técnicas de *clustering* son: el clasificador k-medias o la Mixtura de Gaussianas.

### Clasificador *k*-medias

El algoritmo *k*-medias (MacQueen, 1967) es una técnica sencilla de agrupamiento. Se trata de un algoritmo iterativo cuyo único parámetro es  $K$ , que define el número de agrupamientos o clases a encontrar. El algoritmo comienza con la definición de  $k$  centroides representativos de cada clase (medias). Este es un paso crítico ya que diferentes inicializaciones pueden conducir a diferentes resultados. Es una buena práctica definirlos lo más alejados posible entre sí. A continuación, se clasifican todos los datos con la clase del centroide más cercano. El siguiente paso es recalcular los centroides como las medias de los datos de cada agrupación. Este proceso se repite hasta que las medias no cambian o lo hacen por debajo de una determinada tolerancia.

El método minimiza una función de coste que en este caso es el error cuadrático medio:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (1.1)$$

donde  $\|x_i^j - c_j\|^2$  representa la distancia entre el dato  $x_i^j$  y el centroide  $c_j$ .

Existe una versión de este método que permite la pertenencia de cada muestra a más de una clase de forma fraccional denominado *Fuzzy C*-medias. El algoritmo *Fuzzy C*-medias (Dunn, 1973) es mucho más robusto pero más costoso computacionalmente que el *k*-medias. Se diferencia de éste en que un dato puede pertenecer a varias clases a la vez con una probabilidad determinada por las distancias a cada uno de los centroides.

La función de coste se ve modificada de la siguiente manera:

$$J = \sum_{j=1}^k \sum_{i=1}^n u_{ij}^m \|x_i^j - c_j\|^2, \quad 1 < m < \infty \quad (1.2)$$

donde  $m$  es el factor de borrosidad (*fuzzyness*) cuyo valor suele fijarse en 2 y  $u_{ij}$  es el grado de pertenencia del dato  $i$  a la clase  $j$ .

### Mixtura de Gaussianas

Cuando se conoce la distribución de los datos se pueden aplicar técnicas basadas en modelos como las mixturas de gaussianas donde se modelan los datos como la suma de un conjunto de distribuciones paramétricas, en este caso, gaussianas.

Para calcular los parámetros de las distribuciones se puede utilizar un clasificador de máxima verosimilitud. Uno de los algoritmos más populares para calcular la mixtura que mejor se ajusta a los datos es *Expectation-Maximization* o EM (Dempster et al., 1977).

Se trata de un algoritmo iterativo que consta de dos pasos. Dado un conjunto  $x$  de  $N$  datos (los cuales pueden ser mono o multiespectrales) y sus probabilidades de pertenencia a cada clase que denotamos como  $y$ . Donde  $y_{ik}$  es la probabilidad de que el dato  $x_i$  pertenezca a la clase  $k$ . La distribución de probabilidad de los datos se define como  $f(x|\Phi)$  en la que  $\Phi$  son su media, varianza y frecuencia relativa.

En primer lugar se asignan valores iniciales a los parámetros. A continuación, en el paso E, se calculan las probabilidades condicionales a posteriori a partir de estos parámetros utilizando la regla de Bayes. En el paso M se utilizan las probabilidades a posteriori para realizar una mejor estimación de los parámetros. Estos pasos se repiten para maximizar la verosimilitud.

Este algoritmo tiene un alto coste computacional lo que limita su aplicación a grandes volúmenes de datos como en el caso de imágenes de RMN. Esto se puede solucionar utilizando una reformulación (Jones and McLachlan, 1990) que usa agrupaciones de datos (histogramas) para acelerar el cómputo de los parámetros como puede verse en la siguiente formulación.

**Expectation:** Cálculo de las probabilidades a posteriori:

$$p(y_i|h_i, \Phi_y) = \frac{p(v_i|y_i, \Phi_y)p(y_i)}{\sum_{i=1}^k p(v_i|y_i, \Phi_y)p(y_i)} \quad (1.4)$$

**Maximization:** Actualización de los parámetros para mejorar la clasificación a partir de datos agrupados ( $h$ ):

$$p(y_k) = \frac{1}{N} \sum_{i=1}^B h_i \cdot p(y_i|h_i, \Phi_y) \quad (1.5)$$

$$\mu_k = \frac{\sum_{i=1}^B h_i \cdot p(y_i|h_i, \Phi_y) v_i}{\sum_{i=1}^B h_i \cdot p(y_i|h_i, \Phi_y)} \quad (1.6)$$

$$\Sigma_k = \frac{\sum_{i=1}^B h_i \cdot p(y_i|h_i, \Phi_y) \cdot (v_i - \mu_k) \cdot (v_i - \mu_k)^T}{\sum_{i=1}^B h_i \cdot p(y_i|h_i, \Phi_y)} \quad (1.7)$$

donde  $p(y_k)$  es la frecuencia relativa de aparición de los vóxeles de la clase  $k$ ,  $\mu_k$  es su media,  $\Sigma_k$  es la varianza (o matriz de covarianzas en el caso multiespectral),  $h$  representa el histograma de los datos,  $v$  son todos los posibles valores de intensidad que pueden tener los datos y  $B$  es el número máximo de valores de intensidad.

La distribución de las posibles intensidades sigue una gaussiana de la forma:

$$p(v_i|y_i, \Phi_y) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{(v_i - \mu_k)^2}{2\sigma_k^2}} \quad (1.8)$$



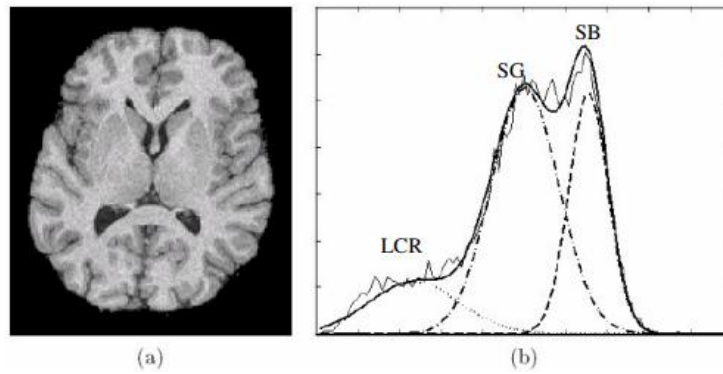


Figura 1.4.3. Ejemplo de modelo usando por mixtura de gaussianas. a) Corte axial de RMN (potenciación T1). b) Histograma que muestra el modelo ajustado usando tres gaussianas (LCR, SG y SB). Fuente: Tesis Koen Van Leemput.

En la figura 1.4.3 se puede ver un ejemplo del uso de mixtura de gaussianas para modelar el histograma de una imagen RMN cerebral potenciada en T1.

### 1.4.3. Estado del arte

En la actualidad existe una gran variedad de métodos de segmentación automática los cuales hemos clasificado en 4 grupos según la tecnología en la que se basan: *single-atlas*, atlas probabilísticos, *multi-atlas* (o MAS) y redes neuronales.

La generación de atlas (o imagen etiquetada) es la base para muchos métodos de segmentación supervisada que se apoyan en casos previamente segmentados. Estos casos se generan mediante segmentación manual o semiautomática por parte de expertos que utilizan un software de visualización, como por ejemplo ITK-SNAP (Yushkevich et al., 2006), y conocimiento anatómico para definir las áreas de interés.

Los métodos *single-atlas* se apoyan en una única imagen de referencia para realizar la segmentación. Generalmente, estos métodos abordan la segmentación como un problema de registro (Pham et al., 2000) estableciendo correspondencias entre las coordenadas del atlas y la imagen a segmentar para posteriormente propagar la segmentación del atlas al caso a segmentar.

El registro es el proceso por el cual una imagen es llevada al espacio de coordenadas de otra (figura 1.4.4). Esto puede hacerse con diferentes grados de libertad siendo el más restrictivo el registro rígido, donde solo se aplican transformaciones de traslación, y rotación (6 grados de libertad en 3D) sobre todos los vóxeles y el más libre el registro no lineal que permite deformaciones independientes (con miles de grados de libertad). Normalmente, a mayor grado de libertad mejor es el ajuste pero también mayor es el tiempo requerido para la estimación del mismo.

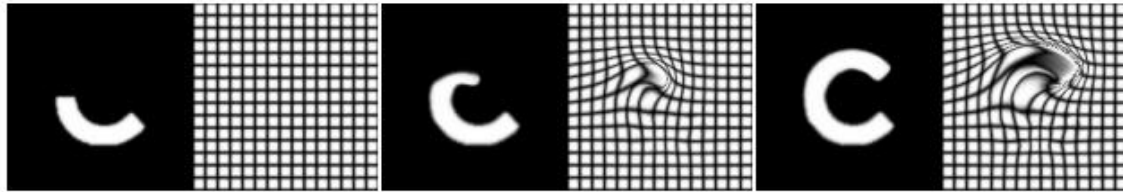


Figura 1.4.4: Ejemplo de deformación no lineal de media letra C a una letra C completa. Fuente: [Avants et al., 2009](#).

Por desgracia, el uso de un único atlas junto con una estrategia de registro no es generalmente suficiente para abarcar la variabilidad entre sujetos. Para mitigar esta limitación, algunos métodos cuentan con varios atlas de los cuales se selecciona el más apropiado con un primer registro (normalmente afín) para posteriormente aplicar un segundo registro no lineal y emplearlo en la segmentación ([Rohlfing et al., 2003](#)) o bien los integran en un único atlas promedio ([Diedrichsen, 2006](#)).

Por otro lado, los atlas probabilísticos permiten captar la variabilidad anatómica de la población en forma de probabilidades de aparición de cada tejido en el espacio. Estos atlas probabilísticos se estiman mediante el promedio de múltiples segmentaciones manuales o semiautomáticas. Un ejemplo de este tipo de atlas es el atlas MNI152 (resultado de promediar 152 sujetos).

Probablemente, la técnica más conocida que usa atlas probabilísticos es la utilizada en el famoso software SPM ([Ashburner and Friston, 2005](#)). El objetivo de SPM es unificar las dos grandes corrientes existentes en segmentación hasta el momento:

- La segmentación no supervisada basada en la clasificación de tejidos en la cual los vóxeles son asignados a una clase según su intensidad. Para ello se necesita conocer la distribución de intensidades por lo que se debe ajustar un modelo basado en mixtura de gaussianas.
- La segmentación basada en atlas donde un cerebro de referencia (plantilla) se deforma para ajustarse al cerebro a segmentar ([Collins et al., 1995](#)). Este enfoque permite incorporar información geométrica (posición) en forma de atlas probabilístico para mejorar la clasificación.

SPM utiliza un modelo generativo que intenta cubrir toda la variabilidad asociada a las imágenes del cerebro. Esto incluye no solo los campos de deformación que registran el atlas probabilístico al caso a segmentar sino también un campo de corrección de inhomogeneidad y los parámetros de cada tejido (medias, varianzas, etc). Este modelo se ajusta mediante un proceso de optimización basado en EM. Precisamente, una de las limitaciones de este método es la posibilidad de quedar atrapado en un mínimo local como ocurre con todos los métodos de optimización. La otra principal limitación de este método es que solo se aplica a la segmentación de tejidos (estructuras homogéneas a nivel de intensidad) y no a estructuras más específicas y heterogéneas.

Por otro lado, los métodos basados en MAS no integran la información de varios atlas en un único modelo sino que cada atlas es consultado y contribuye en su totalidad a la segmentación ([Heckemann et al., 2006](#); [Klein et al., 2005](#)). Esto, permite obtener

mejores resultados al representar mejor la variabilidad entre sujetos. Se basan sobretodo en dos pasos: registro y fusión de etiquetas.

Un enfoque clásico consiste en registrar cada uno de los atlas al caso a segmentar, aplicar esta deformación a las segmentaciones y combinarlas utilizando alguna estrategia como por ejemplo *majority voting* o el vecino más próximo (Artaechevarria et al., 2009). Este enfoque presenta una desventaja importante y es la necesidad de usar un registro no lineal ya que establece una correspondencia directa (a nivel de vóxel) entre los atlas y la imagen a segmentar. Para mitigar la dependencia del registro se pueden utilizar métodos de fusión de etiquetas más elaborados como *Joint label fusion* (Wang et al., 2013) o *Nonlocal label fusion* (Coupé et al., 2011). Este último incorpora dos grandes mejoras. En primer lugar, añade información estructural comparando pequeñas regiones de cada atlas y la imagen a segmentar (parches) para obtener así un peso que permite hacer un voto ponderado en función de la similitud de los patrones del caso a segmentar y los patrones de los casos de ejemplo de la librería de atlas. Por otro lado, la posibilidad de realizar una búsqueda de los parches más similares de forma no local añade un gran número de grados de libertad lo que minimiza los problemas asociados a un registro subóptimo no asumiendo por lo tanto una correspondencia directa entre atlas y caso a segmentar.

Otros métodos abordan el problema de la correspondencia entre atlas y caso a segmentar mediante estrategias de registro como por ejemplo MAGeT (Chakravarty et al., 2013; Park et al., 2014) donde se estiman una gran cantidad de registros no lineales para incrementar el número de segmentaciones candidatas. Con esto, se consiguen resultados razonablemente buenos con un reducido número de casos segmentados a mano a expensas de un alto coste computacional. La combinación de una estrategia de registro junto con *Nonlocal label fusion* da origen a métodos como RASCAL (Weier et al., 2014) donde varios atlas segmentados se registran no linealmente al caso a segmentar de forma eficiente y posteriormente se combinan mediante comparación de parches.

Un enfoque completamente diferente es el de los métodos basados en el uso de clasificadores como *random forest* (Breiman, 2001) o redes neuronales (Hertz, 1991). En el marco de esta tesis hemos utilizado redes neuronales como postproceso pero es un hecho que, aunque el estado del arte en segmentación siguen siendo los métodos basados en MAS, las redes neuronales (concretamente el uso de redes profundas, también llamado *Deep Learning*) están ganando terreno a pasos agigantados ofreciendo resultados cada vez mejores con un coste computacional muy reducido (Zhang et al., 2015).

Las redes neuronales artificiales son un modelo computacional que funciona de forma análoga al comportamiento de un sistema neuronal biológico. Están compuestas por unidades simples (neuronas artificiales) que se conectan entre sí organizándose en capas como muestra la figura 1.4.5.

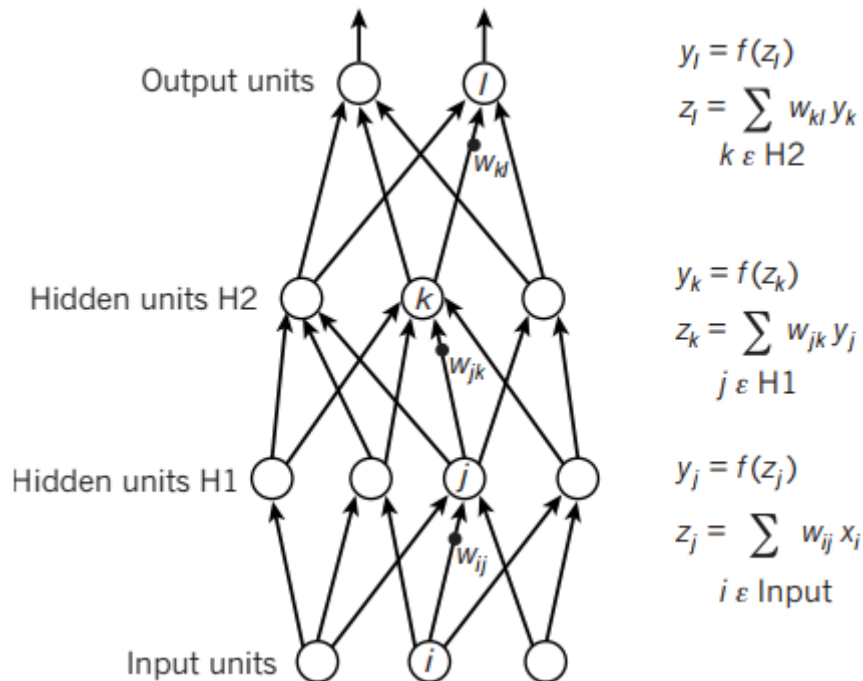


Figura 1.4.5: Ejemplo de una red neuronal sencilla con dos capas ocultas. Fuente: [LeCun et al., 2015](#).

Cada neurona actúa integrando mediante una función suma varias señales de entrada con diferentes pesos asociados y propagando una señal de salida. Puede existir una función umbral que limite la propagación de la señal.

Las redes neuronales supervisadas aprenden de forma implícita mediante ejemplos en lugar de ser programadas de forma explícita. Esto es lo que se conoce como proceso de aprendizaje o entrenamiento durante el cual se ajustan los pesos de las conexiones entre neuronas para conseguir que los datos de entrada den como resultado la respuesta correcta. Para ello, se utiliza una función de coste que calcula la diferencia entre la respuesta de la red y la respuesta correcta (error). Este error se calcula para la respuesta de cada capa empezando por la capa de salida y propagándose hacia la capa de entrada ([LeCun, 1985](#); [Parker, 1985](#); [Rumelhart et al., 1986](#); [Werbos, 1975](#)). Este método de aprendizaje permite que las redes ajustasen sus pesos de forma automática. No obstante, limita el número de capas que pueden utilizarse debido a un fenómeno llamado *gradient vanishing* que consiste en que el módulo del gradiente utilizado para modificar el valor de los pesos se hace cada vez más pequeño a medida que se propaga a través de las capas llegando a un punto en el que su efecto sobre los pesos es despreciable. Este efecto, junto con otros factores como las limitaciones en el hardware y la teoría de que el aprendizaje era prono a converger en mínimos locales hicieron que a finales de los 90 las redes neuronales perdieran fuerza en el campo de la visión por computador.

En 2006, las redes neuronales volvieron a despertar interés gracias al desarrollo de nuevas formas de entrenar la redes neuronales ([Bengio et al., 2007](#); [Hinton et al., 2006](#); [Ranzato et al., 2006](#)). Sin embargo, el hecho que llamo la atención de toda la comunidad científica y que relanzo definitivamente el interés por la redes neuronales

fue la victoria de Alex Krizhevsky en 2012 en la competición ImageNet utilizando una red neuronal profunda (Krizhevsky et al., 2012).

El Dr. Krizhevsky no solo gano la competición sino que lo hizo de forma aplastante con un error un 10% menor que sus competidores. El Dr. Krizhevsky utilizo una versión modificada de una red convolucional llamada LeNet (LeCun, 1985) en combinación con el uso de funciones de activación de tipo ReLU (que minimizan el problema del *gradient vanishing*). Para entrenar esta red utilizo la potencia de computación de las tarjetas gráficas de la época (GPUs) lo que le permitió factores de aceleración muy importantes respecto al entrenamiento en CPU. Actualmente, la inmensa mayoría de la investigación en *Deep learning* se realiza mediante el uso de modernas GPUs con una gran potencia de cómputo.

En la actualidad, una gran variedad de métodos basados en redes neuronales de diversas topologías copan el campo de la imagen médica rivalizando con los consolidados métodos MAS y comienzan a publicarse versiones híbridas que buscan unir las ventajas de ambas tecnologías (Sanroma et al., 2018).

#### 1.4.4. El método Nonlocal label fusion

Como hemos comentado antes, una de las alternativas más utilizadas a la hora de desarrollar un método de segmentación automática es la basada en la tecnología MAS. Estas técnicas, se apoyan en una librería de imágenes previamente segmentadas a mano cuyas segmentaciones se trasladan al caso a segmentar mediante un algoritmo. Uno de estos algoritmos es el presentado por (Coupé et al., 2011) llamado *Nonlocal label fusion*. Este, a su vez, se basa en un método de filtrado de imágenes denominado *Non-local means* propuesto por Antoni Buades (Buades et al., 2005). A continuación, describiremos el método Non-local means ya que es la base de todos los métodos non-locales que veremos en esta tesis.

En el filtrado el objetivo es recuperar una imagen real a partir de una adquisición ruidosa:

$$v(i) = u(i) + n(i) \quad (1.9)$$

donde  $v(i)$  es el valor observado,  $u(i)$  es el valor real y  $n(i)$  representa al ruido para cada pixel  $i$  de la imagen.

En el método de Buades, dada una imagen ruidosa  $v = \{v(i) \mid i \in I\}$  el valor estimado  $NL[v](i)$  se computa como la media ponderada del resto de píxeles de la imagen:

$$NL[v](i) = \sum_{j \in I} w(i, j) v(j) \quad (1.10)$$

donde el peso, representado por  $w(i, j)$  dependen de la similitud entre los píxeles  $i$  y  $j$ , la cual, se computa en base a la distancia Euclídea entre los vectores  $v(N_i)$  y  $v(N_j)$ , donde  $N_k$  representa un parche o región de tamaño fijo centrado en el pixel  $k$  con lo que los

pixeles con vecindarios similares tienen mayor peso en el promedio. Estos pesos se calculan de la siguiente forma:

$$w(i, j) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_2^2}{h^2}} \quad (1.11)$$

donde  $Z(i)$  es una constante de normalización:

$$Z(i) = \sum_{\forall j} e^{-\frac{\|v(N_i) - v(N_j)\|_2^2}{h^2}} \quad (1.12)$$

y el parámetro  $h$  controla el grado de filtrado que aplicamos a la imagen y está relacionado con el nivel de ruido de la imagen. En la figura 1.4.6 se muestra un ejemplo del proceso de comparación de parches que permite realizar la media ponderada de los valores de la imagen. A grandes rasgos, el método *Non-local means* nos dice que para reducir el ruido de una imagen de forma efectiva solo es necesario promediar cada píxel con píxeles con valores similares. El filtro *Non-local means* se basa por lo tanto en la redundancia natural de patrones en la imagen y en el hecho de que esta redundancia no tiene por qué ser necesariamente local. Por otro lado, el hecho de comparar regiones en lugar de píxeles da robustez estadística al cálculo de las similitudes entre regiones.

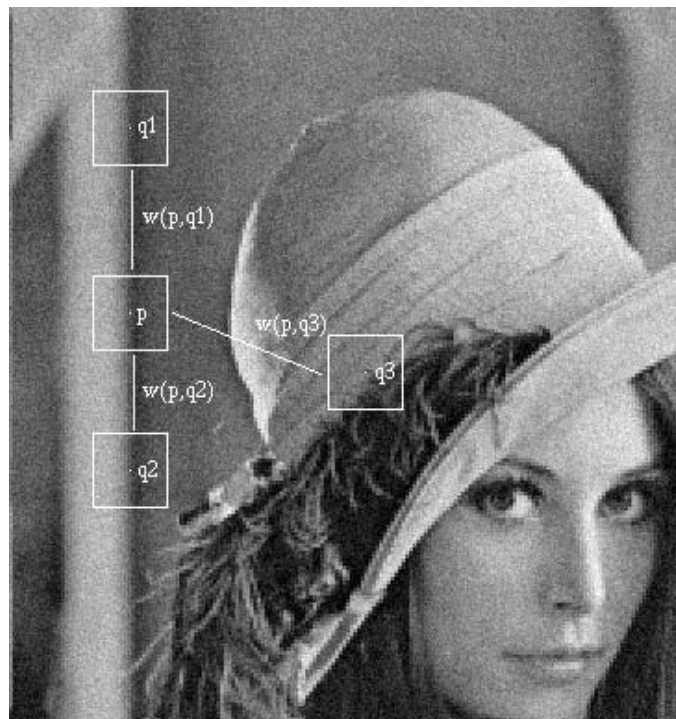


Figura 1.4.6: Ejemplo de esquema del proceso de filtrado mediante *Non-local means*. El parche alrededor del punto  $p$  se compara con otros parches de la imagen, algunos son similares ( $q1$  y  $q2$ ) y otros no ( $q3$ ). Fuente: [Buades et al., 2005](#).

El filtro *Non-local means* ha ganado una gran popularidad en la última década y se ha utilizado en la resolución de múltiples problemas del procesamiento digital de imagen (filtrado, superresolución, segmentación, clasificación, etc.).

En (Coupé et al., 2011) se aplicó la idea del filtro *Non-local means* a la segmentación del hipocampo y los ventrículos laterales en imágenes de RMN tridimensionales (3D). En este caso, para asignar una clase a un vóxel determinado se toma su parche correspondiente y se compara con parches de imágenes de una librería de imágenes segmentadas (limitando la búsqueda a una región definida). La similitud de dichos parches se utiliza para calcular un peso que tendrá la clase acompañada a dicho parche. Finalmente se asigna la clase con mayor peso.

Para un vóxel  $x_i$  de una imagen a segmentar el método estima su etiqueta final  $L(x_i)$  asignando pesos a las etiquetas asociadas a los parches que rodean al vóxel a segmentar dentro de la región de búsqueda  $V_i$  a lo largo de los  $N$  sujetos de la librería. Para cada parche se computa un peso  $v(x_i)$  que representa un voto ponderado en función de la similitud de patrones de todos los parches de la librería consultados:

$$v(x_i) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) y_{s,j}}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (1.13)$$

donde  $y_{s,j}$  es una etiqueta candidata asociada al vóxel  $x_{s,j}$  en la posición  $j$  del sujeto  $s$  y  $w(x_i, x_{s,j})$  es el peso calculado como la similitud entre los parches  $x_i$  y  $x_{s,j}$ :

$$w(x_i, x_{s,j}) = \begin{cases} \exp \frac{-\|P(x_i) - P(x_{s,j})\|_2^2}{h^2}, & ss < th \\ 0, & otherwise \end{cases} \quad (1.14)$$

donde  $P(x_i)$  representa el parche que rodea al vóxel  $x_i$ ,  $P(x_{s,j})$  el parche que rodea al vóxel  $x_j$  en el sujeto  $s$  de la librería y  $\|\cdot\|_2$  es norma L2 normalizada por el número de elementos. La diferencia entre ambos parches consiste en la distancia (en intensidad) entre cada par de vóxeles de ambos parches.

El parámetro  $h$  se estima como la mínima distancia entre todos los parches dentro del área de búsqueda:

$$h(x_i) = \operatorname{argmin}_{x_{s,j}} \|P(x_i) - P(x_{s,j})\|_2 + \varepsilon \quad (1.15)$$

Donde  $\varepsilon$  es una pequeña constante para evitar divisiones por cero.

Para ahorrar tiempo de cómputo se realiza una preselección, solo se comparan parches cuyo *structural similarity index* (Wang et al., 2004) o  $ss$  supera un determinado umbral (en este caso fijado a  $th = 0.95$ ). El valor de  $ss$  representa una medida previa de similitud entre parches que se calcula de la siguiente manera:

$$ss = \frac{2\mu_i\mu_{s,j}}{\mu_i^2 + \mu_{s,j}^2} \cdot \frac{2\sigma_i\sigma_{s,j}}{\sigma_i^2 + \sigma_{s,j}^2} \quad (1.16)$$

donde  $\mu_i$  y  $\sigma_i$  se corresponden con la media y la desviación estándar del parche  $P(x_i)$  que rodea el vóxel  $x_i$ .

Finalmente la etiqueta para un vóxel se computa de la siguiente forma:

$$L(x_i) = \begin{cases} 1 & v(x_i) \geq 0.5 \\ 0 & v(x_i) < 0.5 \end{cases} \quad (1.17)$$

En la figura 1.4.7 se muestra de forma esquemática todo el proceso.

El método propuesto por Coupé fue evaluado comparándolo con otros dos enfoques del estado del arte, segmentación basada en single atlas (seleccionando el mejor de entre toda la librería o *Best template*) (Collins and Pruessner, 2010) y segmentación *Appearance-based* (Duchesne et al., 2002).

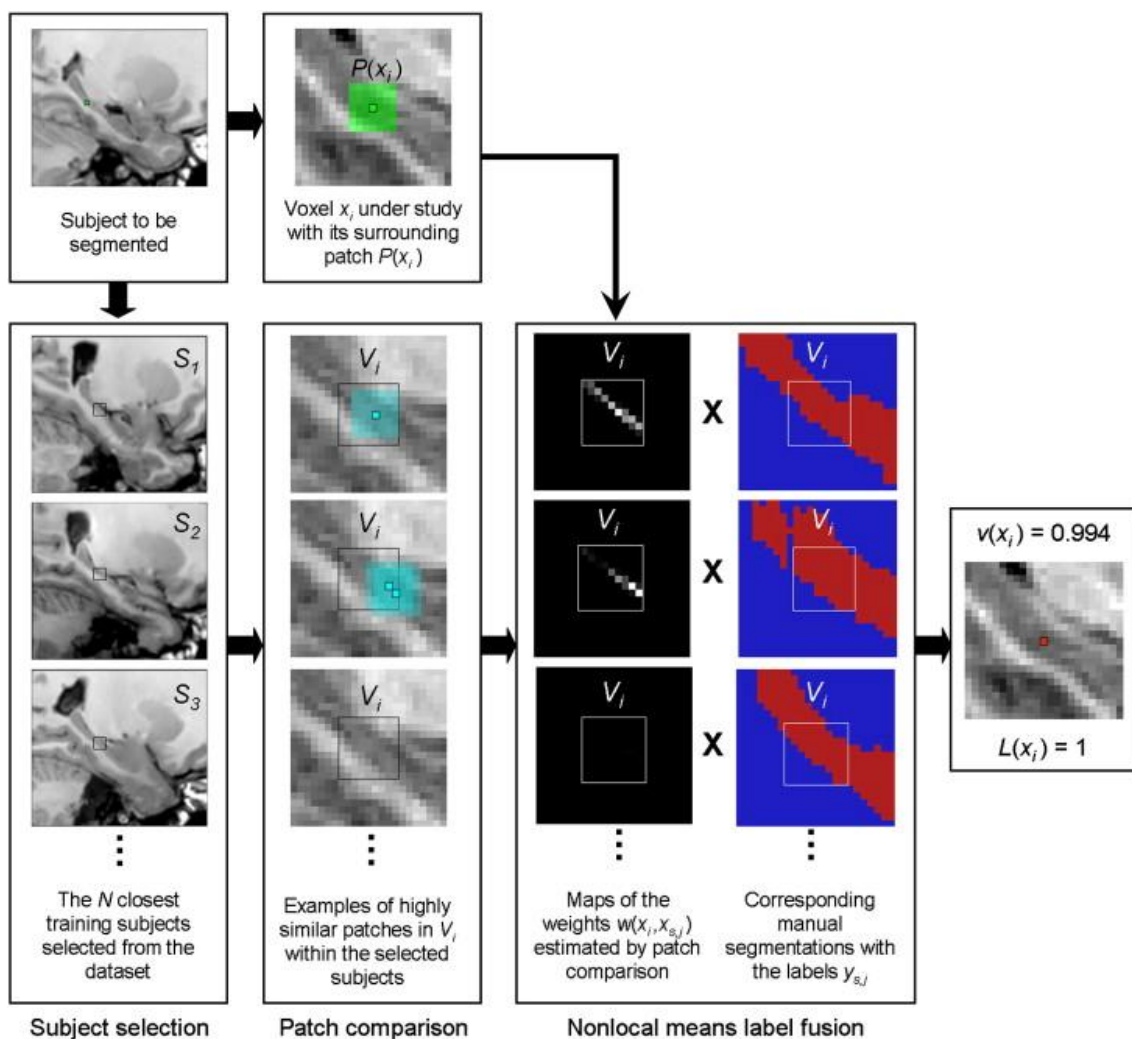


Figura 1.4.7: Esquema del proceso de segmentación mediante Nonlocal label fusión. Nótese que el ejemplo se muestra en 2D por simplicidad pero el proceso real se realiza utilizando parches 3D. Cada parche de la imagen a segmentar es comparado con miles de parches de las imágenes de la librería y sus similitudes se utilizan para ponderar los votos de cada parche. Fuente: Coupé et al., 2011.



Para evaluar la precisión de los diferentes enfoques se calculó el índice Kappa (Zijdenbos et al., 1994) también llamado índice DICE (ecuación 1.18). Se evaluó la calidad de los resultados aplicando el método a la segmentación del hipocampo sobre sujetos sanos y a la segmentación de los ventrículos laterales sobre sujetos con Alzheimer.

$$Kappa(A, B) = \frac{2|A \cap B|}{|A \cup B|} \quad (1.18)$$

Donde  $A$  y  $B$  representan el conjunto de vóxeles de la segmentación manual y automática respectivamente.

Para la segmentación del hipocampo, el enfoque *Appearance-based* obtuvo un Kappa de 0,800, el enfoque *Best template* un Kappa de 0,837 y el método *Nonlocal label fusion* (utilizando 20 casos segmentados a mano como librería) obtuvo un índice Kappa de 0,882. Para la segmentación de los ventrículos laterales, el enfoque *Appearance-based* obtuvo un Kappa de 0,788, el enfoque *Best template* obtuvo un Kappa de 0,909 y el método *Nonlocal label fusion* un Kappa de 0,957 (figura 1.4.8). El método *Nonlocal label fusión* obtuvo una mejora estadísticamente significativa de acuerdo con el resultado del test de Kruskal-Wallis con una  $p < 0,001$  para ambas segmentaciones.

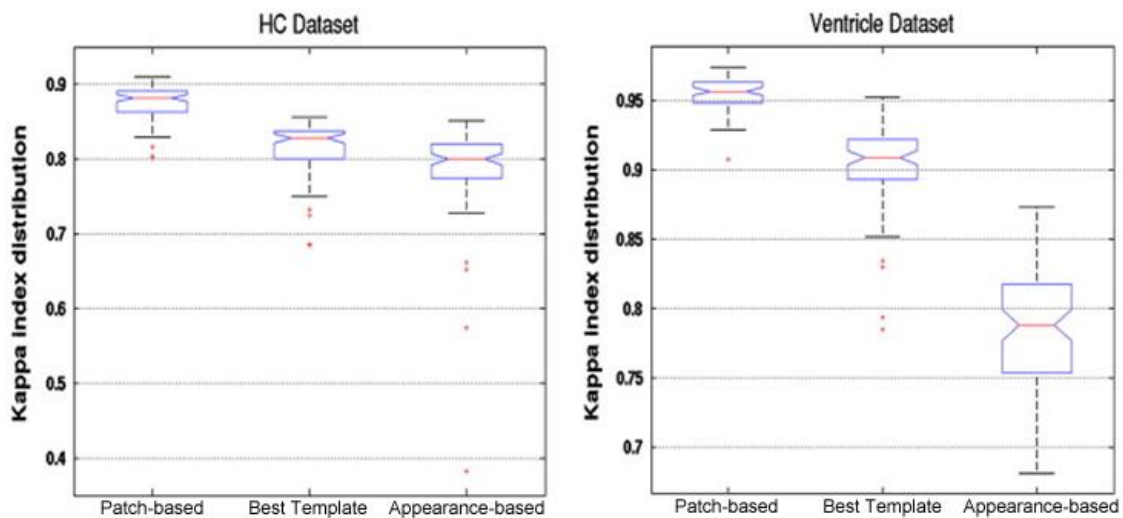


Figura 1.4.8: Distribución del índice Kappa obtenido por los tres métodos para ambas segmentaciones hipocampo y ventrículos laterales.

En resumen, el método *Nonlocal label fusion* tiene un alto rendimiento el cual ha sido validado para la segmentación del hipocampo y los ventrículos laterales comparándolo con otros dos enfoques del estado del arte. Sin embargo, el principal problema del enfoque *nonlocal* es su alto coste computacional, el cual esta principalmente asociado al proceso de comparación de parches. En los capítulos 3, 4 y 5 se describirán tres nuevos métodos de segmentación que se basan en modificaciones de este algoritmo, su combinación con técnicas de preprocesado como registro no-lineal y técnicas de postprocesado. Todas estas modificaciones han tenido como objetivo principal reducir el coste computacional de esta metodología y mejorar la precisión de la misma.

## Capítulo 2: Hipótesis y aportaciones

### 2.1. Hipótesis

Esta tesis se apoya en cuatro hipótesis que consisten en la mejora de la eficiencia y la calidad de los métodos de segmentación *multi-atlas* (MAS) basados en parches:

1. En primer lugar, proponemos mejorar sustancialmente la eficiencia temporal de los métodos de segmentación basados en parches utilizando el conocimiento anatómico del cerebro para evitar procesar zonas que siempre corresponden al mismo tejido entre sujetos.
2. En segundo lugar proponemos una mejora en la calidad de la segmentación en los métodos que utilizan enfoques multiescala incorporando una combinación de escalas dependiente de cada estructura (o etiqueta).
3. También proponemos mejorar la calidad de las segmentaciones producidas utilizando una nueva métrica multiespectral para calcular la similitud de los parches utilizando varios canales.
4. Por último proponemos añadir un post proceso basado en redes neuronales que aprenda y elimine los errores sistemáticos cometidos por el método de segmentación.

### 2.2. Aportaciones

Se han desarrollado tres métodos que permiten segmentar distintas regiones/estructuras cerebrales. En cada una de ellas se han aplicado las hipótesis de esta tesis a fin de mejorar tanto la calidad como la eficiencia de los métodos de segmentación:

1. NABS: Desarrollo de un algoritmo de segmentación basado en *Nonlocal label fusion* que divide el cerebro en 5 regiones (hemisferio cerebral derecho, hemisferio cerebral izquierdo, hemisferio cerebelar izquierdo, hemisferio cerebelar derecho y bulbo raquídeo) aplicando información a priori para obtener una mejora sustancial en la eficiencia.
2. CERES: Desarrollo de un algoritmo para la segmentación de los lobulillos cerebelares adaptando el método OPAL e introduciendo un coeficiente de mezcla de escalas dependiente de las estructuras para obtener un aumento sustancial de la calidad de las segmentaciones.
3. HIPS: Desarrollo de un método de segmentación de las subregiones del hipocampo utilizando dos de los protocolos de segmentación públicamente disponibles usando imágenes de alta resolución y adaptándolo a su uso en imágenes de resolución estándar mediante métodos de superresolución. En este método, incorporamos una nueva métrica de multi-contraste y un postproceso basado en redes neuronales para la corrección de errores sistemáticos con el que se aumenta sustancialmente la calidad de las segmentaciones.

### **2.3. Organización**

En el capítulo 3 se presenta el método NABS que segmenta el cerebro en 5 regiones utilizando una optimización para reducir el coste temporal asociado a los métodos basados en parche.

En el capítulo 4 se presenta el método CERES que segmenta los lobulillos del cerebelo introduciendo un enfoque multiescala adaptativo dependiente de las etiquetas que incrementa la calidad de los resultados.

En el capítulo 5 se presenta el método HIPS que segmenta el hipocampo en subregiones utilizando dos protocolos de segmentación distintos. en este método se propone una nueva métrica multi-contraste y el uso de redes neuronales para la corrección de errores sistemáticos.

Finalmente, en el capítulo 6 se discuten las conclusiones generales de la tesis así como las líneas de trabajo futuras.

## Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

En este capítulo se presenta un nuevo método para la segmentación del encéfalo en 5 subregiones usando imágenes de Resonancia Magnética Nuclear. Este método fue publicado en 2015 (Romero et al., 2015) en la revista Magnetic Resonance in Medicine y una versión adaptada del mismo se utiliza actualmente dentro del sistema online volBrain (<http://volbrain.upv.es>).

### 3.1. Segmentación del cerebro

El encéfalo es un órgano que regula una gran cantidad de procesos de nuestro organismo por no mencionar que es donde tiene origen nuestra consciencia. Se encuentra alojado en el interior del cráneo flotando en un fluido denominado líquido cefalorraquídeo o LCR (también CSF por sus siglas en inglés, *CerebroSpinal Fluid*).

La subdivisión del volumen encefálico en hemisferios derecho e izquierdo, cerebelo y bulbo raquídeo es una aplicación de gran interés dada su utilidad en aplicaciones como la planificación de intervenciones o la estimación de la asimetría cerebral/cerebelar (Zhao et al., 2010). En la figura 3.1.1, se puede observar el etiquetado de una imagen de RMN cerebral en estas 5 regiones anatómicas.

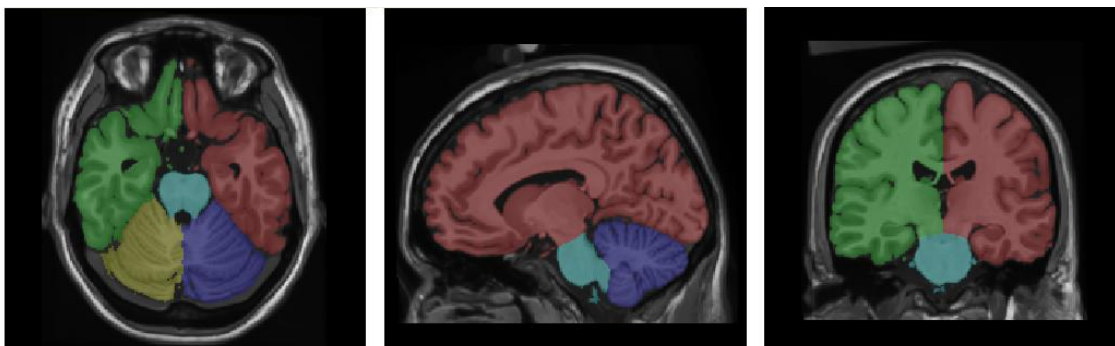


Figura 3.1.1: Ejemplo de etiquetado manual de un cerebro en 5 subregiones (hemisferio cerebral izquierdo e derecho, hemisferio cerebelar izquierdo e derecho y bulbo raquídeo).

Una aproximación clásica para separar los hemisferios cerebrales/cerebelares es el uso del plano sagital medio que se define como el plano que mejor se ajusta a la fisura que separa ambos hemisferios del cerebro (Brummer, 1991) o el plano que maximiza la simetría bilateral (Liu et al., 2001). Estos métodos se basan en la hipótesis de que el cerebro es "bastante" simétrico pero en la práctica esto no siempre es así por lo que la potencia de estos métodos es muy limitada.

Los métodos de segmentación basados en *compartimentos* dividen el encéfalo en sus dos hemisferios y el bulbo raquídeo. Podemos encontrar varios métodos en la literatura basados en reconstrucción que establecen un vóxel inicial dentro del

compartimento a segmentar al que llaman semilla y a partir del cual hacen crecer el área de selección hasta que se alcanza la frontera con otra estructura como por ejemplo la frontera entre sustancia gris y blanca. Los métodos que usan este enfoque suelen diferir en el método para establecer las semillas como FreeSurfer (Dale et al., 1999) o BrainVoyager (Kriegeskorte and Goebel, 2001) en sus primeras versiones.

Otro tipo de métodos son los basados en morfología matemática como BrainVisa (Mangin et al., 2004) que se basan en la geometría de “cuello de botella” que presentan determinadas estructuras. Este método, aplica erosiones sucesivas sobre una máscara inicial para desconectar las diferentes partes del cerebro y obtener así la segmentación. Dicha segmentación se puede refinar con el uso de superficies deformables ajustándolas a los límites de la estructura de interés mediante un criterio basado en intensidades (Liang et al., 2007).

Una desventaja de estos métodos es su dificultad para identificar con precisión la frontera de los compartimentos que definen debido a que son sensibles al efecto de volumen parcial. Este efecto se deriva de la propia adquisición ya que la intensidad de un vóxel en IRM es el promedio de la información de los átomos contenidos en él, pudiendo contener mezcla de tejidos y por tanto pertenecer a varias estructuras (Manjón et al., 2010c). Además, estos métodos, suelen fallar a la hora de separar el bulbo raquídeo del cerebelo debido a la complejidad de esta unión.

Para mitigar este problema, el método basado en cuellos de botella se ha combinado con modelos basados en ecuaciones diferenciales parciales (Zhao et al., 2010) que simulan una transferencia de información entre los dos compartimentos separados por un cuello de botella (sin embargo este método tiene un elevado coste computacional).

En este capítulo, se propone un nuevo método preciso y eficiente para segmentar el cerebro y el cerebelo en sus hemisferios derecho e izquierdo y bulbo raquídeo usando una adaptación de la tecnología *Multi-Atlas Label Fusion* (MALF).

## 3.2. Materiales y métodos

El método que proponemos utiliza una librería de casos segmentados manualmente como base de conocimiento. Esta librería se ha construido utilizando imágenes de RMN de acceso público procedentes del *IXI dataset* (<http://www.brain-development.org/>) que contiene imágenes de casi 600 sujetos sanos procedentes de diferentes hospitales de Londres.

### 3.2.1. Creación de la librería

Para generar la librería se seleccionaron al azar 30 imágenes de entre las casi 600 antes mencionadas. Se trata de imágenes potenciadas en T1 adquiridas en máquinas de 1.5 T y 3 T de diferentes fabricantes. Las imágenes tienen una resolución de  $1 \times 1 \times 1.2 \text{ mm}^3$  y un tamaño de  $256 \times 256 \times 130$  vóxeles. El rango de edades de los

### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

sujetos sanos seleccionados es de 24 a 75 años. Todas las imágenes fueron preprocesadas como se describe a continuación

**Filtrado:** Para reducir el ruido y mejorar la calidad de las imágenes se utilizó el filtro SANLM (*Spatially Adaptive Non-Local Means*) (Manjón et al., 2010b) que puede trabajar con niveles de ruido variables.

**Registro:** Para realizar la segmentación se necesita situar las imágenes en el mismo espacio de coordenadas. Para esto aplicamos una normalización espacial al espacio definido por el Montreal Neurological Institute (MNI152) utilizando SPM8 (Penny et al., 2007). Tras esto, las imágenes tienen un tamaño de 181x217x181 vóxeles y una resolución de 1 mm<sup>3</sup>.

**Extracción intracraneal y corrección de inhomogeneidad:** Dado que solo vamos a segmentar el tejido cerebral eliminamos todos los vóxeles que pertenecen a otros tejidos como pueda ser el cráneo. Para ello utilizamos VBM8 (<http://dbm.neuro.uni-jena.de/>). El mismo software se utilizó para corregir la inhomogeneidad.

**Normalización en intensidad:** El método de segmentación propuesto estima la similitud entre imágenes utilizando medidas basadas en intensidad por lo que todas las imágenes fueron normalizadas en intensidad para que los diferentes tejidos (SB, SG y LCR) tengan valores de intensidad consistentes entre imágenes. Para esto se utilizó el método TMS (Manjón et al., 2010c) para estimar el valor medio de los tejidos y después se aplicó un mapeo lineal (Lötjönen et al., 2010) de tal modo que la SB tuviera un valor medio de 250, la SG un valor medio de 150 y el LCR un valor medio de 50 (ver figura 3.2.1).

**Segmentación manual:** Se realizó la segmentación manual de las imágenes en cinco clases (hemisferios derecho e izquierdo, cerebelo y bulbo raquídeo) por un experto en anatomía cerebral utilizando el software ITK-SNAP (Yushkevich et al., 2006). Para acelerar el proceso se obtuvo una segmentación inicial utilizando el método ADisc (Zhao et al., 2010) que fue corregida manualmente (la figura 3.2.2 muestra un resumen de todo el preproceso).

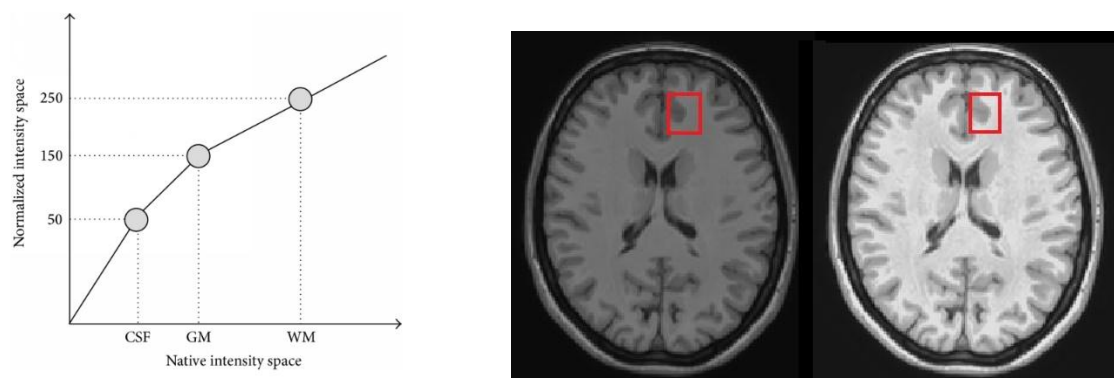


Figura 3.2.1: Ejemplo de una imagen antes y después de aplicar una normalización de intensidad respectivamente. Se puede observar la diferencia en el contraste entre diferentes estructuras.

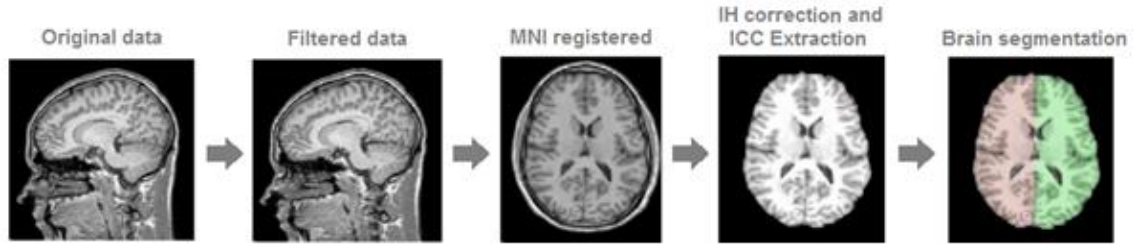


Figura 3.2.2: Pasos para la construcción de la librería.

### 3.2.2. Block-wise Nonlocal label fusion aplicado a múltiples clases

Para poder aplicar el método *Nonlocal label fusion* a un caso con múltiples etiquetas (o clases) es necesario modificar la ecuación 1.13:

$$v(x_i, k) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) \delta(k, y_{s,j})}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (3.1)$$

donde  $\delta$  es la función delta de Kronecker y  $k = \{0, 1, \dots, m\}$  representando el conjunto de etiquetas que vamos a considerar.

Finalmente la etiqueta para un vóxel se calcula como:

$$L(x_i) = \arg \max_k v(x_i, k) \quad (3.2)$$

La carga computacional de este proceso puede reducirse considerablemente si usamos un enfoque orientado a bloque en lugar de un enfoque orientado a vóxel. Es decir, en lugar de computar la etiqueta de un solo vóxel, podemos utilizar el peso calculado para asignarlo a los vóxeles de todo el parche. En (Rousseau et al., 2011) se aplicó este enfoque a la segmentación ya propuesto por (Coupe et al., 2008) para el filtrado obteniendo unos resultados ligeramente superiores a los obtenidos usando un enfoque basado en vóxel. Para obtener reducción del tiempo de computo, la implementación por bloques en lugar de procesar todos los vóxeles del volumen de estudio solo procesa vóxeles en posiciones impares (esto reduce en teoría el tiempo de computo aproximadamente en un factor 8 en 3D).

De hecho, el enfoque basado en bloque no solo reduce el tiempo de cómputo sino que también introduce una regularización implícita debida al solapamiento entre los parches que resulta en un mayor número de parches contribuyendo a los votos por cada vóxel.

Por tanto el cálculo del voto para un bloque se define así:

$$v(B(x_i), k) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) \delta(k, B(y_{s,j}))}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (3.3)$$

donde  $B(x_i)$  representa un bloque de  $N \times N \times N$  vóxeles que van a ser etiquetados simultáneamente. En la fórmula 3.3 hemos abusado de la notación siendo que  $\delta(k, B(y_{s,i}))$  devuelve un bloque de  $N \times N \times N$  valores donde cada valor es comparado con  $k$  tratando cada vóxel por separado y  $v(B(x_i), k)$  es un bloque de  $N \times N \times N$  votos.

Finalmente, el recuento de votos  $v(x_i, k)$  para cada vóxel se obtiene sumando los votos de cada bloque que contiene el vóxel  $x_i$ :

$$v(x_i, k) = \sum_{x_i \in B(z)} [v(B(z), k)]_i \quad (3.4)$$

donde  $[v(B(z), k)]_i$  hace referencia al elemento del bloque  $v(B(z), k)$  correspondiente al vóxel  $x_i$ . La etiqueta final para cada vóxel se computa como muestra la ecuación 3.2.

### 3.2.3. Reducción del área de proceso usando información a priori

Incluso utilizando el enfoque por bloques el coste computacional del método propuesto sigue siendo alto debido a que se tienen que clasificar todos los vóxeles del cerebro. No obstante, podemos aprovechar el hecho de que todas las imágenes se encuentran en el mismo espacio de coordenadas (MNI152). Esto implica que los vóxeles que se encuentran alejados de las fronteras entre clases pueden ser fácilmente clasificados con la etiqueta más frecuente de los casos de la librería para las coordenadas de ese vóxel. De esta forma solo habría que aplicar el método a los vóxeles cercanos las fronteras entre clases.

Para conseguir esto generamos una máscara a partir de los 30 casos de nuestra librería (figura 3.2.3). Dilatando cada una de las etiquetas y calculando su intersección obtenemos lo que llamamos área de incerteza la cual contiene los vóxeles que serán visitados. El resto de vóxeles se etiquetarán con el valor de la moda obtenida de la distribución de etiquetas de los casos de la librería.

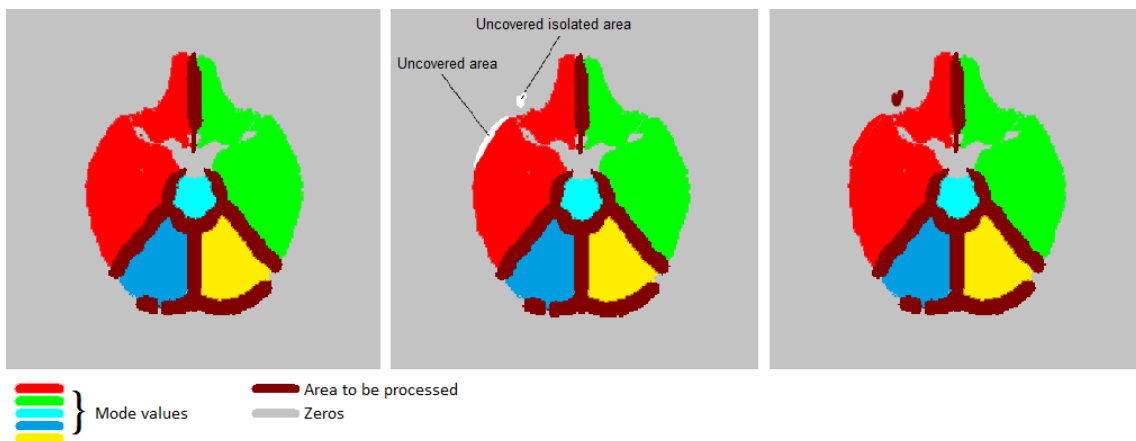


Figura 3.2.3: De izquierda a derecha: Máscara obtenida a partir de la librería, máscara sobre un caso nuevo a segmentar (nótese las aéreas sin cubrir) y máscara adaptada al caso a segmentar.



### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

Al aplicarse esta máscara a un nuevo caso a segmentar esta puede no cubrir todos los vóxeles a clasificar. Por esto la máscara debe adaptarse a cada caso. Para ello distinguimos dos casos: los vóxeles que están en contacto con un área etiquetada se clasifican con la clase de esa misma área lo cual se consigue mediante un crecimiento de regiones. Por otro lado, los vóxeles aislados se añaden al área de incerteza para que sean visitados. La figura 3.2.4 muestra una visión general del preproceso y la segmentación.

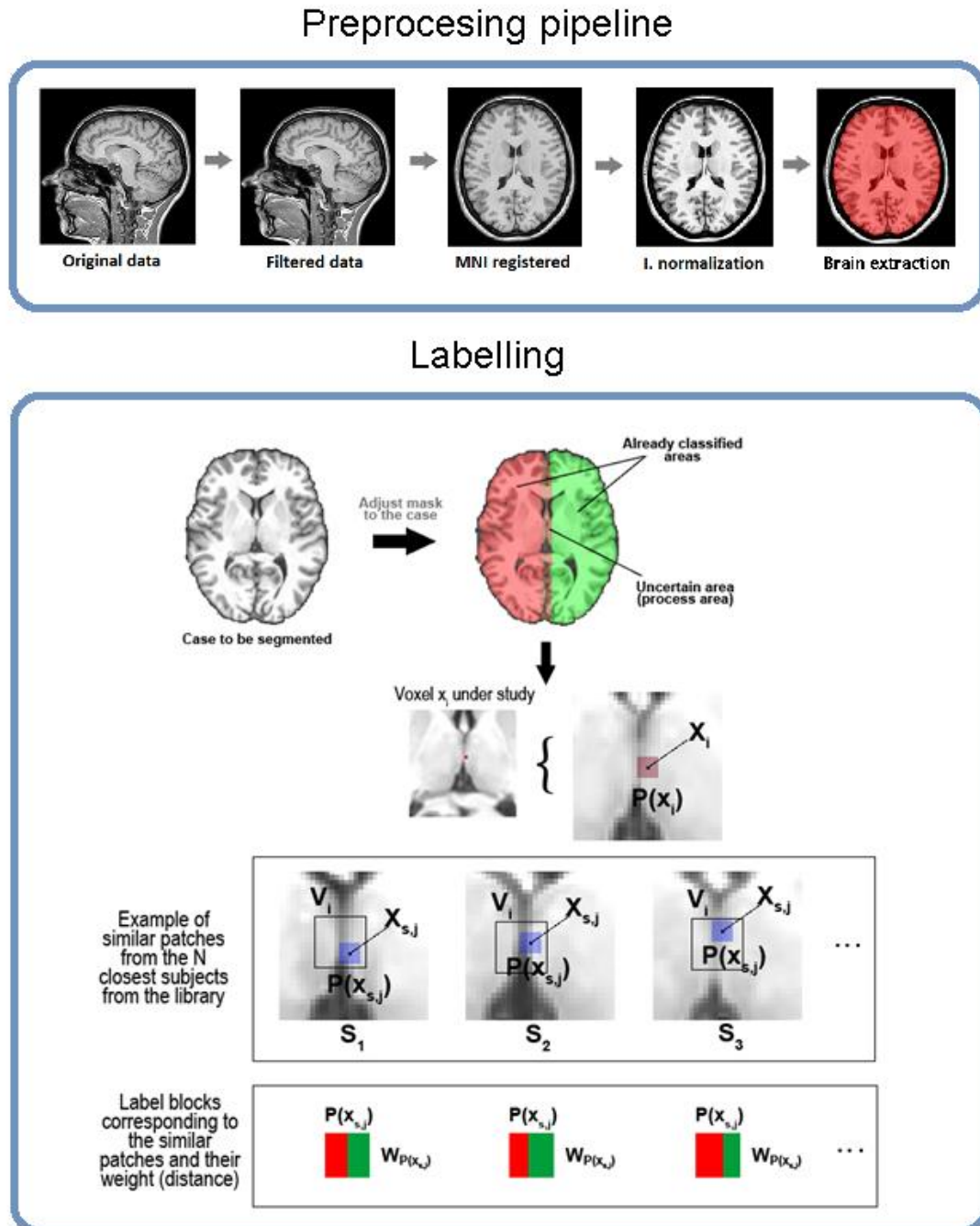


Figura 3.2.4: Visión general del preproceso y la segmentación resumiendo los pasos implicados en la selección de bloques  $P(x_{s,j})$  del área de búsqueda  $V_i$  en varios vasos de la librería. El peso  $w$  almacena el valor de la similitud entre  $P(x_{s,j})$  y  $P(x_i)$  el cual será usando como voto.

### 3.3. Experimentos y resultados

Para validar el método propuesto realizamos una validación cruzada *Leave-One-Out* (LOO) utilizando los 30 casos de la librería. Este proceso consiste en extraer casos de la librería (uno cada vez) y procesarlo como si de un caso sin segmentar se tratase. Para medir la precisión de las segmentaciones generadas por el método las comparamos con las segmentaciones manuales. Para ello utilizamos el índice Kappa de Cohen (Cohen, 1960) que permite evaluar la precisión de múltiples etiquetas. El cálculo del índice Kappa solo tienen en cuenta los vóxeles correspondientes al tejido cerebral. Todos los experimentos fueron realizados utilizando MATLAB 2009b (Mathworks Inc).

#### 3.3.1. Tamaño del área de búsqueda

El método de segmentación busca correspondencias entre patrones dentro de un área de búsqueda definida. El tamaño de esta área tiene un impacto importante sobre la precisión de las segmentaciones por lo que realizamos un experimento para encontrar su valor óptimo. La figura 3.3.1 muestra el índice Kappa promedio para las 30 imágenes de la librería en función del tamaño del área de búsqueda que se define como un cubo cuyos lados miden  $2 \times \text{radio} + 1$  vóxeles (en todos los experimentos se utilizó un tamaño de parche de  $5 \times 5 \times 5$  vóxeles).

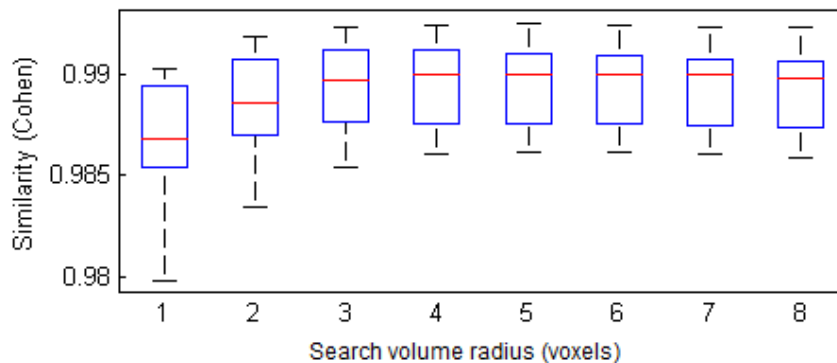


Figura 3.3.1: Impacto del tamaño del área de búsqueda sobre la precisión.

Como puede verse la precisión máxima se alcanza con un valor de radio = 4 (volumen de  $9 \times 9 \times 9$  vóxeles) lo cual es consistente con los hallazgos en (Coupé et al., 2011). Por tanto para el resto de experimentos utilizamos dicho valor.

#### 3.3.2. Tamaño de la librería

El número de casos que se utilizan en la segmentación tiene un impacto considerable en el coste computacional y en la precisión por lo que realizamos este experimento para encontrar el número óptimo de casos a utilizar. Para ello hacemos una preselección de los casos más parecidos a nuestro caso a segmentar utilizando como medida de similitud la media del valor absoluto de la diferencia o *Mean Absolute Difference* (MAD):

$$MAD(s) = \frac{1}{|\Omega|} \sum_{i=1}^{|\Omega|} |x_{s,i} - x_i| \quad (3.5)$$

donde  $|\Omega|$  representa el número de vóxeles.

La figura 3.3.2 muestra el Kappa promedio en función del tamaño de la librería.

Como puede apreciarse la calidad del método se estabiliza en torno a 20 casos. Este resultado está en consonancia con otros métodos que usan MALF (Aljabar et al., 2009; Coupé et al., 2011) donde también se estabilizan los resultados en torno a 20 casos. Por tanto, para el resto de experimentos utilizamos este valor.

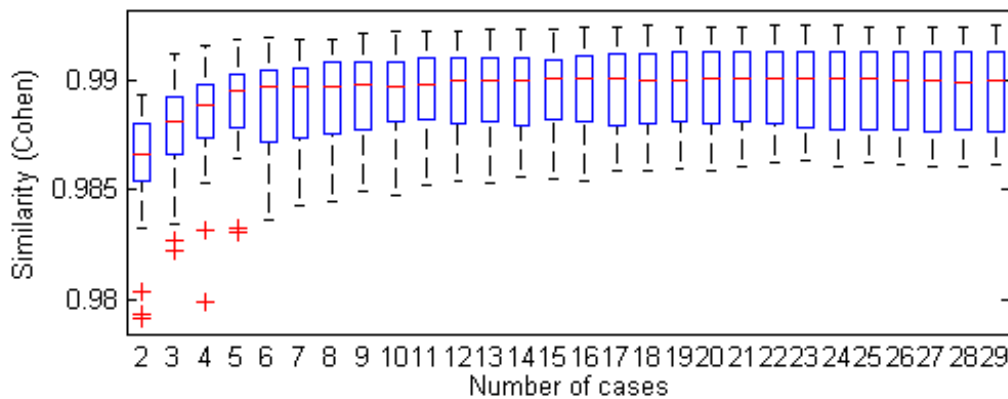


Figura 3.3.2. Impacto del tamaño de la librería en la precisión.

### 3.3.3. Implementación Block-wise

Como se ha explicado, se ha introducido una modificación en el método *Nonlocal label fusion* basada en bloques que mejora la precisión de las segmentaciones y reduce drásticamente la carga computacional. Se han comparado ambas versiones del método, enfoque basado en vóxel y enfoque basado en bloque, para evaluar las diferencias. La precisión en promedio obtenida por la versión basada en vóxel fue de  $0.9889 \pm 0.023$  mientras que la versión basada en bloque fue de  $0.9895 \pm 0.0020$  (ver figura 3.3.3). La precisión obtenida por ambas implementaciones resultó ser significativamente diferente con una  $p = 0.0019$  (t-test pareado). El tiempo de ejecución también se vio sustancialmente reducido siendo de 185 segundos para la implementación enfocada a vóxel y 60 segundos para la implementación enfocada a bloques. El factor de aceleración fue solo de 3 en contraste con el teórico 8. Esto se debe a la sobrecarga de otras partes de algoritmo como la carga de la librería y el proceso de fusión final.

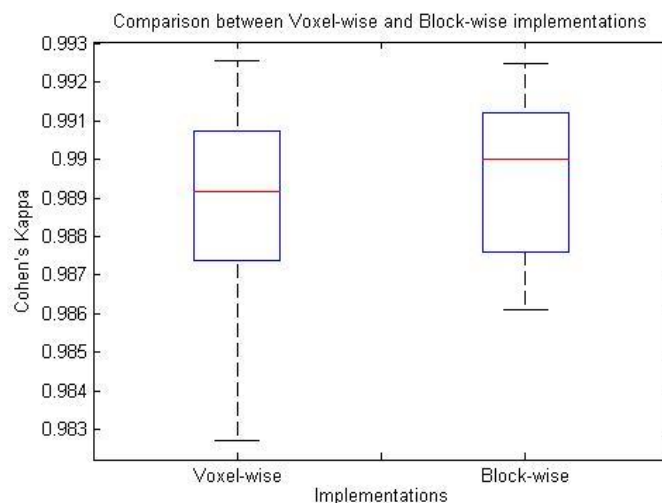


Figura 3.3.3. Comparación entre implementaciones voxel-wise y block-wise en términos de precisión.

### 3.3.4. Comparativa con el método Adaptive Disconnection

Para comparar el método propuesto con el *estado del arte* hemos realizado una comparativa con el método *Adaptive Disconnection* (ADisc) (Zhao et al., 2010) el cual ha demostrado obtener los mejores resultados en bibliografía.

Para que la comparativa fuera justa aplicamos el mismo preprocesado a las imágenes para ambos métodos. El código de *Adaptive Disconnection* fue obtenido en <http://www.cs.tut.fi/~jupeto/software.html>.

El método ADisc se basa en el uso iterativo del algoritmo *shape bottlenecks* (Mangin et al., 1996). Este algoritmo modela una transferencia de información en una imagen binaria utilizando ecuaciones de Laplace y puede detectar conexiones a modo de puente entre dos partes de un objeto complejo. Mediante este algoritmo, ADisc detecta las conexiones ente hemisferios cerebrales, cerebelo y bulbo raquídeo en la sustancia blanca. Después reconstruye tres compartimentos utilizando un crecimiento de regiones. Finalmente los hemisferios y el cerebelo son separados en izquierdo y derecho utilizando de nuevo el algoritmo *shape bottlenecks*.

Para realizar la comparativa hemos realizado una serie de experimentos con ambos métodos, NABS y ADisc. En primer lugar, hemos medido la precisión de las segmentaciones de ambos métodos (figura 3.3.4). Para que la comparativa fuera justa solo se han tenido en cuenta vóxeles etiquetados como cerebro por ambos métodos ya los métodos difieren en cuanto lo se considera cerebro y lo que no (esto puede apreciarse en la figura 3.3.5). De ahí que NABS obtenga resultados diferentes a los del experimento anterior. Se presenta la precisión global (todos los vóxeles) y sobre la región de interés o ROI (de sus siglas en inglés "region of interest") definida como la región de incerteza (ver figura 3.2.3)

### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

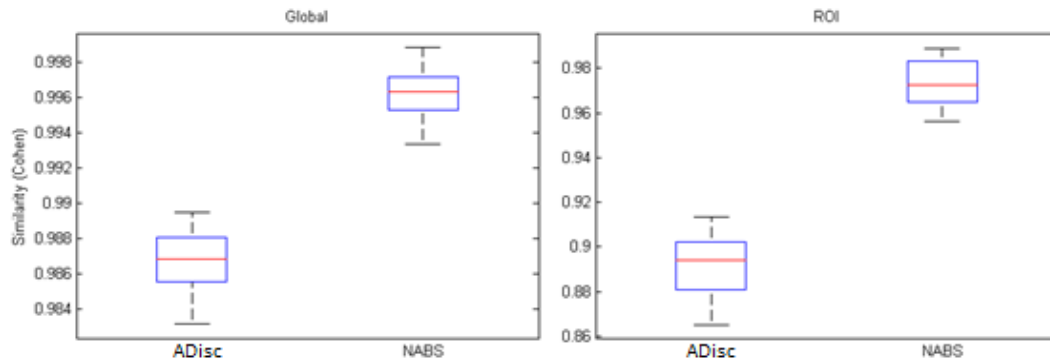


Figura 3.3.4: Precisión obtenida por ADisc y NABS teniendo en cuenta todos los vóxeles del cerebro (Global) y teniendo en cuenta solo los vóxeles visitados por NABS (ROI).

En la comparativa global, ADisc obtuvo una precisión de  $0.9868 \pm 0.0017$  mientras que NABS obtuvo un valor de  $0.9962 \pm 0.0014$ . En la comparativa sobre la ROI ADisc obtuvo una precisión de  $0.8912 \pm 0.0137$  y NABS obtuvo una precisión de  $0.9729 \pm 0.0103$ . Las diferencias entre ambos métodos son estadísticamente significativas (con una  $p = 3.7 \times 10^{-20}$  en la comparativa global y una  $p = 1.6 \times 10^{-22}$  en la comparativa sobre la ROI aplicando un t-test pareado).

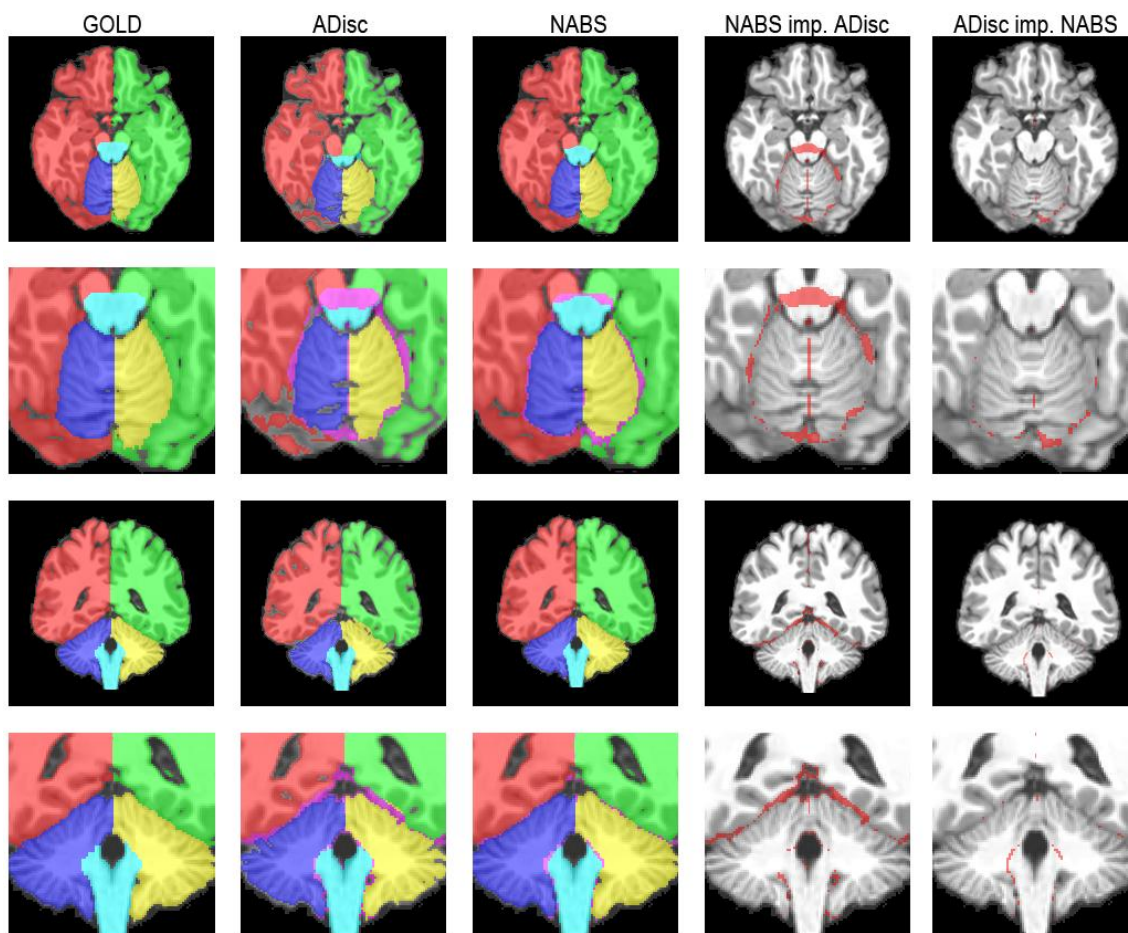


Figura 3.3.5: Ejemplo de resultados. La columna GOLD muestra la segmentación manual. Las columnas ADisc y NABS muestran las segmentaciones de ADisc y NABS respectivamente con los errores resaltados en morado. Las columnas NABS imp. ADisc and ADisc imp. NABS muestran las regiones donde NABS mejora frente a ADisc y viceversa.

### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

Para entender mejor las diferencias entre ambos métodos se ha medido la precisión para cada estructura de forma independiente. Como en el experimento anterior se han medido los resultados globalmente y sobre la ROI. Esta vez utilizamos el índice DICE (Zijdenbos et al., 1994) ya que comparamos una única clase.

Tabla 3.3.1: Dice promedio por cada estructura para ambos métodos. HC = Hemisferio cerebral, HCB = Hemisferio cerebelar, BR = Bulbo raquídeo. \*Todas las diferencias son estadísticamente significativas con una  $p < 0.05$ .

	Método	HC Izq.	HC Der.	HCB Der.	HCB Izq.	BR
Global	ADisc	0.9961 ±0.0007	0.9959 ±0.0006	0.9717 ±0.0074	0.9705 ±0.0068	0.9315 ±0.0213
	NABS	<b>0.9985 ±0.0006</b>	<b>0.9984 ±0.0006</b>	<b>0.9906 ±0.0048</b>	<b>0.9900 ±0.0053</b>	<b>0.9817 ±0.0074</b>
ROI	ADisc	0.9467 ±0.0092	0.9490 ±0.0112	0.8698 ±0.0476	0.8698 ±0.0444	0.8189 ±0.0691
	NABS	<b>0.9818 ±0.0042</b>	<b>0.9817 ±0.0057</b>	<b>0.9648 ±0.0157</b>	<b>0.9617 ±0.016</b>	<b>0.9445 ±0.0267</b>

Como puede verse en la tabla 3.3.1, NABS presenta una gran mejora para el cerebelo y bulbo raquídeo tanto globalmente como sobre la ROI. Esto explica el mejor funcionamiento de NABS en áreas afectadas por el volumen parcial como los límites del cerebelo.

#### 3.3.5. Estimación de la asimetría

Ya que una de las principales aplicaciones de la segmentación de los hemisferios es la estimación de la asimetría se ha comparado el impacto de usar ambos métodos para este fin. Definimos la asimetría como:

$$\text{Asimetría} = \frac{R - L}{R + L} \quad (3.6)$$

donde R y L son los volúmenes correspondientes las mitades derecha e izquierda de la estructura de interés.

Se han calculado los índices de asimetría para ambos cerebro y cerebelo utilizando las segmentaciones producidas por ambos métodos y comparándolas con el índice de asimetría derivado de la segmentación manual. Los resultados se muestran en la figura 3.3.6. Puede verse que ADisc sub-estima ligeramente la asimetría para ambas estructuras mientras el error promedio de NABS es cercano a cero. Además, la varianza del error es menor para NABS. El error promedio para ADisc fue de  $-0.0016 \pm 0.0029$  y para NABS fue de  $0.00004 \pm 0.0012$ .

### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

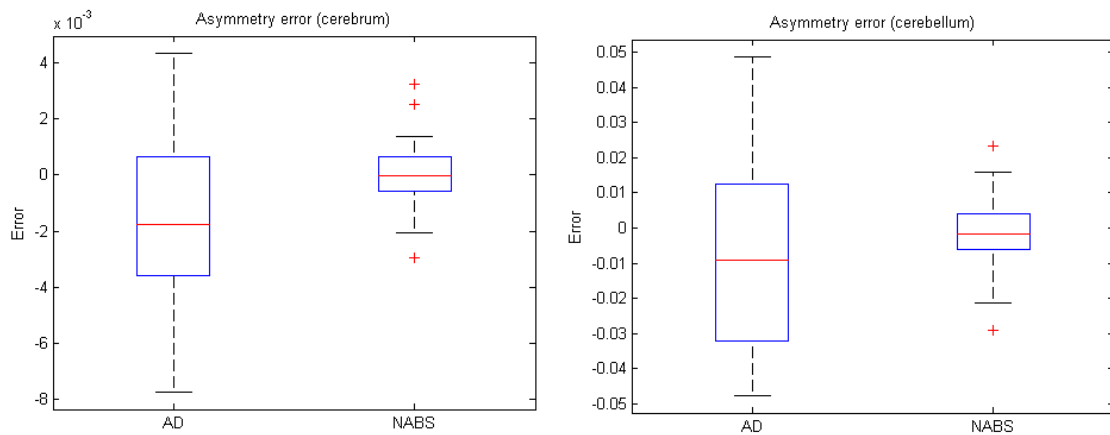


Figura 3.3.6: Error de asimetría sobre cerebro y cerebelo para ambos métodos.

#### 3.3.6. Tolerancia a lesiones y atrofia

Con el fin de comparar la robustez de ambos métodos cuando se enfrentan a casos patológicos se ha realizado un experimento introduciendo una lesión simulada en el hemisferio cerebral derecho de una de las imágenes de la librería. Para ello se ha reducido la intensidad en una zona delimitada al 50% de su valor real. La lesión simulada tenía un volumen de  $12 \text{ cm}^3$ . La figura 3.3.7 muestra el resultado de ambos métodos.

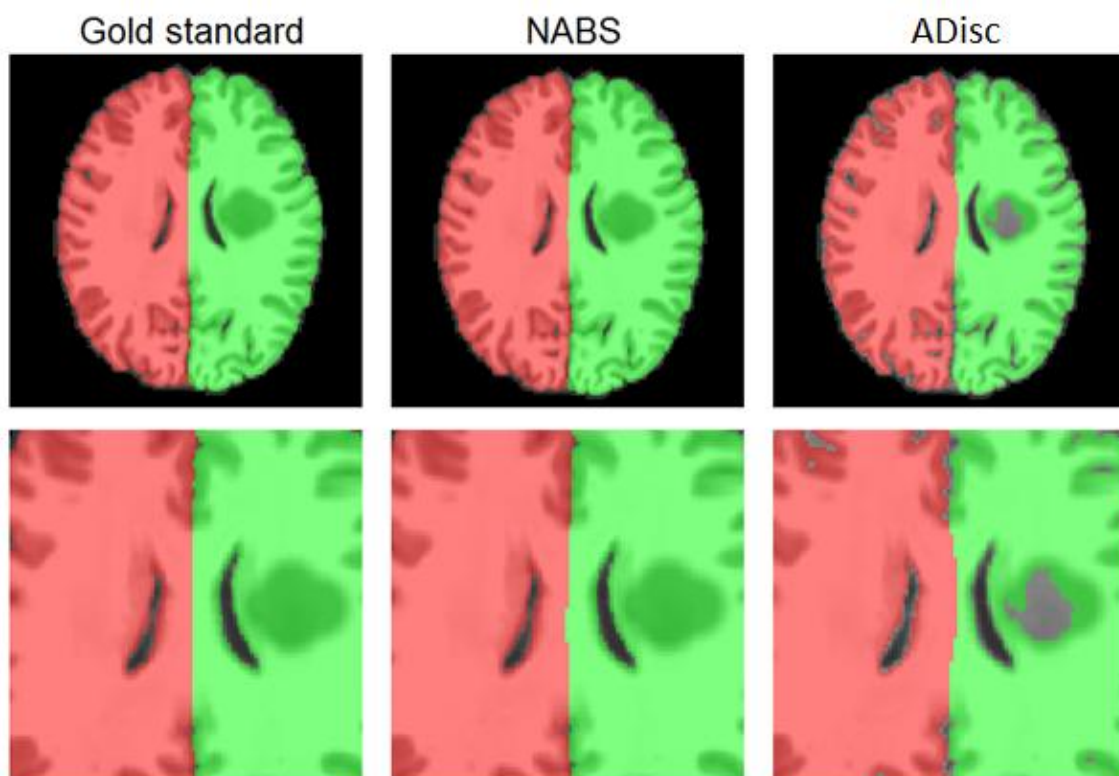


Figura 3.3.7: De izquierda a derecha: segmentación manual, segmentación de NABS y segmentación de ADisc.

Podemos observar que la segmentación de NABS no se ve afectada por la presencia de la lesión mientras que la segmentación de ADisc no solo falla en el núcleo de la lesión sino que también aparece una perturbación en la frontera entre hemisferios. La precisión obtenida por ADisc fue 0.8961 mientras que la obtenida por NABS fue 0.9656 (sobre la ROI).

#### 3.3.7. Estimación de la asimetría en pacientes con demencia

Se ha observado que pacientes con Alzheimer presentan una mayor asimetría cerebral que pacientes con un envejecimiento normal (Thompson et al., 1998). Para demostrar el valor clínico de NABS se ha realizado un experimento comparando la estimación de la asimetría obtenida con ambos métodos sobre sujetos normales y pacientes con demencia.

Para este experimento se han utilizado imágenes potenciadas en T1 de 40 sujetos diestros entre 60 y 90 años de edad (un grupo de control de 10 hombres y 10 mujeres considerados como no-dementes (CRD = 0) y un grupo de 10 hombres y 10 mujeres diagnosticados con demencia leve (CRD = 1). Estas imágenes se obtuvieron de la base de datos OASIS (Marcus et al., 2007). Se han procesado las imágenes con ambos métodos y se ha calculado la asimetría para cerebro y cerebelo. Finalmente, se analizaron las diferencias entre controles y dementes. Se ha estudiado el valor absoluto de la asimetría en lugar de la lateralización de la atrofia cerebral (Derflinger et al., 2011) (ver figura 3.3.8). Nótese que se ha usado la misma librería de 30 casos normales para la segmentación con NABS.

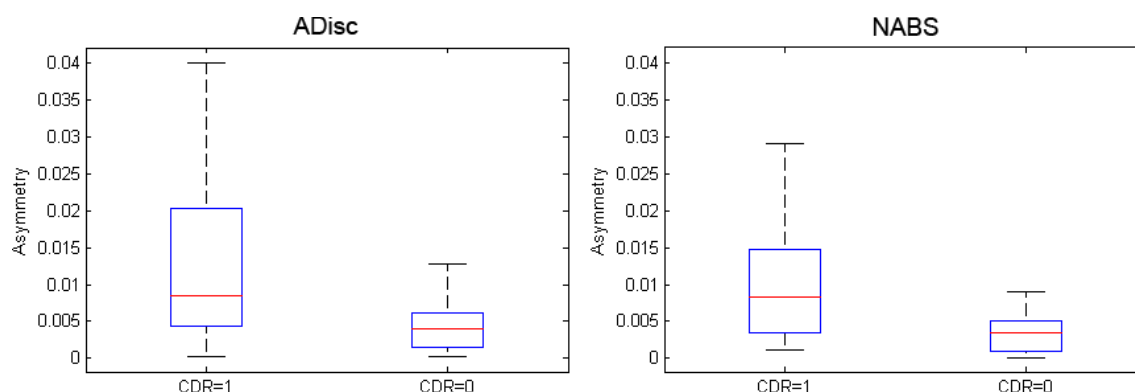


Figura 3.3.8: Asimetría cerebral en función de los grupos de control (CRD=0) y el grupo de dementes (CRD=1) calculada con Adisc y NABS.

Se han evaluado los resultados realizando un t-test pareado. Ambos métodos presentaron diferencias significativas entre ambos grupos ( $p = 0,0012$  para ADisc y  $p = 0,0011$  para NABS) lo cual es consistente con publicaciones previas (Thompson et al., 1998). Para profundizar más en los resultados y comprender mejor las diferencias entre grupos se han calculado las curvas ROC (de sus siglas en inglés Receiver-Operating Characteristic) para la estimación de la asimetría como predictor de demencia (ver figura 3.3.9) y se ha estimado el área bajo la curva ROC (AUCs) como índice que resume la información contenida en las curvas ROC (Hanley and McNeil, 1982). El AUC para las asimetrías estimadas por NABS fue de 0,80 para el cerebro y 0,70 para el cerebelo y la AUC para la asimetría estimada por ADisc fue de 0,77 para el cerebro y de 0.53 para el cerebelo.



### Capítulo 3. Segmentación de hemisferios cerebrales (NABS)

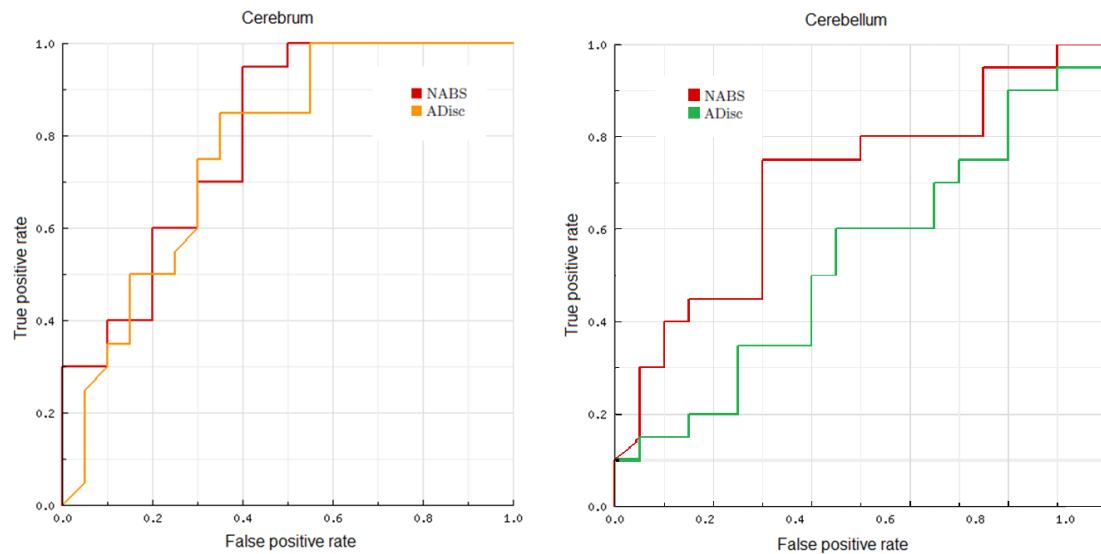


Figura 3.3.9: Comparación de curvas ROC a partir de las asimetrías calculadas por NABS y ADisc para cerebro y cerebelo. Sensibilidad y especificidad en el umbral óptimo fueron 0,95 y 0,60 (Cerebro con NABS); 0,85 y 0,65 (Cerebro con ADisc); 0,75 y 0,70 (Cerebelo con NABS); 0,60 y 0,55 (Cerebelo con ADisc. Para computar las curvas se utilizó (Vergara et al., 2008).

El test de DeLong (Vergara et al., 2008) no mostró diferencias significativas entre las AUCs de ambos métodos lo cual podría deberse al reducido tamaño de la muestra. Sin embargo las curvas ROC sugieren que la asimetría estimada por NABS es un mejor predictor de demencia.

Para ilustrar esto, en la figura 3.3.10 se puede observar una alteración en la frontera de los hemisferios cerebrales en la segmentación de ADisc. Esto puede ser causado por la diferencia de volumen de ambos hemisferios lo cual afecta al resultado de ADisc.

Se realizó un último experimento para ilustrar la robustez de NABS a la presencia de anomalías/atrofias cerebrales extremas. Se ha procesado un caso de OASIS con una enorme atrofia con ambos métodos. El sujeto era un varón de 86 años de edad diagnosticado con Alzheimer (CDR = 2). Como se puede ver en la figura 3.3.11 NABS produce una segmentación aceptable mientras que ADisc falla en varias regiones (especialmente en el cerebelo).

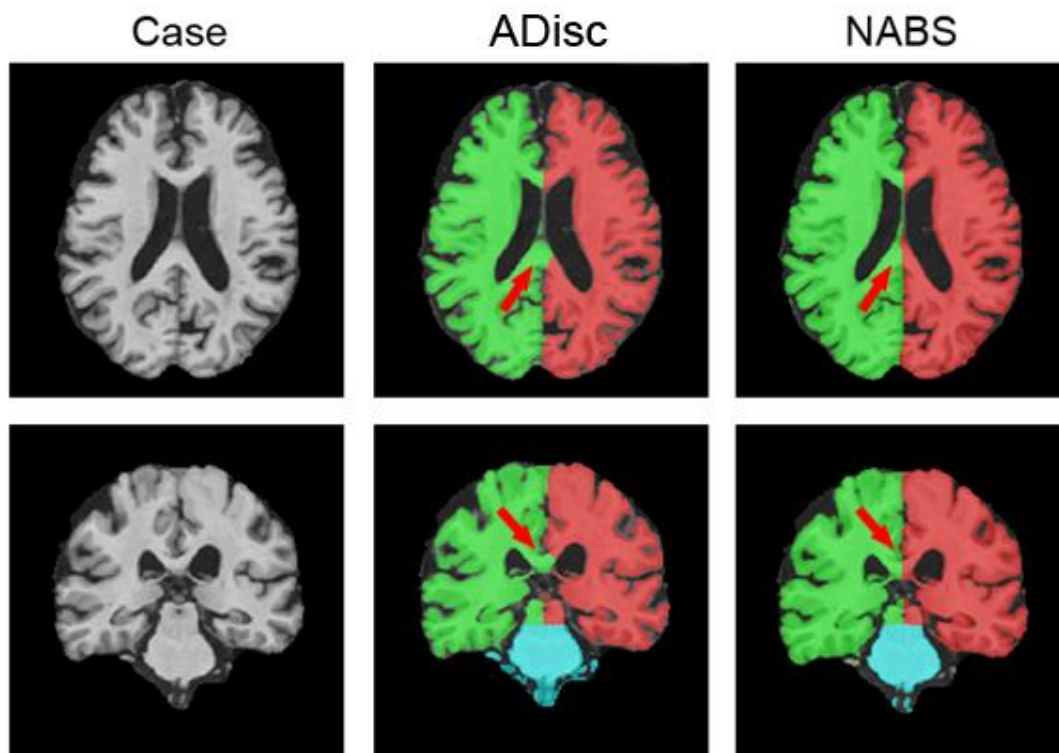


Figura 3.3.10: De izquierda a derecha: Imagen preprocesada, segmentación de ADisc y segmentación de NABS. Las flechas señalan puntos donde ADisc falla mientras que NABS clasifica los vóxeles correctamente.

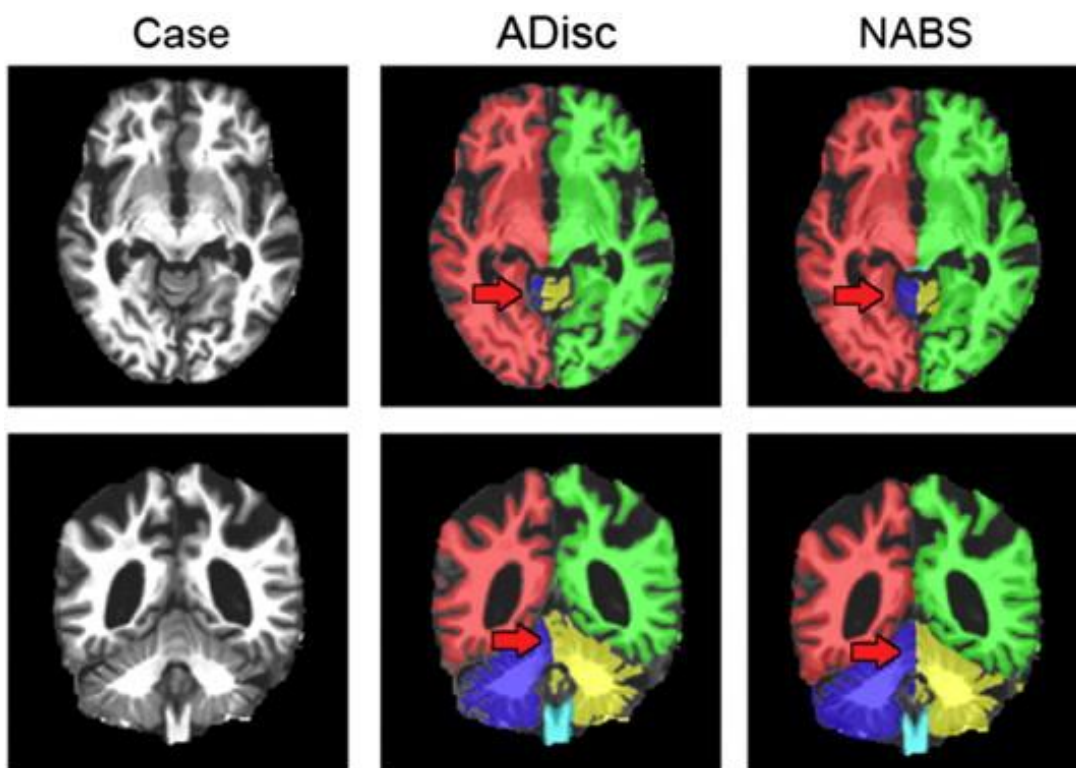


Figura 3.3.11: Ejemplo de un sujeto de la base de datos OASIS con una atrofia importante. De izquierda a derecha: Imagen preprocesada, segmentación de ADisc y segmentación de NABS. Las flechas señalan las regiones donde ADisc falla mientras NABS clasifica los vóxeles correctamente.

### 3.3.8. Coste computacional

El tiempo de ejecución requerido por ADisc es de unos 5 minutos por caso en promedio mientras que NABS realiza la segmentación en 1 minuto por caso en promedio. Todos los experimentos incluida la medición del tiempo de cómputo fueron realizados en una maquina con un procesador Intel core i7 y 8 GB de memoria RAM ejecutando MATLAB 2009b en un sistema Windows 7 Profesional 64 bit.

## 3.4. Conclusión y discusión

En este capítulo hemos presentado un método de segmentación *multi-atlas* basado en parches que es preciso y rápido al cual hemos llamado NABS y que segmenta el cerebro en cinco estructuras: hemisferios cerebrales derecho e izquierdo, hemisferios cerebelares izquierdo y derecho y bulbo raquídeo. La principal novedad de NABS es la adaptación del método *Nonlocal label fusion* para segmentar múltiples estructuras, la segmentación por bloques completos y el uso de una máscara que incorpora información a priori obtenida de la librería que reduce el coste computacional al reducir el área sobre la que se debe computar la segmentación. Estas optimizaciones redujeron el tiempo de ejecución de una hora a tan solo un minuto.

Hemos demostrado que NABS es capaz de segmentar con precisión las estructuras cerebrales en sujetos sanos para un amplio rango de edades. Además, hemos presentado evidencias de que NABS seguiría siendo preciso a la hora de procesar sujetos atróficos o con lesiones incluso con una librería compuesta solo por sujetos sanos. Esto resulta interesante ya que denota que no es necesario (aunque si interesante) construir librerías específicas de cada patología para utilizar NABS.

Evidentemente no se sabe los resultados que podría ofrecer el método con sujetos de otras patologías. No obstante, pese a que NABS pudiera fallar en algunos casos atípicos, estos podrían ser incluidos en la librería para aumentar su robustez.

Hemos realizado también una comparativa con el método *Adaptive Disconnection* el cual representa el estado del arte en cuanto a segmentación de hemisferios y ha sido validado contra varios métodos de segmentación. Esta comparativa ha demostrado que NABS es significativamente más preciso que ADisc. Como aplicación práctica, se ha estudiado la asimetría cerebral utilizando las segmentaciones producidas por ambos métodos. Tanto ADisc como NABS detectaron un aumento de la asimetría significativo en casos diagnosticados con demencia respecto a controles sanos. No obstante el tamaño de efecto fue mayor (y los p-valores menores) para NABS que para ADisc lo cual puede ser indicativo de que la segmentación de NABS es más precisa/específica.

Recientemente, NABS ha incorporado algunas mejoras resultantes del desarrollo de otros métodos similares dentro de la plataforma volBrain. En concreto, se ha sustituido el registro del SPM por uno más robusto utilizando la herramienta ANTs ([Avants et al., 2009](#)). Además, se ha mejorado la extracción intracraneal utilizando el método NICE ([Manjón et al., 2014](#)) y finalmente se ha incorporado el método de corrección de inhomogeneidades N4 ([Tustison et al., 2010](#)).

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

En este capítulo se presenta un nuevo método para la segmentación de los lóbulos del cerebelo usando imágenes de Resonancia Magnética Nuclear. Este método fue publicado en 2017 ([Romero et al., 2017a](#)) en la revista Neuroimage y se ha hecho público para la comunidad científica a través de la plataforma online volBrain (<http://volbrain.upv.es>).

### 4.1. Segmentación del cerebelo

El cerebelo es una estructura neuroanatómica situada debajo del cerebro y detrás del bulbo raquídeo al que está conectado a través de los pedúnculos cerebelares. Pese a que representa un 10% del volumen intracraneal total, juega un papel clave en la coordinación y el aprendizaje motor ([Manto et al., 2013](#)). El cerebelo actúa como modulador de los impulsos de la corteza motora para hacer posible movimientos precisos. Esto es posible ya que el cerebelo integra las señales de la corteza somatosensorial con las de la corteza cerebelar. Se ha establecido que las lesiones en el cerebelo implican disfunciones motoras importantes como la ataxia de miembros, alteraciones del equilibrio y otras deficiencias en la coordinación ([Compston and Coles, 2008](#); [Davie et al., 1995](#); [Kase et al., 1993](#); [Klockgether, 2008](#)). Además, estudios recientes apuntan a que el cerebelo está implicado en muchas funciones cognitivas ([Stoodley, 2012](#)) como la atención, las emociones o el comportamiento ([Timmann et al., 2010](#)).

Pese a las evidencias de la implicación del cerebelo en muchas funciones cerebrales, este ha sido muy poco estudiado hasta hace relativamente poco tiempo, cuando aparecieron estudios enfocados a confirmar la relación entre el cerebelo y varias enfermedades neuropsiquiátricas. Por ejemplo, los pacientes con esquizofrenia muestran una reducción del vermis ([Okugawa et al., 2002](#)) y sujetos con Alzheimer presentan una reducción de los lóbulos posteriores ([Thomann et al., 2008](#)). Sin embargo, existen publicaciones con resultados inconcluyentes al respecto ([James et al., 2004](#); [Nenadic et al., 2010](#)). Esto podría explicarse con el hecho de que la segmentación del cerebelo es un problema muy difícil de resolver debido a su gran complejidad anatómica.

La anatomía del cerebelo consiste en una estructura de sustancia blanca que crece desde el bulbo raquídeo en forma de dos árboles. Cada uno da origen a un hemisferio cerebelar. Esta sustancia blanca está envuelta por una capa de sustancia gris forma pliegues los cuales se agrupan en lóbulos y a su vez en lobulillos. El tamaño, posición y número (visible) de estos lobulillos es altamente variable entre sujetos lo cual hace de la segmentación del cerebelo una tarea compleja.

La segmentación manual por expertos requiere de un gran esfuerzo pues requiere de mucho tiempo al ser un proceso tedioso. La segmentación manual por no-expertos es

más rápida pero sigue sin ser aplicable a grandes conjuntos de datos.

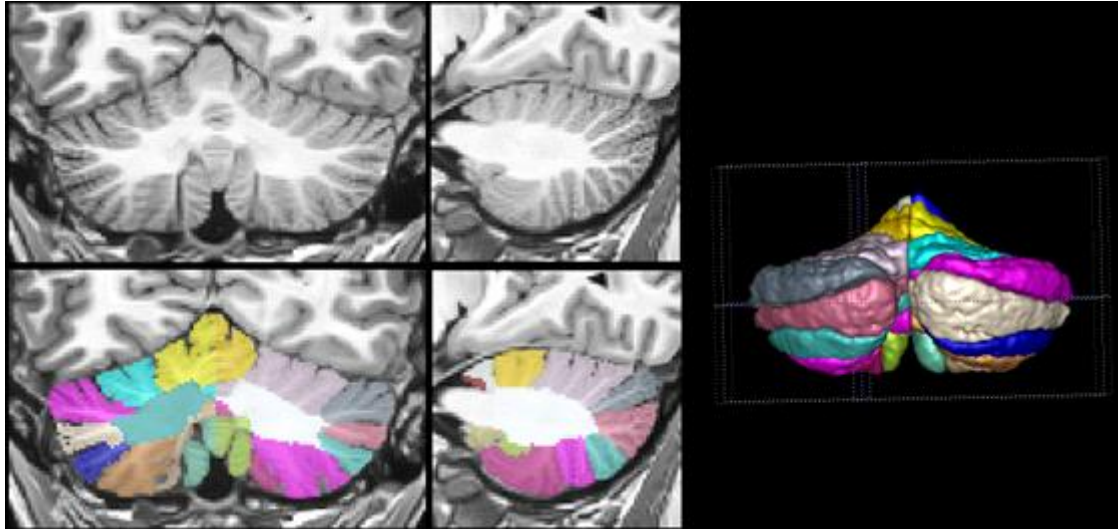


Figura 4.1.1: Izquierda: Ejemplo de una imagen de RMN del cerebelo y sus segmentación. Derecha: Representación 3D de los lobulillos del celébrelo.

Además, ambas opciones implican lidiar con la variabilidad inter e intra-observador. Algunos métodos semi-automáticos han sido desarrollados ofreciendo segmentación de grandes estructuras como el bulbo raquídeo, sustancia blanca y sustancia gris (Weier et al., 2012). Estos métodos buscan reducir el tiempo necesario para realizar una segmentación manual pero aun así resulta costosa. En la figura 4.1.1 se muestra a modo de ejemplo una imagen de RMN de un cerebelo así como la subdivisión manual de sus lobulillos.

Una alternativa a la segmentación manual son los métodos MALF (Aljabar et al., 2009) que como ya se ha comentado, consisten en transferir y combinar múltiples segmentaciones manuales a un caso sin segmentar. Uno de los primeros métodos automáticos desarrollados específicamente para el cerebelo es SUIT (Diedrichsen, 2006). En este método, un conjunto de imágenes segmentadas a mano se utilizó para crear un atlas probabilístico que es deformado a la imagen a segmentar. Sin embargo, SUIT no es capaz de abarcar toda la variabilidad anatómica debido al uso de un único atlas. Además, el método incluye un paso de identificación del cerebelo, el cual, se sugiere que sea revisado a mano por lo que podríamos considerarlo un método semi-automático.

En los últimos años se han publicado métodos totalmente automáticos como ACCLAIM (Bogovic et al., 2013) que usa modelos deformables en combinación con un clasificador *random forest* para clasificar los vóxeles en las fronteras entre estructuras. Recientemente, Yang (Yang et al., 2016) propuso un método de segmentación utilizando MALF junto con clasificación de tejidos (y fronteras entre estructuras) utilizando *Graph-cuts* para abarcar las variabilidad de la anatomía del cerebelo logrando buenos resultados a expensas de un coste computacional muy alto (varias horas de proceso por caso). Otro método de segmentación basado en MALF para la segmentación de los lobulillos cerebelares es MAGeT (Chakravarty et al., 2013; Park et al., 2014). Este método, se enfoca en resolver un problema muy común en MALF, la poca disponibilidad de segmentaciones manuales. Se trata de una estrategia que se

basa en la estimación de un elevado número de registros no lineales para incrementar el número de segmentaciones candidatas en la etapa de fusión de etiquetas. MAGeT demostró obtener mejores resultados que las técnicas que utilizan un único atlas pero con un coste temporal y de memoria mucho mayor. RASCAL (Weier et al., 2014) también es un método MALF que utiliza una librería de casos segmentados a mano. Esta librería se mapea sobre el caso a segmentar mediante registro no lineal para realizar una fusión de etiquetas basada en parches (Coupé et al., 2011).

A continuación, se presentara un nuevo método de segmentación de los lobulillos del cerebelo y la sustancia blanca (SB) denominado CERES (*CEREbellum Segmentation*) basado en la adaptación de un método de segmentación reciente llamado *Optimized PatchMach Label fusion* (OPAL) (Giraud et al., 2016; Ta et al., 2014). OPAL es un método basado en MALF que utiliza una versión eficiente del *Nonlocal label fusion* para producir una segmentación rápida y precisa utilizando una librería de casos segmentados a mano. El método original que fue desarrollado para la segmentación del hipocampo y los ventrículos laterales (Coupé et al., 2011) ha sido adaptado para segmentar el cerebelo utilizando una librería de casos registrados no linealmente en lugar de usar solo un registro afín. Además, se ha añadido un paso adicional a la segmentación para potenciar la regularidad de los contornos. También, dado que solo disponemos de 5 casos segmentados a mano hemos usando una técnica de extensión de la librería añadiendo casos segmentados automáticamente.

## 4.2. Material y métodos

En las siguientes secciones, se describen los datos utilizados en el desarrollo del método así como los detalles del método propuesto.

### 4.2.1. Datos

En el desarrollo y la validación de este método se han utilizado dos conjuntos de datos diferentes. Se han utilizado 5 imágenes de RMN potenciadas en T1 de alta resolución segmentadas a mano para construir la librería. Además se han utilizado 20 imágenes de RM potenciadas en T1 de resolución estándar segmentadas a mano para realizar una comparativa con otros métodos. Las imágenes de alta resolución son de acceso público y se encuentran disponibles en la web de CoBrALab (<http://cobralab.ca/atlas/Cerebellum.html>). Ambos conjuntos de imágenes y sus segmentaciones manuales son los mismos que los utilizados en (Park et al., 2014).

#### *Datos de alta resolución*

Los datos de alta resolución son adquisiciones de 5 voluntarios sanos (2 hombres y 3 mujeres de entre 29 y 57 años) sobre las cuales se realizó una segmentación manual de los lobulillos del cerebelo y su sustancia blanca. Las imágenes fueron adquiridas utilizando un equipo GE Discovery MR 750 3T (General Electric, Waukesha, WI) utilizando una bobina de 8 canales. Para ello, se utilizó una secuencia *3D inversion-prepared fast spoiled gradient-recalled echo acquisition*, FSPGR-BRAVO con un tiempo de escaneo de aproximadamente 20 minutos con los siguientes parámetros:

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

TE/TR = 4.3 ms / 9.2 ms, TI = 650 ms,  $\alpha = 8^\circ$ , 2NEX, FOV = 22 cm, grosor de corte = 0.6 mm, resolución en plano de 384 x 384 píxeles para obtener un vóxel isotrópico de 0.6 mm<sup>3</sup>. Se usaron los filtros de reconstrucción ZIPX2 y ZIP512 obtener una resolución final de 0.3 mm<sup>3</sup> con un tamaño de matriz de 489x734x503 vóxeles (Park et al., 2014). Estas imágenes fueron segmentadas manualmente por dos expertos lo cual permitió estimar la precisión inter e intra-observador para evaluar la consistencia de la segmentación manual. El cerebelo se dividió en 26 estructuras: sustancia blanca y lobulillos I-II, III, IV, V, VI, Crus I, Crus II, VIIb, VIIa, VIIIa, VIIIb, IX y X distinguiendo hemisferios derecho e izquierdo como puede verse en la figura 4.2.1.

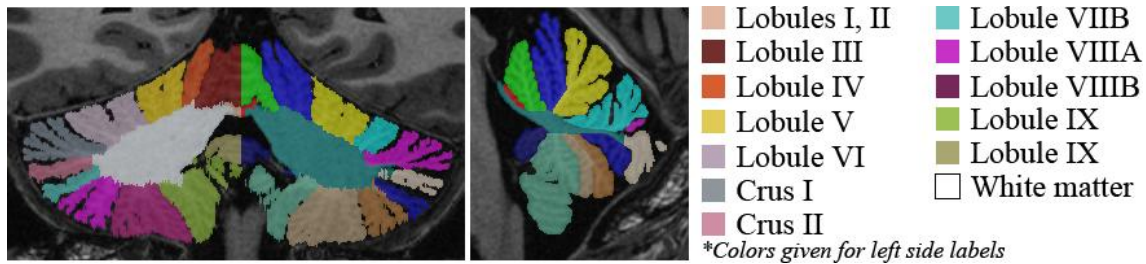


Figura 4.2.1: Ejemplo de imagen de alta resolución y su correspondiente segmentación manual.

### Datos de resolución estándar

Este conjunto de datos consiste en 10 imágenes de controles sanos y 10 imágenes de pacientes con esquizofrenia (15 hombres y 5 mujeres en un rango de edad de 25 a 35 años con una media de 30,1 años). Las imágenes fueron adquiridas utilizando un equipo GE Echospeed (General Electric Medical Systems, Milwaukee, WI) que permite una máxima amplitud de gradiente de 40 mT/m. Para ello, se utilizó una secuencia *Axial inversion recovery prepared spoiled gradient recall* con los siguientes parámetros: TE = 5.3 ms, TR = 12.3 ms, TI = 300 ms, ángulo = 20°, NEX1, grosor de corte = 1.5 mm con una resolución en plano de 0.78 x 0.78 mm. Estas imágenes fueron registradas al espacio MNI utilizando el software MRITotal (Collins et al., 1994) resultando en una resolución de 1 mm<sup>3</sup> con un tamaño de matriz de 181x217x181 vóxeles. El cerebelo como única estructura se segmentó manualmente en los 20 casos. Los lobulillos del hemisferio izquierdo se segmentaron sobre 10 imágenes mientras que los lobulillos del hemisferio derecho se segmentaron sobre las 10 restantes (24 lobulillos: I-II, III, IV, V, VI, Crus I, Crus II, VIIb, VIIIa, VIIIb, IX, y X).

### 4.2.2. Preprocesado

Todas las imágenes fueron preprocesadas para mejorar su calidad y llevarlas a un espacio de coordenadas común como se describe a continuación:

**Filtrado:** Para reducir el ruido y mejorar la calidad de las imágenes se utilizó el filtro SANLM (*Spatially Adaptive Non-Local Means*) (Manjón et al., 2010b) que puede trabajar con niveles de ruido variables.

**Corrección de inhomogeneidad:** Como paso previo para mejorar el registro se aplicó una corrección de inhomogeneidad preliminar utilizando el software N4 (Tustison et al., 2010). Esta imagen corregida se utiliza solo para mejorar la estimación del registro.

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

Registro lineal: Todas las imágenes fueron registradas a un espacio de coordenadas común para que la anatomía entre casos de la librería fuera consistente. Para ello, se aplicó a todas las imágenes un registro lineal (afín) para llevarlas al espacio MNI utilizando la plantilla MNI152. Esto se hizo utilizando el software *Advanced Normalization Tools* (ANTs) (Avants et al., 2009).

Corrección de inhomogeneidad en el espacio MNI: Tras el registro realizamos la corrección de inhomogeneidad definitiva utilizando la máscara intracraneal asociada a la plantilla MNI152 para una mejor estimación de la inhomogeneidad. Para ello utilizamos también el software N4 (Tustison et al., 2010).

Recorte de la región de interés: Para reducir el coste computacional se recortaron todas las imágenes de modo que solo apareciese la región del cerebelo. Para calcular esta región se realizó la suma de las segmentaciones manuales de todas las imágenes de la librería y se calcularon sus límites a los que se añadió un margen de 5 vóxeles en cada dirección.

Registro no-lineal de baja dimensionalidad: Para conseguir una mejor correspondencia anatómica entre la imagen a segmentar y la librería, el área recortada de ambos (imagen y librería) se registró no linealmente a la plantilla MNI152 (también recortada).

Aunque podríamos realizar la segmentación en este espacio no lineal, esto requeriría aplicar una transformación inversa para llevar la segmentación resultante al espacio lineal. Debido a que en esta transformación inversa aplica una interpolación los datos resultantes pierden calidad.

Para evitar esto, se ha optado por seguir una estrategia diferente. Se han estimado las transformaciones no-lineales de la librería al espacio MNI152 pero no se han aplicado sino que se han almacenado para su uso posterior. A la hora de segmentar un caso nuevo también se estima su transformación no-lineal al espacio MNI152. Finalmente se traslada toda la librería al espacio de la imagen a segmentar concatenando las transformaciones directas de cada caso con la transformación inversa del caso a segmentar. De este modo realizamos la segmentación en el espacio MNI152 lineal usando una librería registrada no-linealmente específica de cada caso.

Este paso de registro no lineal tiene un gran impacto en la segmentación. Por esta razón, se han evaluado tres niveles de resolución diferentes en el registro no lineal (baja =  $4 \text{ mm}^3$ , media =  $2 \text{ mm}^3$  y alta =  $1 \text{ mm}^3$ ). Generalmente se espera que con registros de alta dimensionalidad ( $1 \text{ mm}^3$ ) la anatomía obtenga la mejor correspondencia y por tanto la segmentación obtenga la mayor calidad. No obstante esto es muy dependiente de los datos con que se trabaja y en nuestros experimentos, un registro de baja dimensionalidad ( $4 \text{ mm}^3$ ) permitía obtener la mejor calidad de segmentación siendo además mucho más rápido de computar. El registro no-lineal se realizó utilizando ANTs (Avants et al., 2009).

Normalización en intensidad: Puesto que el método de segmentación encuentra correspondencias entre parches utilizando una métrica basada en los niveles de intensidad es necesario normalizar las imágenes de modo que los tejidos (sustancia blanca y sustancia gris del cerebelo) tengan los mismos niveles de intensidad entre sujetos. Para ello se ha aplicado una normalización local basada en regiones de



interés (ROIs) de forma similar a la propuesta (Asman and Landman, 2012). Específicamente, se ha tomado como referencia la plantilla MNI152 y se ha llevado al espacio de la imagen a segmentar utilizando su transformación inversa. Una vez registrada la plantilla y la librería al caso a segmentar calculamos las ROIs a partir de la librería como la etiqueta más frecuente para cada vóxel (*majority vote*). Después, se obtiene el valor mediano de intensidad para cada ROI para ambas imágenes (la imagen a segmentar y la plantilla MNI152). Con estos valores se ajusta un modelo polinomial de segundo orden. Finalmente, las intensidades del caso a segmentar se normalizan evaluando el modelo polinomial para el valor de intensidad de cada vóxel.

### 4.2.3. Creación de la librería

Para construir la librería, todas las imágenes de alta resolución fueron preprocesadas como se describe en el apartado anterior. Las segmentaciones manuales en el espacio nativo fueron linealmente registradas al espacio MNI152. Entonces, las imágenes y sus etiquetas fueron invertidas (izquierda-derecha) para aumentar el número de casos de la librería de 5 a 10. El hecho de que el cerebelo sea casi simétrico permite que el proceso de segmentación se beneficie de esto al aumentar el número de patrones localmente. Seguidamente las imágenes y sus etiquetas fueron recortadas en torno al cerebelo. Después, sus transformaciones no lineales al espacio MNI152 fueron estimadas. Estas transformaciones se usarán, como ya se ha comentado, para registrar toda la librería al espacio de la imagen a segmentar de forma similar a como se hace en (Weier et al., 2014). Finalmente, las imágenes fueron normalizadas en intensidad de la forma ya descrita. Nótese que la librería resultante en el espacio MNI tiene una resolución de  $1 \text{ mm}^3$  siendo la original de  $0.3 \text{ mm}^3$ . Esto se hizo para obtener segmentaciones a una resolución similar a la obtenida actualmente en entornos de investigación y clínicos.

### 4.2.3. Optimized PatchMatch label fusión

Uno de los principales problemas de la técnica *Nonlocal label fusion* es su elevado coste computacional. Este coste viene dado en su mayor parte por el proceso de búsqueda de parches similares (*patch matching*). Afortunadamente, recientes avances en el campo de la visión por computador han dado como resultado métodos eficientes para realizar búsquedas aproximadas de parches similares como por ejemplo el algoritmo PatchMatch.

#### *El algoritmo PatchMatch*

El algoritmo PatchMatch (Barnes et al., 2009) es una estrategia eficiente para encontrar correspondencias entre parches de dos imágenes (originalmente en 2D). Se basa en la hipótesis de que dado un parche  $p(i,j)$  perteneciente a una imagen  $A$  el cual se corresponde con un parche  $p(i',j')$  perteneciente a una imagen  $B$ , entonces los parches adyacentes a  $p(i,j)$  tienen una alta probabilidad de corresponderse con los parches adyacentes a  $p(i',j')$ .

El algoritmo se implementa en 3 pasos: inicialización, propagación y búsqueda aleatoria. En primer lugar se inicializa cada parche de  $A$  asignando aleatoriamente una

correspondencia con un parche de B. Después los parches adyacentes se evalúan para mejorar la similitud de cada correspondencia. Finalmente, con objeto de evitar mínimos locales, se realiza una búsqueda aleatoria y se evalúan varios parches para comprobar si mejoran la similitud actual. Los pasos 2 y 3 se repiten de forma iterativa para mejorar las correspondencias lo cual converge rápidamente obteniendo un conjunto de correspondencias muy cercano al óptimo denominado *Aproximated Nearest Neighbor Field* (ANNF). En la figura 4.2.2 se muestra el esquema del proceso.

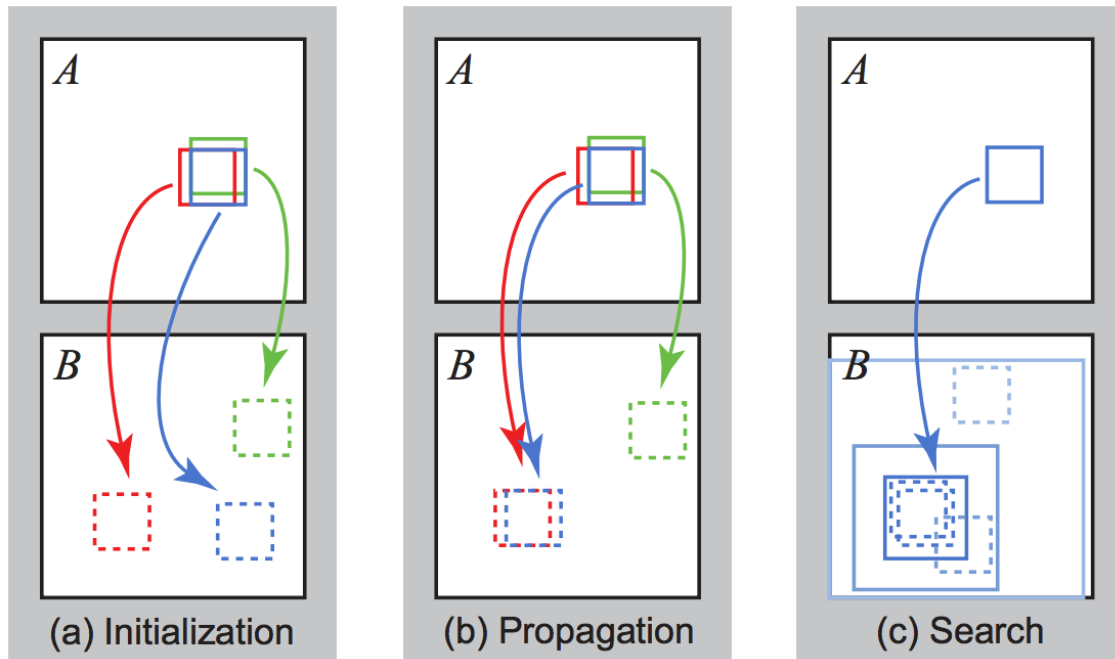


Figura 4.2.2: Funcionamiento del algoritmo PatchMatch: a) Inicialización aleatoria en la cual se asignan correspondencias a cada parche (las flechas indican la asignación de correspondencias), b) Propagación en la cual se evalúa la calidad de las correspondencias para parches adyacentes y c) Búsqueda aleatoria en la cual se evalúan nuevas correspondencias aleatorias. Fuente: Barnes et al., 2009.

### Optimized PatchMatch Label Fusion (OPAL)

La extensión a imágenes 3D y adaptado al problema de fusión de etiquetas fue propuesta por (Ta et al., 2014). El método, que fue denominado OPAL (*Optimized Patchmatch Label fusion*) aplica PatchMatch a imágenes de resonancia magnética 3D y encuentra correspondencias entre los parches de una imagen dada S y una librería L. Uno de los puntos fuertes de PatchMatch es que su complejidad computacional solo depende del tamaño de la imagen S y no de la librería L debido a su carácter probabilístico. Esto, permite utilizar una librería muy extensa sin tener que realizar ningún tipo de preselección a un coste prácticamente constante (sin tener en cuenta el tiempo de carga de la librería).

En comparación con PatchMatch, OPAL implementa una inicialización restringida a un área definida ya que habitualmente, el sujeto a segmentar y la librería se encuentran en el mismo espacio geométrico (MNI152) y por lo tanto se espera que las correspondencias para un parche dado se encuentren cerca de sus coordenadas. El

método OPAL fue extendido en 2016 por Giraud y colaboradores para incluir también el uso de imágenes multi-espectrales/multi-característica y multi-parámetro combinado con un enfoque llamado de *late fusion* (Snoek et al., 2005) para combinar los mapas de probabilidad de cada etiqueta (Giraud et al., 2016). La figura 4.2.3 muestra un esquema del funcionamiento de OPAL.

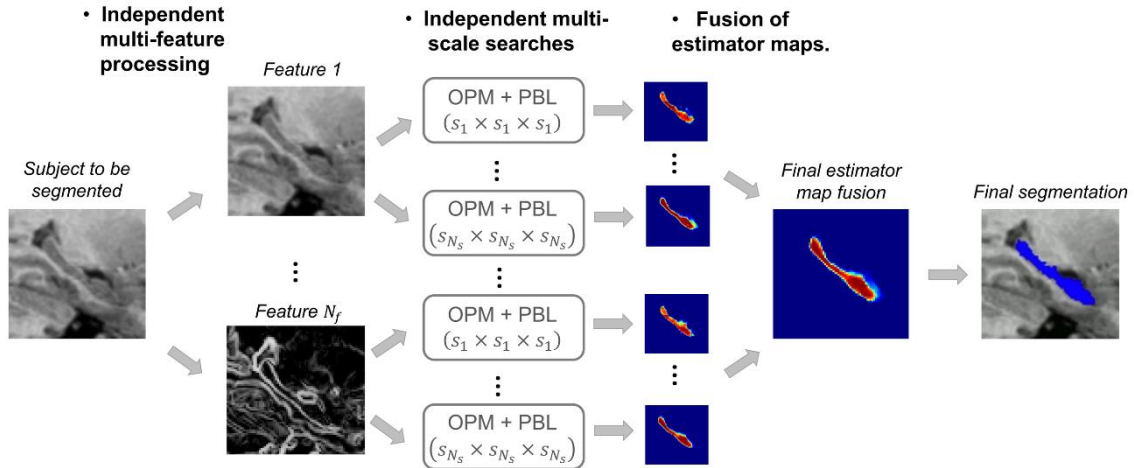


Figura 4.2.3: Esquema de pasos de OPAL. En este ejemplo, se usan imágenes y sus gradientes a distintas escalas y los mapas de probabilidad resultantes son promediados para obtener el mapa de probabilidad final.

### OPAL multiescala

En segmentación basada en parches el tamaño del parche es un parámetro clave que está fuertemente relacionado con la estructura de interés a segmentar y la resolución de la imagen. Este parámetro puede ser ajustado empíricamente para un conjunto de datos concreto. No obstante, obtener una solución de compromiso resulta complicado por lo que un enfoque multiescala con diferentes tamaños de parche proporciona mejores resultados como puede verse en la literatura (Manjón et al., 2014). En OPAL (Giraud et al., 2016) se computan simultáneamente diferentes escalas y características. En este trabajo se utilizan múltiples escalas pero únicamente se tienen en cuenta las intensidades de los vóxeles como característica. Como resultado se computan varios ANNFs para diferentes tamaños de parche, es decir, que tenemos varios mapas de probabilidad para cada etiqueta que se combinan mediante el proceso de *late fusion*. En el trabajo original de OPAL todos los mapas contribuyen con el mismo peso en el resultado final. En CERES hemos propuesto un enfoque adaptativo en función de la etiqueta:

$$p(l) = \alpha(l)p_1(l) + (1 - \alpha(l))p_2(l) \quad (4.1)$$

donde  $p_1(l)$  es el mapa de probabilidades de la escala 1 para la etiqueta  $l$ ,  $p_2(l)$  es el mapa de probabilidades de la escala 2 para la etiqueta  $l$ ,  $p(l)$  es el mapa de probabilidades final para la etiqueta  $l$  y  $\alpha(l) \in [0, 1]$  es el coeficiente de mezcla para la etiqueta  $l$  optimizado usando la técnica de *gradient descent*.

En este proceso de optimización se minimiza el error de segmentación (maximizando el coeficiente DICE). Dado que el conjunto de imágenes de entrenamiento es muy pequeño (5 sujetos que se traducen en 10 si usamos su versión en espejo) hemos utilizado el conjunto de validación usando una validación cruzada LOO. Con esto realizamos 20 experimentos excluyendo cada vez un sujeto diferente (y su versión en espejo).

#### 4.2 4. Regularización de etiquetas

Como consecuencia del reducido número de casos en la librería las segmentaciones producidas pueden presentar pequeñas irregularidades. Para mitigar este problema hemos propuesto un postproceso que regulariza las etiquetas mediante una convolución usando un *kernel* de 5x5x5 vóxeles. Este proceso penaliza las estructuras con bordes irregulares lo cual está de acuerdo con la definición de los lobulillos cerebelares. Para evitar que estructuras muy pequeñas desaparezcan, el mapa original de cada etiqueta se combina con el mapa convolucionado usando un coeficiente de mezcla  $\beta$ :

$$p_f(l) = p(l) + \beta(l) \cdot H \otimes p(l) \quad (4.2)$$

donde  $p_f(l)$  representa el peso final de la etiqueta  $l$ ,  $H$  es el *kernel* de convolución y  $\beta(l)$  es el coeficiente de mezcla para la etiqueta  $l$ . Los coeficientes de mezcla se optimizaron de forma ortogonal realizando una búsqueda dicotómica para encontrar su valor óptimo. Para ello utilizamos el conjunto de datos de validación realizando una validación cruzada LOO como la comentada anteriormente.

La figura 4.2.4 muestra un esquema general de todo el proceso de segmentación de CERES.

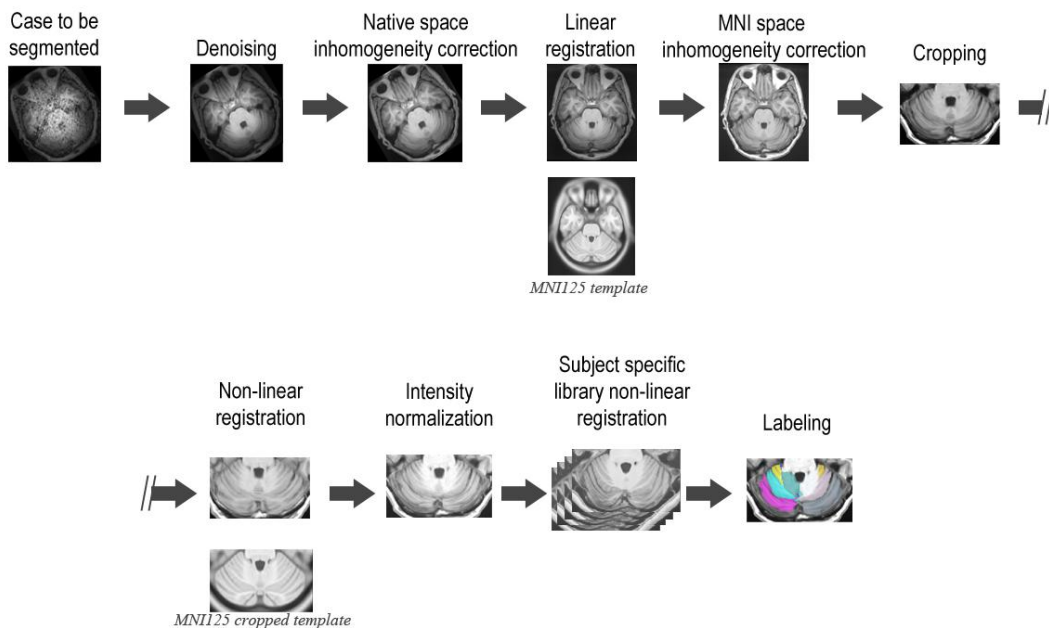


Figura 4.2.4: Esquema general del proceso de segmentación.

#### 4.2.5. Extensión de la librería.

Por último, como ya se ha mencionado, la librería de CERES contiene muy pocos casos segmentados a mano ( $N = 5$ ). Para reducir el impacto negativo que este hecho tiene sobre la segmentación realizamos una extensión de librería consistente en aplicar el proceso de segmentación antes descrito a 100 imágenes sin segmentar para después añadirlas junto con las segmentaciones a la librería como se ha hecho en otros trabajos previos (Eskildsen et al., 2012; Giraud et al., 2016). Estas 100 imágenes corresponden a sujetos sanos obtenidos del *IXI dataset* mencionado en el capítulo 3.

### 4.3. Experimentos y resultados

Se realizaron varios experimentos para estimar los parámetros óptimos del método y para comparar el rendimiento del método propuesto en términos de precisión y tiempo de cómputo con otros métodos relacionados del estado del arte.

#### 4.3.1. Parámetros del método

Puesto que OPAL (Giraud et al., 2016) fue desarrollado y optimizado para segmentar el hipocampo es necesario adaptar sus parámetros para obtener un resultado óptimo en su aplicación a la segmentación del cerebelo. Para medir la precisión del método hemos usado el índice DICE (Zijdenbos et al., 1994) para las 13 estructuras en las que dividimos el cerebelo (no se ha hecho distinción entre izquierda y derecha). Los resultados de los experimentos se han calculado sobre el conjunto de datos de validación. Nótese que debido a que en estos datos solo medio cerebelo ha sido segmentado a mano los resultados se obtuvieron comparando la segmentación con el área segmentada a mano. En todos los experimentos los tamaños de parche ha sido fijado a  $3 \times 3 \times 3$  y  $9 \times 9 \times 9$  vóxeles. El área de búsqueda restringida se fijó a  $7 \times 7 \times 7$  vóxeles. El número de Patch Matches en OPAL fue fijado en 32 y el número de iteraciones de OPAL fue fijado a 4.

#### 4.3.2. Construcción de la librería

Como ya se ha explicado, CERES utiliza una librería de casos segmentados manualmente a los que se les aplica un preprocesado. En este preprocesado las imágenes son registradas de forma lineal (afín) al espacio MNI152. Posteriormente son recortadas y registradas de nuevo para conseguir una mejor correspondencia anatómica entre sujetos. Este segundo registro tiene un gran impacto en la calidad de la segmentación (sobre todo debido al pequeño tamaño de la librería). Por tanto, hemos evaluado tres enfoques posibles. El primer enfoque consiste en realizar solo un registro lineal al espacio MNI recortado. En el segundo enfoque usamos un registro no lineal y en el tercer enfoque creamos una librería específica para cada caso a segmentar registrando no-linealmente los casos de la librería al caso a segmentar en el espacio MNI152 lineal. En el segundo enfoque la segmentación se realiza en el espacio destino mientras que en el tercero se realiza en el espacio MNI152 original. En la tabla 4.3.1 se pueden ver los resultados de los tres diferentes enfoques siendo la librería específica la mejor opción para todas las estructuras.

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

Como puede apreciarse, el registro no lineal mejora considerablemente los resultados. Este resultado tiene mucho sentido debido a este proceso minimiza el problema de tener una librería con pocos casos. Las diferencias entre la opción 2 y 3 se deben fundamentalmente a la minimización de los problemas de interpolación de las etiquetas. En este experimento, aplicamos el método OPAL sin ninguna de las optimizaciones propuestas para medir solo la influencia del enfoque de construcción de la biblioteca.

Tabla 4.3.1: DICE medio y desviación típica para la segmentación de CERES sobre el conjunto de validación. Las columnas muestran la media y la desviación típica para cada lobulillo y el cerebelo completo. Resultados calculados en el espacio nativo. Mejor resultado en negrita.

<i>Estructura</i>	<i>Reg. lineal</i>	<i>Reg. no-lineal</i>	<i>Lib. específica</i>
<b>PROMEDIO</b>	0.7207 ± 0.0574	0.7549 ± 0.0452	<b>0.7595 ± 0.0474</b>
<b>Lobulillo I-II</b>	0.4022 ± 0.2157	0.4165 ± 0.1993	<b>0.4226 ± 0.2110</b>
<b>Lobulillo III</b>	0.6644 ± 0.2039	0.7044 ± 0.1728	<b>0.7068 ± 0.1714</b>
<b>Lobulillo IV</b>	0.6879 ± 0.1625	0.7156 ± 0.1301	<b>0.7213 ± 0.1374</b>
<b>Lobulillo V</b>	0.7263 ± 0.1424	0.7448 ± 0.1315	<b>0.7514 ± 0.1376</b>
<b>Lobulillo VI</b>	0.8392 ± 0.0437	0.8620 ± 0.0322	<b>0.8667 ± 0.0322</b>
<b>Lobulillo Crus I</b>	0.8678 ± 0.0297	0.8937 ± 0.0174	<b>0.8970 ± 0.0161</b>
<b>Lobulillo Crus II</b>	0.7564 ± 0.0668	0.7973 ± 0.0545	<b>0.8024 ± 0.0553</b>
<b>Lobulillo VII B</b>	0.6076 ± 0.1354	0.6667 ± 0.1107	<b>0.6742 ± 0.1107</b>
<b>Lobulillo VIII A</b>	0.7277 ± 0.0705	0.7783 ± 0.0691	<b>0.7830 ± 0.0692</b>
<b>Lobulillo VIII B</b>	0.8012 ± 0.0541	0.8456 ± 0.0383	<b>0.8499 ± 0.0358</b>
<b>Lobulillo IX</b>	0.8623 ± 0.0383	0.8813 ± 0.0322	<b>0.8849 ± 0.0313</b>
<b>Lobulillo X</b>	0.7053 ± 0.0830	0.7529 ± 0.0438	<b>0.7540 ± 0.0473</b>
<b>Cerebelo</b>	0.9274 ± 0.0107	0.9346 ± 0.0091	<b>0.9358 ± 0.0091</b>

### 4.3.3. Multi-escala adaptativo

OPAL trabaja obteniendo correspondencias entre parches y lo hace a diferentes escalas que finalmente se combinan (*late fusion*). Hemos comparado el enfoque original de OPAL consistente en promediar uniformemente ambas escalas con nuestra propuesta de combinación adaptativa. La tabla 4.3.2 muestra los resultados. Nótese que los coeficientes de mezcla se optimizaron siguiendo una validación cruzada LOO por lo que se obtuvo un conjunto de coeficientes para cada caso del conjunto de imágenes de validación. Por simplicidad, en la tabla 4.3.3 se muestra un promedio de los coeficientes. No obstante, en los experimentos se utilizó el conjunto específico de cada caso para evitar que el caso bajo estudio influya en su propia segmentación.

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

Tabla 4.3.2: DICE promedio para la segmentación de CERES sobre el conjunto de validación. Las columnas muestran los valores promedio para los Lobulillos y el cerebelo completo... Mejores resultados en negrita.

<i>Estructura</i>	$\alpha = 0.5$	$\alpha$ adaptivo (LOOCV)
<b>PROMEDIO</b>	0.7595	<b>0.7640</b>
<i>Lobulillo I-II</i>	0.4226	<b>0.4364</b>
<i>Lobulillo III</i>	0.7068	<b>0.7237</b>
<i>Lobulillo IV</i>	0.7213	<b>0.7281</b>
<i>Lobulillo V</i>	0.7514	<b>0.7517</b>
<i>Lobulillo VI</i>	0.8667	<b>0.8673</b>
<i>Lobulillo Crus I</i>	0.8970	<b>0.9000</b>
<i>Lobulillo Crus II</i>	0.8024	<b>0.8034</b>
<i>Lobulillo VII B</i>	0.6742	<b>0.6756</b>
<i>Lobulillo VIII A</i>	0.7830	<b>0.7848</b>
<i>Lobulillo VIII B</i>	0.8499	<b>0.8506</b>
<i>Lobulillo IX</i>	0.8849	<b>0.8855</b>
<i>Lobulillo X</i>	0.7540	<b>0.7604</b>
<b>Cerebelo</b>	0.9358	<b>0.9370</b>

Tabla 4.3.3: Valores promedio de  $\alpha$  adaptativo.

<i>Estructura</i>	$\alpha$ coeficiente	<i>Estructura</i>	$\alpha$ coeficiente
<i>Lobulillo I-II</i>	0.82	<i>Lobulillo VII B</i>	0.59
<i>Lobulillo III</i>	0.81	<i>Lobulillo VIII A</i>	0.59
<i>Lobulillo IV</i>	0.71	<i>Lobulillo VIII B</i>	0.61
<i>Lobulillo V</i>	0.59	<i>Lobulillo IX</i>	0.60
<i>Lobulillo VI</i>	0.50	<i>Lobulillo X</i>	0.79
<i>Lobulillo Crus I</i>	0.81	<b>SB</b>	0.51
<i>Lobulillo Crus II</i>	0.60	<b>Fondo</b>	0.82

### 4.3.4. Regularización de etiquetas

El siguiente experimento tiene como objetivo estudiar el efecto de la regularización de etiquetas sobre la calidad de la segmentación. La tabla 4.3.4 muestra los resultados de la segmentación de CERES con y sin regularización sobre los 20 casos del conjunto de validación. Nótese que los coeficientes de regularización se optimizaron siguiendo una validación cruzada LOO por lo que se obtuvo un conjunto de coeficientes para cada caso del conjunto de imágenes de validación. En la tabla 4.3.4 se muestra el DICE promedio así como la distancia Hausdorff (Beauchemin et al., 1998) para cada estructura y el cerebelo completo. Pese a que las diferencias no son significativas se observa que la regularización tiende a producir mejores índices DICE y distancias de Hausdorff menores.

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

Tabla 4.3.4: DICE medio y distancia de Hausdorff de CERES sobre el conjunto de validación. Las columnas muestran los valores promedio para cada estructura y el cerebelo completo. Los valores se calcularon en el espacio nativo. Mejores resultados en negrita.

Estructura	Sin regularización		Con regularización	
	DICE	Hausdorff	DICE	Hausdorff
<b>PROMEDIO</b>	0.7640	6.3992	<b>0.7702</b>	<b>6.3302</b>
<b>Lobulillo I-II</b>	0.4364	3.2214	<b>0.5056</b>	<b>2.8571</b>
<b>Lobulillo III</b>	<b>0.7237</b>	6.2210	0.7157	<b>6.0854</b>
<b>Lobulillo IV</b>	0.7281	5.5987	<b>0.7320</b>	<b>5.4904</b>
<b>Lobulillo V</b>	0.7517	6.7155	<b>0.7535</b>	<b>6.7022</b>
<b>Lobulillo VI</b>	0.8673	6.1831	0.8678	<b>6.1557</b>
<b>Lobulillo Crus I</b>	0.9000	6.5719	<b>0.9005</b>	<b>6.5583</b>
<b>Lobulillo Crus II</b>	0.8034	10.9254	<b>0.8043</b>	<b>10.8393</b>
<b>Lobulillo VIIB</b>	0.6756	<b>9.8729</b>	<b>0.6778</b>	9.9934
<b>Lobulillo VIIIA</b>	0.7848	8.4833	<b>0.7862</b>	<b>8.3133</b>
<b>Lobulillo VIIIB</b>	0.8506	4.7480	<b>0.8531</b>	<b>4.7222</b>
<b>Lobulillo IX</b>	0.8855	<b>5.0390</b>	<b>0.8858</b>	5.0499
<b>Lobulillo X</b>	<b>0.7604</b>	3.2097	0.7603	<b>3.1952</b>
<b>Cerebelo</b>	0.9370	-	<b>0.9372</b>	-

### 4.3.5. Extensión de la librería

En este experimento hemos evaluado el impacto del número de sujetos de la librería extendida usados para producir la segmentación. La librería extendida incluye los 10 casos manualmente segmentados y 100 casos segmentados automáticamente. En cada ronda se seleccionan los N casos más similares al caso segmentar de entre los 100 segmentados automáticamente utilizando para ello el error cuadrático medio. En la figura 4.3.1 se puede observar que se alcanza un máximo cuando N=50 casos. También se observa en la gráfica una importante oscilación para valores por encima de 50. Esto se explica debido al funcionamiento de PatchMatch el cual implica una búsqueda aleatoria que se repite de forma iterativa. El proceso converge muy rápido ya que no explora toda la librería y por tanto obtenemos un resultado muy próximo al óptimo (pudiendo serlo o no). Esto implica que para diferentes tamaños de librería el algoritmo la explora de forma diferente aunque la mayoría de casos sean los mismos. Por esta razón, la calidad oscila una vez superado el valor óptimo (en este caso 50 casos). El uso de la extensión de librería permitió incrementar el resultado de un DICE promedio de 0.7702 a 0.7729. Por tanto fijamos N=50 para el resto de experimentos.

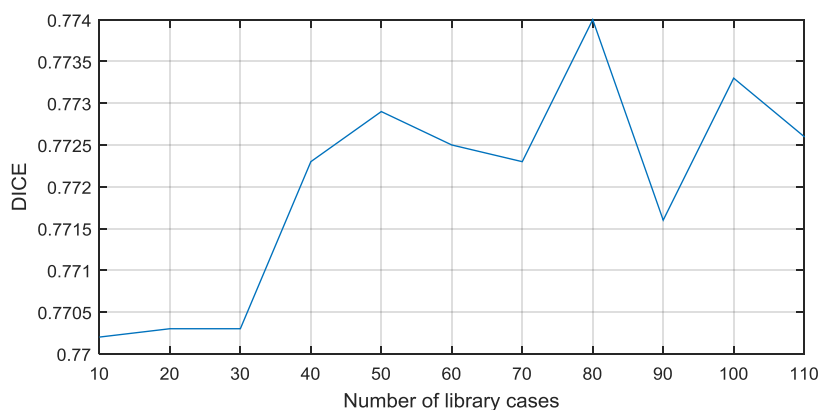


Figura 4.3.1: Evolución del DICE promedio en función del número de casos tomados de la librería extendida.



### 4.3.6. Comparativa de métodos

Tras todas las optimizaciones y la extensión de la librería hemos comparado nuestro método CERES con otros 3 métodos representativos del estado del arte: SUIT (Diedrichsen, 2006), MAGeT (Chakravarty et al., 2013; Park et al., 2014) y RASCAL (Weier et al., 2014). Los métodos fueron aplicados al conjunto de validación para evaluar su rendimiento. En la tabla 4.3.5 se compara el índice DICE promedio obtenido por tres de los cuatro métodos para las 12 estructuras y el cerebelo completo. Como puede observarse, CERES supera a ambos métodos a excepción del lobulillo VI para el cual MAGeT obtiene un valor ligeramente superior. La significación estadística se evaluó utilizando un test de Wilcoxon no paramétrico debido a que los valores de DICE obtenidos no siguen una distribución normal.

Tabla 4.3.5: DICE promedio y desviación típica para MAGeT, RASCAL y CERES. La última columna suena la precisión intra-observador como referencia de la calidad alcanzable por un humano. Mejores resultados en negrita. Las diferencias significativas ( $p < 0,05$ ) entre CERES y MAGeT están marcadas con \* y las diferencias entre RASCAL y CERES con †.

Estructura	MAGeT	RASCAL	CERES	Intraobservador
<b>Promedio</b>	0.7320 ± 0.0568	0.6890 ± 0.0524	<b>0.7729 ± 0.0427†</b>	0.833
<b>Lobulillo I-II</b>	0.3960 ± 0.1424	0.3260 ± 0.2178	<b>0.5201 ± 0.1555*†</b>	0.639
<b>Lobulillo III</b>	0.6800 ± 0.1741	0.6379 ± 0.2165	<b>0.7213 ± 0.1572</b>	0.751
<b>Lobulillo IV</b>	0.6980 ± 0.1440	0.6627 ± 0.1611	<b>0.7271 ± 0.1346</b>	0.818
<b>Lobulillo V</b>	0.7320 ± 0.1398	0.6666 ± 0.1560	<b>0.7561 ± 0.1332†</b>	0.881
<b>Lobulillo VI</b>	<b>0.8710 ± 0.0359</b>	0.7969 ± 0.0523	0.8695 ± 0.0316†	0.912
<b>Lobulillo Crus I</b>	0.8870 ± 0.0257	0.8383 ± 0.0351	<b>0.9007 ± 0.0152†</b>	0.904
<b>Lobulillo Crus II</b>	0.7780 ± 0.0679	0.7340 ± 0.0667	<b>0.8096 ± 0.0569†</b>	0.900
<b>Lobulillo VII B</b>	0.5990 ± 0.1487	0.5820 ± 0.1137	<b>0.6850 ± 0.1205*†</b>	0.863
<b>Lobulillo VIII A</b>	0.7300 ± 0.0934	0.6757 ± 0.1426	<b>0.7926 ± 0.0759*†</b>	0.860
<b>Lobulillo VIII B</b>	0.7970 ± 0.0607	0.7783 ± 0.0931	<b>0.8533 ± 0.0390*†</b>	0.833
<b>Lobulillo IX</b>	0.8560 ± 0.0384	0.8460 ± 0.0545	<b>0.8849 ± 0.0327†</b>	0.874
<b>Lobulillo X</b>	0.7540 ± 0.0490	0.7237 ± 0.0680	<b>0.7548 ± 0.0469*</b>	0.760
<b>Cerebelo</b>	0.9250 ± 0.0094	0.9349 ± 0.0089	<b>0.9377 ± 0.0090†</b>	0.941

Puesto que SUIT divide el cerebelo en un conjunto diferente de estructuras hemos tenido que adaptar las segmentaciones de MAGeT, RASCAL y CERES para poder compararlas. Para ello hemos fusionado los lobulillos I-II, III y IV como se hizo en (Park et al., 2014). Los lobulillos Crus I y IX fueron excluidos de la comparativa debido a diferencias en su definición.

Hemos procesado las imágenes del conjunto de validación usando SUIT como se indica en su página web ([http://www.icn.ucl.ac.uk/motorcontrol/imaging/suit\\_function.htm](http://www.icn.ucl.ac.uk/motorcontrol/imaging/suit_function.htm)) sin aplicar ninguna corrección manual. En la tabla 4.3.6 se muestra el índice DICE promedio de los 4 métodos para las diferentes estructuras. CERES obtiene el mejor resultado en promedio mientras que MAGeT segmenta mejor los

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

lobulillos I-IV y X pese a que no hay diferencias significativas. La figura 4.3.2 muestra algunos ejemplos de segmentación de los cuatro métodos.

Tabla 4.3.6: DICE promedio y desviación típica para SUIT, MAGeT, RASCAL y CERES considerando el protocolo de segmentación de SUIT. Mejores resultados en negrita. Las diferencias significativas ( $p < 0,05$ ) entre SUIT y CERES marcadas con \*, entre MAGeT y CERES con † y entre RASCAL y CERES con ‡.

Estructura	SUIT	MAGeT	RASCAL	CERES
<b>Promedio</b>	0.7097 ± 0.0689	0.7639 ± 0.0792	0.7140 ± 0.0487	<b>0.7888 ± 0.0409 *‡</b>
<b>Lobulillo I-IV</b>	0.7435 ± 0.0880	<b>0.8055 ± 0.0964</b>	0.7703 ± 0.1107	0.7898 ± 0.1021*
<b>Lobulillo V</b>	0.6598 ± 0.1026	0.7429 ± 0.1369	0.6730 ± 0.1560	<b>0.7561 ± 0.1332 *‡</b>
<b>Lobulillo VI</b>	0.7800 ± 0.0543	<b>0.8762 ± 0.0365</b>	0.7994 ± 0.0523	0.8695 ± 0.0316 †‡
<b>Lobulillo Crus II</b>	0.7430 ± 0.0631	0.7787 ± 0.0678	0.7300 ± 0.0667	<b>0.8096 ± 0.0569 *‡</b>
<b>Lobulillo VII B</b>	0.5701 ± 0.1572	0.6013 ± 0.1476	0.5761 ± 0.1137	<b>0.6850 ± 0.1205 *‡</b>
<b>Lobulillo VIIIA</b>	0.7134 ± 0.0996	0.7330 ± 0.0928	0.6701 ± 0.1426	<b>0.7926 ± 0.0759 *†‡</b>
<b>Lobulillo VIIIB</b>	0.7721 ± 0.0596	0.8012 ± 0.0607	0.7654 ± 0.0931	<b>0.8533 ± 0.0390 *†‡</b>
<b>Lobulillo X</b>	0.6955 ± 0.0512	<b>0.7721 ± 0.0475</b>	0.7275 ± 0.0680	0.7548 ± 0.0469*

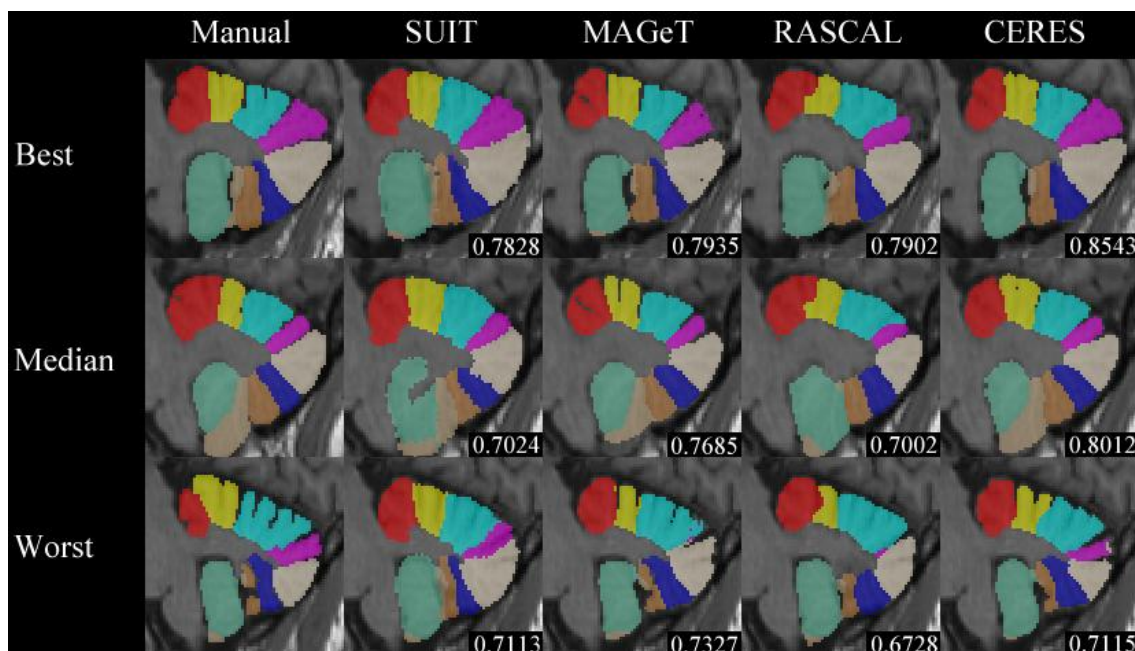


Figura 4.3.2: Ejemplos de segmentación manual y automática usando SUIT, MAGeT, RASCAL y CERES sobre los casos con mejor, mediano y peor DICE. El índice DICE de cada caso está indicado junto a la captura. Nótese como SUIT tiende a sobre-segmentar y siempre rellena las fisuras entre lobulillos. También se puede observar que RASCAL produce una frontera irregular entre los lobulillos I-IV (rojo) y el V (amarillo) en los tres casos. Finalmente, en el peor caso se puede ver que para MAGeT y RASCAL el lobulillo VIIA (morado) es apenas visible.

### 4.3.7. Tiempo de ejecución

Una de las ventajas de CERES es su eficiencia puesto que produce resultados competitivos en términos de precisión con un tiempo de ejecución reducido. El método más lento de la comparativa fue MAGeT (aproximadamente 4 horas por caso) debido a la gran cantidad de registros no-lineales que requiere.

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

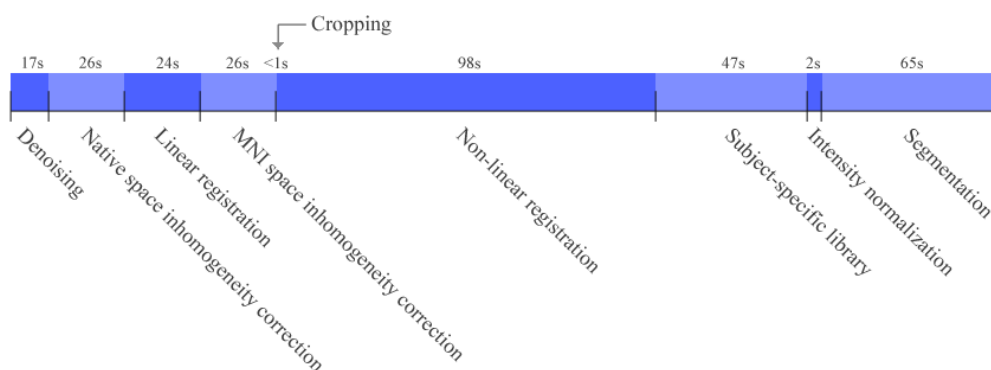


Fig. 4.3.3. Tiempo de ejecución en segundos para cada una de las partes del método propuesto.

RASCAL necesita en torno a 110 minutos para genera la segmentación. Finalmente SUIT (sin realizar correcciones a mano) y CERES finalizan en aproximadamente 5 minutos. Todos los experimentos incluida la medición del tiempo de cómputo fueron realizados en una maquina con un procesador Intel core i7 y 8 GB de memoria RAM ejecutando MATLAB 2014a en un sistema Windows 7 Profesional 64 bit.

### 4.4. Conclusión y discusión

Hemos presentado un método para segmentar los lobulillos del cerebelo al cual hemos llamado CERES que funciona de forma totalmente automática y es capaz de producir segmentaciones precisas con un bajo coste computacional. Hemos mostrado que CERES produce resultados al nivel del estado del arte superando a métodos actuales como SUIT, MAGeT y RASCAL. Las contribuciones principales de este capítulo son: 1) el uso de una librería específica para cada caso a segmentar, 2) una implementación multiescala adaptativa del método OPAL y 3) una regularización de etiquetas.

En términos de precisión, CERES obtuvo los mejores resultados para la mayoría de las estructuras. Cada una de las mejoras propuestas en el método permitió mejorar consistentemente los resultados siendo la librería específica el que tuvo mayor impacto (pasando de 0.7207 a 0.7595). El mezclado de mapas de probabilidad multiescala adaptativo y la regularización de etiquetas consiguieron aumentar la precisión hasta un DICE de 0.7729. Resulta muy interesante que CERES genere segmentaciones con una precisión cercana a la del intra-observador para el cerebelo completo (CERES = 0.9377 vs. humano = 0.941) (Park et al., 2014). Se debe mencionar que esta alta precisión obtenida por CERES se consigue con tan solo 5 casos segmentados a mano (10 si contamos con su versión en espejo). Se espera que los resultados con un número mayor de casos segmentados a mano sean significativamente mejores pudiendo representar mejor la variabilidad anatómica del cerebelo.

Desde el punto de vista de la eficiencia, CERES produce la segmentación en un tiempo muy reducido (5 minutos por caso) siendo el más eficiente de los método comparados. Esto se explica por dos razones principales. En primer lugar, los métodos

## Capítulo 4: Segmentación de los lóbulos del cerebelo (CERES)

que se basan en MALF suelen requerir el cómputo de múltiples registros no lineales. CERES solo requiere estimar una transformación no-lineal ya que las transformaciones asociadas a los casos de la librería se computan *offline*. Además este registro no-lineal es de baja dimensionalidad y se aplica un volumen recortado con lo que requiere menos tiempo de lo habitual. En segundo lugar una de las partes con mayor carga computacional de los métodos MALF es el proceso de encontrar correspondencias entre parches. En CERES, este tiempo se ve muy reducido gracias al uso de OPAL haciendo que establecer correspondencias entre parches ya no sea un cuello de botella en el proceso de segmentación.

Otra característica importante de CERES es su adaptabilidad. Puesto que utiliza una librería de casos segmentados a mano como base de conocimiento puede aprender nuevas anatomías simplemente añadiendo casos a la librería. Como se ha visto en los experimentos, la mayor limitación de CERES es el reducido tamaño de su librería (5 casos) ya que otros métodos *multi-atlas* encuentran resultados óptimos utilizando en torno a 20 casos. Por tanto, es de esperar que CERES mejore con una librería más extensa lo cual sucederá a medida que más casos etiquetados a mano se hagan disponibles.

Finalmente, para hacer CERES accesible a la comunidad científica se ha integrado en nuestra plataforma online de acceso gratuito volBrain (<http://volbrain.upv.es>). En el momento de redactar este documento CERES ya ha procesado más de 15.000 casos con una tasa de error cercana al 1%, esto es, casos en los que la plataforma terminó prematuramente no pudiendo completar la segmentación.

Para finalizar nos parece importante comentar que recientemente, una versión mejorada de CERES (CERES2) participó en el *challenge* ENIGMA de la conferencia MICCAI2017 en Quebec (Canadá). En este *challenge*, se compararon 9 métodos distintos de segmentación de lóbulos del cerebelo (incluyendo algunos métodos del estado del arte basados en *Deep learning*). CERES2 ganó el *challenge* de forma clara, quedando en primera posición en los dos *datasets* utilizados. Actualmente, se ha preparado una publicación con los resultados que ha sido enviada a la revista Neuroimage (Carass et al., 2018). CERES2 introduce una nueva normalización en intensidad basada en *histogram matching* y un corrector de errores sistemáticos basado en redes neuronales. Esta nueva versión de CERES se incorporará a nuestro sistema online volBrain (<http://volBrain.upv.es>) para hacerla accesible a la comunidad científica en breve.

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

En este capítulo se presenta un nuevo método para la segmentación de las subregiones del hipocampo usando imágenes de Resonancia Magnética. Este método fue publicado en 2017 (Romero et al., 2017b) en la revista Neuroimage y se está integrando actualmente dentro del sistema online volBrain (<http://volbrain.upv.es>).

### 5.1. Segmentación del hipocampo

El hipocampo (HC) es una pequeña estructura cerebral bilateral situada en el lóbulo temporal medio a ambos lados del bulbo raquídeo, cerca del cerebelo. Su nombre proviene de su parecido con el caballo de mar. Si recorremos el hipocampo desde su extremo superior en el surco hipocámpal encontramos el *dentate gyrus* (DG) seguido por el *Cornu Ammonis* (CA) el cual está subdividido en tramos consecutivos (del CA4 a CA1) llegando finalmente al *Subiculum* en el extremo inferior. El CA está a su vez dividido en seis capas llamadas *strata*. Estas capas son el *Stratum Oriens* (SO), *Stratum Pyramidale* (SP), *Stratum LUCidum* (SLU), *Stratum Radiatum* (SR), *Stratum Lacunosum* (SL) y *Stratum Moleculare* (SM). En la figura 5.1.1 puede verse una representación gráfica del hipocampo.

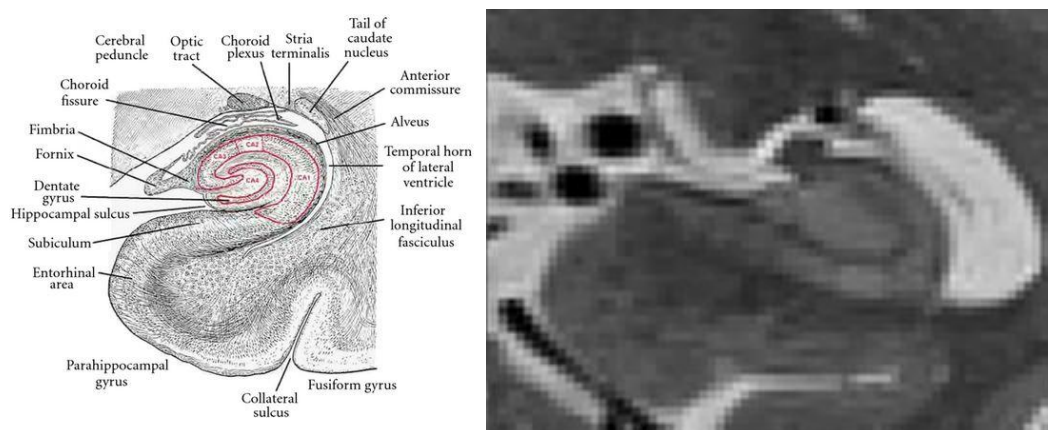


Figura 5.1.1: De izquierda a derecha: Anatomía del hipocampo y una imagen de ejemplo de RMN (T2) de alta resolución.

El hipocampo está implicado en varias funciones cerebrales como la memoria y el razonamiento espacial (Milner, 1958; Schmajuk, 1990). Varios estudios muestran que posee un papel muy importante en varias enfermedades neurodegenerativas como el Alzheimer (AD) (Braak and Braak, 1991) o la esquizofrenia (Altshuler et al., 1998). Por esto, el estudio del volumen del hipocampo es de gran interés ya que es una valiosa herramienta para el seguimiento de enfermedades y el ajuste de tratamientos (Dickerson and Sperling, 2005; Jack et al., 2005, 2000). Sin embargo, la anatomía del

hipocampo es compleja y variable y los límites entre subestructuras han sido descritos en la literatura utilizando características citoarquitectónicas que requieren de tinción histológica y resolución microscópica para visualizarlas (Insausti and Amaral, 2012).

Dada la importancia del hipocampo, varios protocolos y métodos de segmentación han sido desarrollados (Barnes et al., 2008; Collins and Pruessner, 2010; Coupé et al., 2011). Sin embargo, uno de los principales problemas en el avance de este campo es la disparidad de las definiciones del HC y la falta de recursos como por ejemplo casos segmentados a mano. Recientemente, se ha propuesto un protocolo armonizado para el hipocampo completo (junto con 120 casos segmentados a mano a una resolución de 1 mm<sup>3</sup>) el cual será un referente para el desarrollo y la comparación de nuevos métodos de segmentación (Boccardi et al., 2015).

Hasta ahora, debido a las limitaciones en la resolución de las imágenes de RM, los estudios se veían restringidos a considerar el hipocampo como una sola estructura (Chupin et al., 2009). Aunque se haya demostrado que el estudio del hipocampo completo es una buena aproximación para el estudio del AD, algunos estudios ex-vivo revelan que el envejecimiento normal y el AD afectan a las subregiones de forma diferente durante la vida del paciente (Braak and Braak, 1991).

En la actualidad, muchos protocolos de segmentación de subregiones del HC han sido desarrollados como respuesta a los avances en las secuencias de RMN que permiten adquirir imágenes de alta resolución haciendo posible dividir el hipocampo en sus partes constituyentes. Sin embargo, todavía no se ha llegado al consenso entre los diferentes protocolos como se muestra en (Yushkevich et al., 2015a) donde 21 protocolos de segmentación manual son comparados. Algunos de estos protocolos han sido utilizados para crear conjuntos de imágenes segmentadas a mano que son un recurso fundamental para el desarrollo de métodos de segmentación.

Por ejemplo, imágenes adquiridas en un equipo de 9.4 Teslas a ultra alta resolución ex-vivo se usaron para crear, mediante segmentación manual, un atlas anatómico (Yushkevich et al., 2009) incluyendo el CA1, CA2-3, DG y el surco hipocampal. En 2013, Winterburn presentó un nuevo atlas in-vivo de alta resolución (Winterburn et al., 2013) que dividía el hipocampo en cinco subregiones: CA1, CA2-3, CA4/DG, *Stratum* y *Subiculum* (junto con 5 casos segmentados manualmente, *dataset* de Winterburn). Más tarde, en 2015, Kulaga-Yoskovitz desarrolló otro protocolo de segmentación (Kulaga-Yoskovitz et al., 2015) consistente en tres estructuras: CA1-3, CA4/DG y *Subiculum* (junto con 25 casos segmentados a mano, *dataset* de Kulaga-Yoskovitz).

Para realizar estudios volumétricos y aplicar estos protocolos se hace necesario el uso de herramientas de segmentación automática. Es bien conocido que la segmentación manual de casos de estudio representa un problema en términos de reproducibilidad. Además, es un trabajo extremadamente tedioso que requiere de mucho tiempo y un alto coste económico (puede llevar hasta 50 horas para un experto segmentar un solo hipocampo (Iglesias et al., 2015)). Dado que la segmentación manual no es una opción asequible, varios métodos automáticos han sido desarrollados en los últimos años. Uno de los primeros métodos de segmentación de subregiones del hipocampo fue propuesto por Van Leemput (Van Leemput et al., 2009) utilizando un modelo generativo de la región del hipocampo. Este modelo se genera utilizando un atlas

probabilístico basado en mallas que contiene información sobre donde es más probable que se encuentren las estructuras anatómicas. Este atlas probabilístico se genera a partir de un conjunto de imágenes de ultra alta resolución. Recientemente, Iglesias ([Iglesias et al., 2015](#)) continuó este trabajo y mejoró el modelo utilizando un atlas más preciso generado a partir de imágenes ex-vivo de ultra-alta resolución así como varias modalidades de imagen (multi-contraste).

Pipitone propuso un método basado en MALF ([Pipitone et al., 2014](#)) utilizando imágenes potenciadas en T2 dirigido a segmentar un conjunto de datos considerablemente grande utilizando unos pocos casos segmentados manualmente. Este método es una adaptación de MAGeT ([Chakravarty et al., 2013](#)) que como vimos en el capítulo 4 tiene un alto coste temporal. En 2015, Yushkevich propuso otro método ([Yushkevich et al., 2015b](#)) empleando también imágenes potenciadas en T2 donde combinaba un enfoque *multi-atlas* con una fusión de etiquetas basada en un voto de la mayoría ponderado por similitud. También utilizó un método de postprocesado de corrección de errores sistemáticos. Por desgracia, este método requiere horas para producir una segmentación debido al uso exhaustivo de registros no-lineales como en el caso de MAGeT.

Recientemente, en 2016, Caldairou presentó un método híbrido ([Caldairou et al., 2016](#)), en el cual, un conjunto de datos de entrenamiento eran registrados no linealmente al caso a segmentar. Entonces, utilizando correspondencias de parches genera un modelo tridimensional basado en mallas a partir de las segmentaciones manuales. Estas correspondencias entre parches se recomputan para cada vértice de la malla minimizando el error para ajustar el modelo deformable al caso a segmentar.

En este capítulo presentamos un nuevo método de segmentación basado en MALF que ha sido validado con dos de los protocolos de segmentación de subregiones de hipocampo disponibles públicamente. El método propuesto utiliza una adaptación de MOPAL ([Romero et al., 2017a](#)), una versión multi-contraste del método de *patch matching* OPAL ([Giraud et al., 2016](#)), para producir segmentaciones rápidas y precisas. El método presentado en este capítulo utiliza imágenes de alta resolución ( $0,5 \times 0,5 \times 0,5 \text{ mm}^3$ ) potenciadas en T1 y en T2. Además, también funciona con imágenes de resolución estándar (tras un proceso de superresolución ([Coupé et al., 2013](#); [Manjón et al., 2010a](#))) lo cual amplía su utilidad a datos retrospectivos. Como paso final, el método propuesto, también incluye un post-proceso para corregir errores sistemáticos basado en redes neuronales.

## 5.2. Materiales y métodos

En las siguientes secciones, se describen los datos utilizados en el desarrollo del método así como los detalles del método propuesto.

### 5.2.1. Datos

En este capítulo hemos usado dos conjuntos de datos diferentes que se corresponden con dos protocolos de segmentación de hipocampo diferentes. Ambos consisten en imágenes de alta resolución potenciadas en T1 y T2. La figura 5.2.1 muestra un ejemplo de estas imágenes.

#### *Dataset de Winterburn*

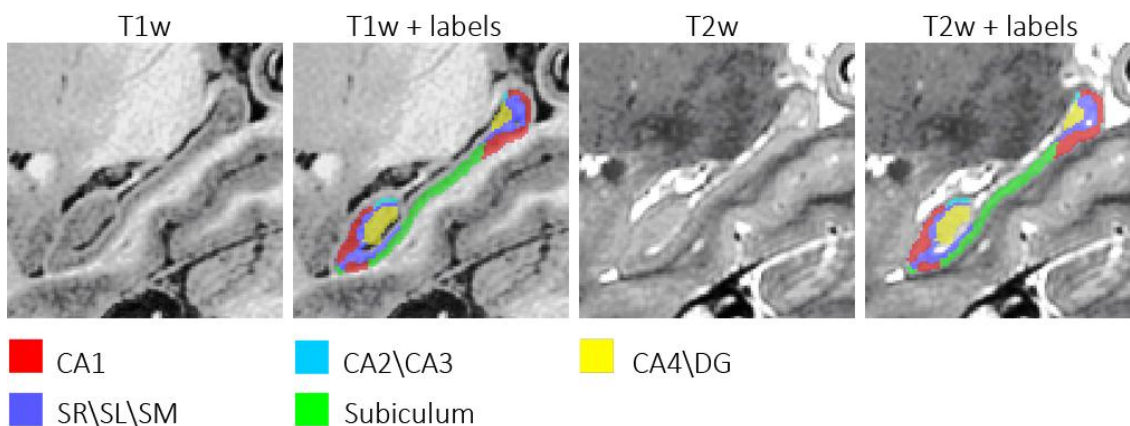
Este conjunto de imágenes contiene 5 sujetos a una resolución de  $0,3 \times 0,3 \times 0,3 \text{ mm}^3$  con secuencias potenciadas en T1 y en T2 obtenidas mediante una doble interpolación de adquisiciones a  $0,6 \times 0,6 \times 0,6 \text{ mm}^3$  y sus correspondientes segmentaciones manuales. Las imágenes pueden obtenerse libremente en el sitio web de CoBrALab (<http://cobralab.ca/atlasses>). Las imágenes fueron tomadas de 5 voluntarios sanos (2 hombres y 3 mujeres de entre 29 y 57 años). Las imágenes T1 fueron adquiridas usando una secuencia *3D inversion-prepared fast spoiled gradient-recalled echo* (TE/TR=4,3 ms / 9,2 ms, TI=650ms,  $\alpha=8^\circ$ , 2-NEX y una resolución isotrópica de  $0,6 \times 0,6 \times 0,6 \text{ mm}^3$ ). Las imágenes T2 fueron adquiridas usando una secuencia *3D fast spin echo*, FSE-CUBE (TE/TR=95,3ms / 2500 ms, ETL=100 ms, 2NEX, y una resolución isotrópica de  $0,6 \text{ mm}^3$ ). Se usaron los filtros de reconstrucción ZIPX2 y ZIP512 para obtener una resolución final de  $0,3 \times 0,3 \times 0,3 \text{ mm}^3$ . El hipocampo y sus subregiones fueron segmentados manualmente por un experto en 5 regiones (CA1, CA2-3, CA4/DG, SR/SL/SM y *Subiculum*). Los detalles sobre el protocolo de segmentación se pueden encontrar en el artículo original ([Winterburn et al., 2013](#)).

#### *Dataset de Kulaga-Yoskovitz*

Este conjunto de imágenes incluye 25 sujetos de un repositorio público (<http://www.nitrc.org/projects/mni-hisub25>) con un rango de edad de  $31 \pm 7$ , 12 hombres y 13 mujeres, con segmentaciones manuales que dividen el hipocampo en tres estructuras (CA1-3, CA4/DG y *Subiculum*). Las imágenes consisten en secuencias MPRAGE potenciadas en T1 ( $0,6 \text{ mm}^3$  isotrópico) y TSE potenciadas en T2 ( $0,4 \times 0,4 \times 2,0 \text{ mm}^3$  anisotrópico). Las imágenes fueron preprocesadas para corregir inhomogeneidades, normalizadas en intensidad y linealmente registradas al espacio MNI152. Ambas potenciaciones en T1 y T2 fueron remuestreadas a una resolución de  $0,4 \times 0,4 \times 0,4 \text{ mm}^3$ . Para minimizar los artefactos de interpolación las imágenes T2 fueron superresueltas utilizando un método de superresolución ([Manjón et al., 2010a](#)). Los detalles sobre el protocolo de segmentación se describen con detalle en la publicación original ([Kulaga-Yoskovitz et al., 2015](#)).



### Winterburn



### Kulaga-Yoskovitz

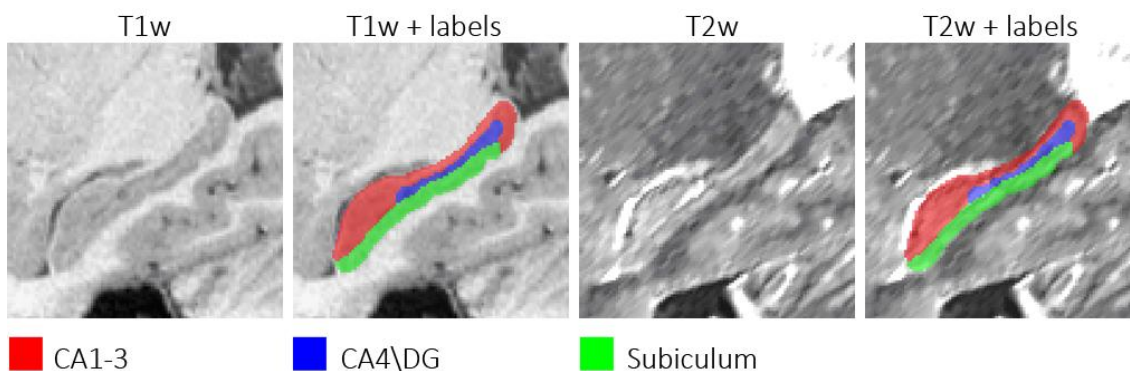


Figura 5.2.1: Ejemplo de los datasets de Kulaga-Yoskovitz y Winterburn. Pueden verse las imágenes T1 y T2 así como sus segmentaciones manuales.

### 5.2.2. Preprocesado

Todas las imágenes fueron preprocesadas para situarlas en un espacio de coordenadas e intensidades común. Para ello se aplicaron los siguientes pasos. Filtrado usando el *Spatially Adaptive Non-local Means Filter* (Manjón et al., 2010b). Corrección de inhomogeneidad usando el método N4 (Tustison et al., 2010). Las imágenes fueron registradas linealmente al espacio MNI152 utilizando ANTs (Avants et al., 2009). Esta transformación fue estimada utilizando la plantilla MNI152 con las imágenes T1 de los *datasets* y fue aplicada tanto a las imágenes T1 como T2. Previo a este registro se estimó una transformación rígida para alinear cada imagen T2 a su correspondiente T1 que después se concatenó con el registro lineal para realizar ambos registros con una única interpolación. Nótese que se procesan imágenes de alta resolución (tamaño de vóxel inferior a  $1 \text{ mm}^3$ ) y se utiliza una versión de alta resolución de la plantilla MNI152 ( $0,5 \times 0,5 \times 0,5 \text{ mm}^3$ ) por lo que la segmentación se realiza siempre a una resolución de  $0,5 \times 0,5 \times 0,5 \text{ mm}^3$ .

Las imágenes fueron normalizadas en intensidad para que los tejidos cerebrales de los diferentes sujetos tuvieran niveles de intensidad similares. Para ello se aplicó la técnica de Nyúl y Udupa (Nyúl and Udupa, 1999). Seguidamente, para reducir el consumo de memoria y el tiempo de ejecución las imágenes fueron recortadas en

torno al área que ocupa el HC. Para esto se calcularon los límites en torno al HC en el espacio MNI152 a partir de la suma de las segmentaciones manuales con un margen de 5 vóxeles en cada dirección.

Nótese que cuando se trabaja con un caso a segmentar cuya resolución es inferior a  $0,5 \times 0,5 \times 0,5 \text{ mm}^3$ , una vez recortado debe ser superresuelto para alcanzar la resolución a la que trabaja el método propuesto. Para ello utilizamos el método de superresolución LASR (Coupé et al., 2013).

Para conseguir una mejor correspondencia anatómica entre sujetos se estimó el registro no lineal de las imágenes recortadas a la plantilla MNI152 también recortada. Para ello se estimó un registro multi-contraste utilizando T1 y T2 teniendo ambas imágenes el mismo peso. El registro se estimó utilizando ANTs (Avants et al., 2009) utilizando la métrica de correlación cruzada y un enfoque piramidal a escalas 8x, 4x, 2x, y 1x con 200, 200, 200 y 0 iteraciones en cada escala respectivamente. Este registro introduce un pequeño emborronamiento en las imágenes debido a la interpolación aplicada lo cual tiene un impacto negativo en la segmentación. Para mitigar este efecto y mejorar las imágenes se aplica un filtro de afilado o *sharpening* sumándole a cada imagen su laplaciano.

### 5.2.3. Construcción de la librería

El método propuesto requiere de la construcción de una librería de sujetos segmentados a mano situados en el mismo espacio geométrico y de intensidad que el caso a segmentar. Para ello se ha construido una librería aplicando el preprocesado antes descrito a todas las imágenes de ambos *datasets* (incluyendo el registro de la etiquetas al espacio MNI152 recortado). Para incrementar el número de casos de la librería, las imágenes fueron invertidas de izquierda a derecha con lo que se dobló el número de casos de la librería.

Para segmentar un nuevo caso, este es preprocesado del mismo modo que los casos de la librería. Una vez terminado el preprocesado el resultado es un conjunto de imágenes recortadas y sus transformaciones no lineales al espacio MNI152 (y sus segmentaciones manuales) y un caso a segmentar recortado con su transformación no-lineal al espacio MNI152 también. A partir de estas transformaciones se genera una librería específica para el caso a segmentar. A cada imagen de la librería se le aplica su transformación al espacio MNI152 concatenada con la transformación inversa del caso a segmentar con lo que finalmente cada caso de la librería se registra no-linealmente al espacio del caso a segmentar tal como se ha descrito en el capítulo anterior en el método CERES (Romero et al., 2017a).

### 5.2.4. Multispectral Optimized PatchMatch (MOPAL)

El método de segmentación propuesto está basado en una modificación del método OPAL (Giraud et al., 2016) ya explicado en el capítulo 4 que consiste en la utilización de una nueva métrica que le permite tomar información de dos canales o características que en este caso son las imágenes potenciadas en T1 y en T2 lo cual se ha demostrado beneficioso para el proceso de *patch matching* (B. Fisher and Oliver, 1998; Xiao et al., 2015).

Como ya se ha explicado, OPAL utiliza PatchMatch para buscar correspondencias entre parches de la imagen a segmentar y la librería. Estas correspondencias se evalúan mediante una métrica basada en las intensidades de los parches. En la versión clásica de OPAL la métrica consistía en una suma de diferencias al cuadrado o SSD de sus siglas en inglés (*Sum of Squared Differences*) como se muestra en la ecuación 5.1 (5.2 para su versión multicanal). Nuestra propuesta es una métrica basada una semi-norma SSD (una SSD por canal) que tiene en cuenta el poder discriminativo de cada canal localmente (ecuación 5.3). Hemos llamado a esta métrica *multispectral semi norm* (MSN):

$$SSD_{i,j,s} = \frac{\|P(A_i) - P(A'_{s,j})\|_2^2}{M} \quad (5.1)$$

$$MSSD_{i,j,s} = \frac{\|P(A_i) - P(A'_{s,j})\|_2^2 + \|P(B_i) - P(B'_{s,j})\|_2^2}{2M} \quad (5.2)$$

$$MSN_{i,j,s} = \frac{\|P(A_i) - P(A'_{s,j})\|_2^2 \cdot \|P(B_i) - P(B'_{s,j})\|_2^2}{M \left( \|P(A_i) - P(A'_{s,j})\|_2^2 + \|P(B_i) - P(B'_{s,j})\|_2^2 \right)} \quad (5.3)$$

donde  $A$  y  $B$  representan los canales T1 y T2 de la imagen a segmentar,  $A'$  y  $B'$  representan los canales T1 y T2 de la librería,  $P(A_i) \in A$  representa un parche de la imagen  $A$  centrado en las coordenadas  $i$ ,  $P(B_j) \in B$  representa un parche de la imagen  $B$  centrado en las coordenadas  $j$  y  $M$  es el número de vóxeles por parche.

### 5.2.5. Multiscale label fusion

Como se comentó en el capítulo anterior, en el método OPAL original, los mapas de probabilidad de etiquetas se estiman utilizando dos procesos independientes con dos tamaños de parche diferentes para dar cuenta de las características de escala múltiple. Estos mapas se promedian de manera uniforme (proceso denominado late fusión) para obtener el mapa de probabilidad final. En esta nueva variante del algoritmo, que llamamos MOPAL, utilizamos también los coeficientes  $\alpha$  de mezcla de múltiples escalas dependientes de la etiqueta para equilibrar las diferentes contribuciones de escala por etiqueta (ecuación 4.2).

### 5.2.6. Corrector de errores sistemáticos

Cualquier método de segmentación está sujeto a errores tanto aleatorios como sistemáticos. Los primeros pueden ser minimizados mediante técnicas de promedio reduciendo la desviación estándar del error de clasificación. Este es el caso de MOPAL donde, debido a su enfoque sobrecompleto, un alto número de votos por vóxel se usan para reducir el error de clasificación. Por desgracia, los errores sistemáticos no pueden minimizarse con esta estrategia. Sin embargo, este sesgo, siendo

sistemático puede aprenderse para corregir/calibrar la segmentación resultante. En 2011, Wang identificó este problema y propuso un método llamado SegAdapter (Wang et al., 2011). Este método, se basa en el uso de un clasificador de tipo *Adaboost* el cual aprende y corrige localmente los errores sistemáticos usando información espacial (coordenadas) y de intensidad (parches).

Inspirados por el trabajo pionero de Wang, proponemos un método de corrección de errores basado en un conjunto de redes neuronales, entrenadas usando una técnica de *boosting* (Schapire, 1990), para aumentar la precisión mediante la reducción de los errores sistemáticos producidos por el método de segmentación propuesto. Hemos llamado a este método PEC (*Patch-based Ensemble Corrector*).

El conjunto de redes ha sido entrenado con un vector de características formado por parches de tamaños 3x3x3 vóxeles (muestreado completo) y 7x7x7 vóxeles (submuestreado saltando dos vóxeles en cada dimensión, es decir, dando como resultado 3x3x3 valores) de las imágenes T1, T2, la segmentación automática, una distancia Euclídea y sus coordenadas en x, y, z en el espacio MNI. El mapa de distancia Euclídea fue calculado a partir de la segmentación del hipocampo completo como la mínima distancia 3D en vóxeles de cada punto al borde del hipocampo (figura 5.2.2).

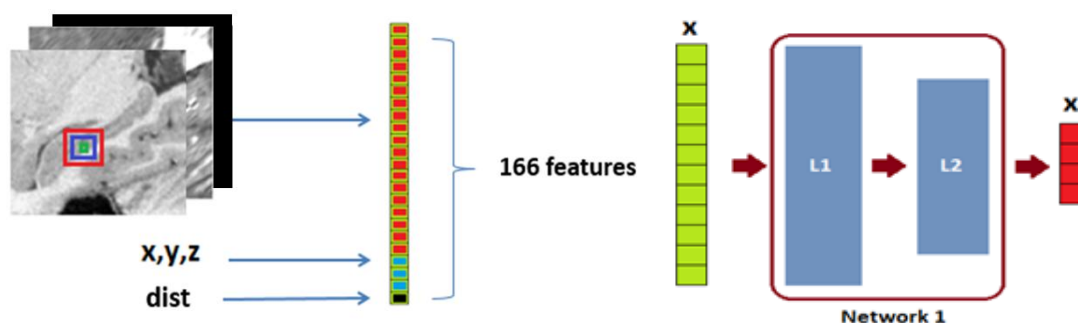


Figura 5.2.2. Izquierda: representación del vector de características de entrada. Derecha: Esquema de la red neuronal utilizada.

Esto supone un vector de 166 características que se mapean a un parche de la segmentación manual de tamaño 3x3x3 vóxeles. Se ha utilizado una clasificación basada en parches sobrecompleta como se propuso en un trabajo previo (Manjón et al., 2016). La red neuronal utilizada consiste en un perceptrón con dos capas ocultas de tamaño 83 y 55 neuronas resultando en una red con una topología de 166x83x55x27 neuronas (figura 5.2.2). Un conjunto de 10 redes neuronales ha sido entrenado usando *boosting* de forma que cada red nueva fue entrenada con un subconjunto diferente de datos seleccionado asignando una mayor probabilidad de aparición a aquellas muestras que fueron erróneamente clasificadas por la red previa (figura 5.2.3).

A diferencia de SegAdapter, nuestro método utiliza dos parches por posición (uno pequeño y uno grande) lo cual le permite ser específico en cada punto al mismo tiempo que incluimos información de contexto e información topológica con el mapa de distancia Euclídea. Otro aspecto importante es que se utiliza un enfoque de predicción estructurada (*patch-wise*) en contraste con la predicción de un solo vóxel (*voxel-wise*)

propuesta por Wang lo cual produce predicciones más regulares. El esquema de corrección sobrecompleto además incrementa el número de votos produciendo a una estimación más robusta. Finalmente, pese a que el número de redes usado en este corrector ( $N = 10$ ) puede parecer pequeño en comparación con los 500 árboles utilizados en SegAdapter se debe tener en cuenta que las redes neuronales son clasificadores mucho más fuertes que los árboles de decisión. La figura 5.3.1 muestra un ejemplo de los resultados de PEC.

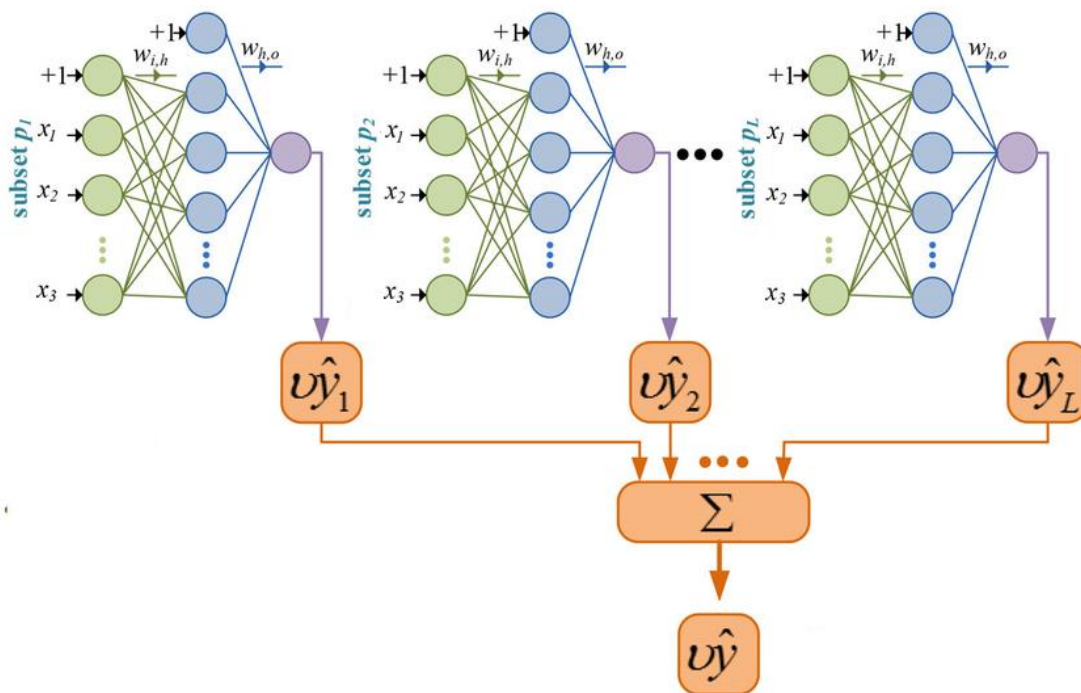


Figura 5.2.3. Esquema de la técnica de boosting usada. Cada red se entrena con un subconjunto de muestras diferente (subset  $p_i$ ) en el cual se incluyen muestras “difíciles” que no han sido clasificadas correctamente por las redes anteriores. La predicción final es un promedio de las predicciones de las múltiples redes entrenadas.

### 5.3. Experimentos y resultados

En esta sección se presentan los parámetros del método propuesto y sus resultados. Los parámetros han sido ajustados independientemente para trabajar con los dos protocolos/datasets y los resultados se han comparado con otros métodos del estado del arte. Para evaluar la calidad de las segmentaciones utilizamos el índice DICE (Zijdenbos et al., 1994). La significación estadística ha sido evaluada utilizando un test de Wilcoxon dado que los valores de DICE obtenidos no siguen una distribución normal.

### 5.3.1. Parámetros de MOPAL

En todos los experimentos se han utilizado parches de tamaño 3x3x3 y 7x7x7 vóxeles para cada escala respectivamente. El área de búsqueda ha sido fijada a 7x7x7 vóxeles. Se han utilizado 64 PatchMatches independientes con 4 iteraciones cada uno. Todos estos parámetros fueron optimizados para ambos *datasets* conjuntamente.

#### *Dataset de Winterburn*

En todos los experimentos con este *dataset* se han utilizado los siguientes coeficientes de mezcla de escalas (5 estructuras + fondo) siendo  $\alpha = [0.4711, 0.3443, 0.3826, 0.3900, 0.8439, 0.7715]$ . Estos coeficientes han sido optimizados utilizando la versión con métrica MSN del método propuesto y realizando una validación cruzada LOO consistente en 5 rondas de optimización excluyendo el caso bajo evaluación y su versión invertida (5 rondas de 8 sujetos de optimización y 1 sujeto de validación). El resultado fueron 5 conjuntos de coeficientes que fueron usados en el resto de experimentos. Por simplicidad, los valores de  $\alpha$  mostrados son el promedio de las 5 rondas de optimización con se hizo en el apartado 4.

La tabla 5.3.1 muestra los resultados del método en sus versiones de un solo contraste (T1 o T2) y multi-contraste (T1+T2) basado en la métrica SSD clásica (ecuación 5.2) y la nueva métrica MSN calculados mediante validación cruzada LOO. Se puede observar que para este *dataset* las imágenes T1 presentan una contribución a la segmentación muy pobre ya que en general no se encuentran diferencias significativas entre las versiones T2 y multi-contraste. Esto tiene sentido puesto que las imágenes T2 de este *dataset* presentan mejor contraste que las T1. De hecho, la segmentación manual fue realizada sobre las imágenes T2. También se puede observar que la nueva métrica MSN propuesta permite explotar mejor la información complementaria de los canales T1 y T2 en la segmentación.

*Tabla 5.3.1: DICE promedio en el espacio MNI y su desviación estándar para la segmentación de cada estructura utilizando imágenes de alta resolución (HR) T1, T2 y multi-contraste (T1+T2) sobre el dataset de Winterburn. Las diferencias significativas se han señalado con \* entre T1w y T2w, con † entre T1w y T1w+T2w MSN, con ‡ entre T2w y T1w+T2w y con  $\phi$  entre T1w+T2w SSD y T1w+T2w MSN ( $p < 0.05$ ). Mejores resultados en negrita.*

<i>Estructura</i>	<i>T1w HR</i>	<i>T2w HR</i>	<i>T1w+T2w HR SSD</i>	<i>T1w+T2w HR MSN</i>
<b>Promedio</b>	0.6222 ± 0.0946	0.6830 ± 0.0727*	<b>0.6803 ± 0.0711</b>	<b>0.6943 ± 0.0689 †</b>
<b>CA1</b>	0.6633 ± 0.0455	0.7394 ± 0.0287 *	0.7321 ± 0.0270	<b>0.7468 ± 0.0285 †</b>
<b>CA2\CA3</b>	0.5186 ± 0.0788	0.5916 ± 0.0511 *	0.5893 ± 0.0494	<b>0.5965 ± 0.0483 †</b>
<b>CA4\DG</b>	0.7242 ± 0.0254	<b>0.7727 ± 0.0277 *</b>	0.7542 ± 0.0282	0.7686 ± 0.0294 †
<b>SR\SL\SM</b>	0.5245 ± 0.0566	0.6604 ± 0.0389 *	0.6229 ± 0.0378	<b>0.6604 ± 0.0373 † <math>\phi</math></b>
<b>Subiculum</b>	0.6805 ± 0.0439	0.6510 ± 0.0629	0.7032 ± 0.0427	<b>0.6992 ± 0.0412 ‡</b>
<b>Hipocampo</b>	0.8717 ± 0.0284	0.8925 ± 0.0105 *	0.9019 ± 0.0133	<b>0.9056 ± 0.0114 ‡</b>

### Dataset de Kulaga-Yoskovitz

Para este *dataset* se volvieron a estimar los valores óptimos para los coeficientes de mezcla de escalas del mismo modo que para el *dataset* de Winterburn. Tenemos 3 estructuras y el fondo resultando en unos valores promedio de  $\alpha = [0.4, 0.5, 0.5, 0.9]$ . Para ello se realizaron 5 rondas de 40 sujetos para optimización y 10 para validación.

La tabla 5.3.2 muestra los resultados del método en sus versiones mono-contraste (T1 o T2) y multi-contraste (T1 y T2) basado en la métrica SSD clásica y la nueva métrica MSN calculados mediante validación cruzada *Leave-five-out*. Se puede observar que para este *dataset* las imágenes T2 presentan una contribución a la segmentación muy pobre ya que no se encuentran diferencias significativas entre las versiones T1 y multi-contraste. Una vez más esto se explica con el hecho de que las imágenes T1 de este *dataset* presentan mejor contraste que las T2. De hecho, la segmentación manual fue realizada sobre las imágenes T1. De nuevo, se puede observar que la nueva métrica MSN propuesta permite explotar mejor la información complementaria de los canales T1 y T2 en la segmentación.

Tabla 5.3.2: DICE promedio en el espacio MNI y su desviación estándar para la segmentación de cada estructura utilizando imágenes de alta resolución (HR) T1, T2 y multi-contraste (T1+T2) sobre el dataset de Kulaga-Yoskovitz. Las diferencias significativas se han señalado con \* entre T1w y T2w, con † entre T1w y T1w+T2w MSN, con ‡ entre T2w y T1w+T2w y con ϕ entre T1w+T2w SSD y T1w+T2w MSN ( $p < 0.05$ ). Mejores resultados en negrita.

Estructura	T1w HR	T2w HR	T1w+T2w HR SSD	T1w+T2w HR MSN
<b>Promedio</b>	0.8797 ± 0.0265	0.8426 ± 0.0304 *	0.8753 ± 0.0228	<b>0.8826 ± 0.0259 ‡</b>
<b>CA1-3</b>	0.9088 ± 0.0153	0.8727 ± 0.0208 *	0.9015 ± 0.0144	<b>0.9115 ± 0.0151 ‡ ϕ</b>
<b>CA4DG</b>	0.8571 ± 0.0321	0.8429 ± 0.0476	0.8600 ± 0.0349	<b>0.8616 ± 0.0339 †‡</b>
<b>Subiculum</b>	0.8733 ± 0.0209	0.8120 ± 0.0381 *	0.8645 ± 0.0238	<b>0.8746 ± 0.0236 ‡ ϕ</b>
<b>Hipocampo</b>	<b>0.9583 ± 0.0073</b>	0.9202 ± 0.0152 *	0.9507 ± 0.0075	0.9581 ± 0.0067 ‡ ϕ

### 5.3.2. Corrector de errores sistemáticos

Finalmente, se ha evaluado el método de corrección de errores propuesto y se ha comparado con SegAdapter (Wang et al., 2011). Para esta comparación se ha usado la versión 1.9 de SegAdapter con sus parámetros óptimos obtenidos empíricamente. Para el *dataset* de Winterburn se ha utilizado un radio de dilatación de 1 para obtener la ROI, una frecuencia de muestreo de 0.15, un tamaño de parche de 7x7x7 vóxeles y las imágenes T1 y T2 como características. Para el *dataset* de Kulaga-Yoskovitz se ha usado también un radio de dilatación de 1, una frecuencia de muestreo de 0.1, un tamaño de parche de 7x7x7 y las imágenes T1 como características (las imágenes T2 no resultaron de ayuda en este *dataset*).

Para el *dataset* de Winterburn se entrenaron ambos SegAdapter y PEC 5 veces usando una validación cruzada LOO excluyendo cada par de hipocampos (izquierdo y derecho) de cada caso de evaluación. Para el *dataset* de Kulaga-Yoskovitz este esquema resultaría en 5 estructuras x 25 sujetos = 125 rondas de entrenamiento lo

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

que supone varias semanas de procesamiento por lo que se realizaron solo dos rondas de entrenamiento dividiendo el *dataset* en dos grupos de 15 y 10 sujetos para realizar la validación cruzada entre ellos.

Las tablas 5.3.3 y 5.3.4 muestran el índice DICE promedio alcanzado por los dos métodos de corrección sobre ambos *datasets*. Nótese que PEC obtuvo buenos resultados en ambos *datasets*. Sin embargo, el incremento obtenido para el *dataset* de Winterburn fue mayor. Esto tiene sentido ya que la segmentación para este *dataset* presenta un mayor margen de mejora (0,6943 frente a 0,8826) y el tamaño de la librería es bastante pequeño. Además, las estructuras del protocolo de Winterburn son más pequeñas en su definición que las de Kulaga-Yoskovitz de modo que los cambios tienen mayor impacto sobre la precisión en el *dataset* de Winterburn.

La mejora obtenida por SegAdapter fue bastante pequeña para el *dataset* de Kulaga-Yoskovitz y obtuvo resultados negativos para el *dataset* de Winterburn. Esto no es un resultado esperable ya que se trata de un método validado. Únicamente podemos suponer que SegAdapter no está bien adaptado para trabajar con datos de alta resolución ya que fue diseñado para aplicarse a datos de resolución estándar y que el reducido número de casos de Winterburn hace la situación aún más desfavorable introduciendo problemas de sobreajuste. Se observó que PEC funciona mejor sobre el *dataset* de Kulaga-Yoskovitz si se excluye la estructura CA4\DG por lo que la corrección no se aplicó para esta estructura.

No se pudo aplicar este tipo de selección al SegAdapter puesto que requiere que todas las etiquetas sean corregidas al mismo tiempo. Todos los resultados que se muestran a continuación en este capítulo se refieren a la versión final de HIPS, es decir, HIPS+PEC. La figura 5.3.1 muestra un ejemplo de aplicación de PEC a la segmentación sobre el *dataset* de Kulaga-Yoskovitz.

Tabla 5.3.3: DICE promedio en el espacio MNI para la segmentación de HIPS aplicando SegAdapter y PEC sobre el *dataset* de Winterburn. Diferencias significativas marcadas con \* entre HIPS y HIPS + PEC, con † entre HIPS + SegAdapter y HIPS + PEC. No se encontraron diferencias entre HIPS y HIPS + SegAdapter.

<b>Estructura</b>	<b>HIPS</b>	<b>HIPS + SegAdapter</b>	<b>HIPS + PEC</b>
<b>Promedio</b>	0.6943 ± 0.0689	0.6822 ± 0.0786	<b>0.7158 ± 0.0652†</b>
<b>CA1</b>	0.7468 ± 0.0285	0.7470 ± 0.0226	<b>0.7762 ± 0.0251*†</b>
<b>CA2\CA3</b>	0.5965 ± 0.0483	0.5683 ± 0.0512	<b>0.6179 ± 0.0630</b>
<b>CA4\DG</b>	0.7686 ± 0.0294	0.7622 ± 0.0317	<b>0.7750 ± 0.0307</b>
<b>SR\SL\ISM</b>	0.6604 ± 0.0373	0.6489 ± 0.0274	<b>0.7018 ± 0.0191*†</b>
<b>Subiculum</b>	0.6992 ± 0.0412	0.6844 ± 0.0418	<b>0.7082 ± 0.0597</b>
<b>Hipocampo</b>	0.9056 ± 0.0114	0.9003 ± 0.0117	<b>0.9111 ± 0.0098†</b>



Tabla 5.3.4: DICE promedio en el espacio MNI para la segmentación de HIPS aplicando SegAdapter y PEC sobre el dataset de Kulaga-Yoskovitz. Diferencias significativas marcadas con \* entre HIPS y HIPS + PEC, con † entre HIPS + SegAdapter y HIPS + PEC. No se encontraron diferencias entre HIPS y HIPS + SegAdapter.

Estructura	HIPS	HIPS + SegAdapter	HIPS + PEC
<b>Promedio</b>	0.8826 ± 0.0259	0.8833 ± 0.0247	<b>0.8879 ± 0.0271</b>
<b>CA1-3</b>	0.9115 ± 0.0151	0.9115 ± 0.0126	<b>0.9158 ± 0.0145</b>
<b>CA4/DG</b>	0.8616 ± 0.0339	<b>0.8656 ± 0.0286</b>	0.8616 ± 0.0339
<b>Subiculum</b>	0.8746 ± 0.0236	0.8727 ± 0.0226	<b>0.8863 ± 0.0206*†</b>
<b>Hipocampo</b>	0.9581 ± 0.0067	0.9573 ± 0.0061	<b>0.9595 ± 0.0064</b>

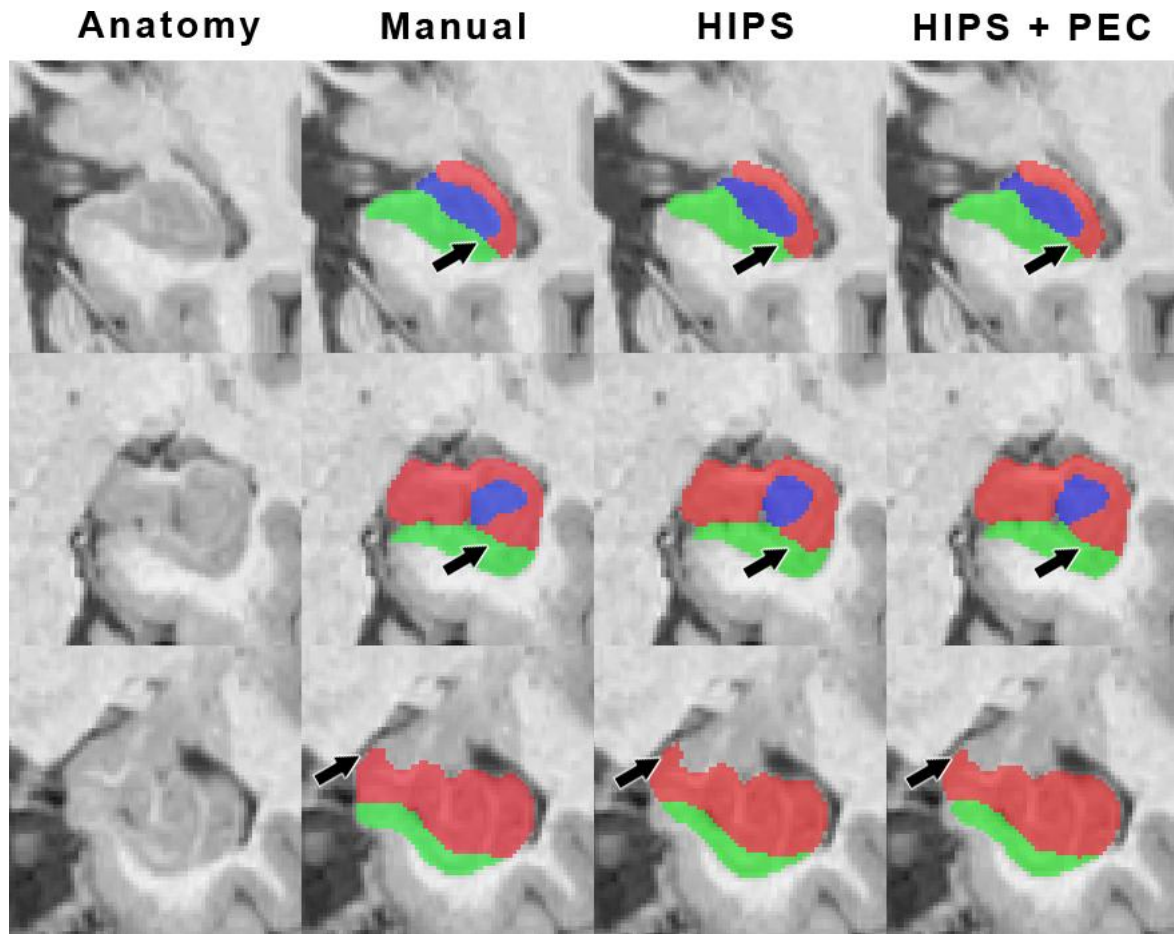


Figura 5.3.1: Ejemplo de aplicación de PEC sobre una segmentación del dataset de Kulaga-Yoskovitz. De izquierda a derecha se muestran la imagen anatómica, la segmentación manual, la segmentación automática y la segmentación corregida. Las flechas señalan puntos donde PEC realizó cambios.

### 5.3.3. Resolución estándar frente a alta resolución

El método propuesto trabaja con imágenes de alta resolución pero estas secuencias no siempre se encuentran disponibles tanto en investigación como en entornos clínicos. Por lo tanto, sería conveniente disponer de una herramienta capaz de analizar datos retrospectivos. Por esto, se ha evaluado el método propuesto sobre imágenes de resolución estándar (1 mm<sup>3</sup>) remuestreadas a 0,5x0,5x0,5 mm<sup>3</sup>. Para ello se han utilizado una interpolación por B-Splines y un método reciente de superresolución

llamado LASR (Coupé et al., 2013). Para simular las imágenes de resolución estándar, se redujo la resolución de las imágenes de alta resolución por un factor 2 para posteriormente volver a aumentarla con los métodos mencionados.

Se ha realizado este experimento con dos configuraciones del método de segmentación: la versión multi-contraste (tablas 5.3.5 y 5.3.7) y la versión mono-contraste utilizando solo T1 (tablas 5.3.6 y 5.3.8) ya que esta configuración representa mejor la aplicación del método a datos retrospectivos que normalmente consisten en secuencias T1.

Las tablas 5.3.5, 5.3.6, 5.3.7 y 5.3.8 muestran los resultados sobre ambos datasets que confirman que HIPS puede producir resultados competitivos incluso utilizando imágenes de resolución estándar. Nótese que los resultados utilizando LASR son mejores que utilizando B-Splines para el *dataset* de Kulaga yoskovitz. Sin embargo, esta mejora es pequeña pese a la significación estadística. Además, la precisión obtenida utilizando imágenes superresueltas con LASR es muy similar a la obtenida con imágenes de alta resolución ya que no se han encontrado diferencias significativas entre ambas opciones. Esto es un hecho relevante ya que demuestra que el método propuesto puede procesar imágenes de  $1 \text{ mm}^3$  (que es la resolución más habitual). Aunque se podría argumentar que el método está utilizando información a priori de la librería en lugar de seguir las referencias anatómicas contenidas en la imagen, el análisis de los resultados sugiere que esto no es así. Basta con observar las imágenes superresueltas para comprobar que estructuras como la "banda oscura" (en T1) se hace más visible tras el proceso de superresolución (ver figuras 3.5.2 y 3.5.3). Es más, las diferencias entre el uso de B-Splines y LASR solo se pueden explicar cómo el efecto del proceso de superresolución. Las figuras 3.5.2 y 3.5.3 muestran un ejemplo de segmentación basada en alta resolución y superresolución y como la anatomía se recupera parcialmente.

Tabla 5.3.5: Resultados sobre el dataset de Winterburn. DICE promedio en el espacio MNI y desviación estándar para cada estructura utilizando imágenes de alta resolución, imágenes interpoladas con B-Splines, imágenes superresueltas con LASR e imágenes de alta resolución (HR). Segmentación producida utilizando la versión multi-contraste del método propuesto. No se encontraron diferencias significativas entre el uso de B-Splines y LASR, B-Splines y HR y LASR y HR. Mejores resultados en negrita.

<i>Estructura</i>	<i>B-Spline T1w + T2w</i>	<i>LASR T1w + T2w</i>	<i>HR T1w + T2w</i>
<i>Promedio</i>	0.7078 ± 0.0659	<b>0.7108 ± 0.0647</b>	0.7158 ± 0.0652
<i>CA1</i>	0.7690 ± 0.0267	<b>0.7707 ± 0.0267</b>	0.7762 ± 0.0251
<i>CA2\CA3</i>	0.6108 ± 0.0741	<b>0.6170 ± 0.0655</b>	0.6179 ± 0.0630
<i>CA4\DG</i>	0.7690 ± 0.0306	<b>0.7732 ± 0.0305</b>	0.7750 ± 0.0307
<i>SR\SL\ISM</i>	0.6871 ± 0.0230	<b>0.6903 ± 0.0216</b>	0.7018 ± 0.0191
<i>Subiculum</i>	<b>0.7030 ± 0.0668</b>	0.7025 ± 0.0614	0.7082 ± 0.0597
<i>Hipocampo</i>	0.9080 ± 0.0089	<b>0.9119 ± 0.0130</b>	0.9111 ± 0.0098

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

Tabla 5.3.6: Resultados sobre el dataset de Winterburn. DICE promedio en el espacio MNI y desviación estándar para cada estructura utilizando imágenes de alta resolución, imágenes interpoladas con B-Splines, imágenes superresueltas con LASR e imágenes de alta resolución (HR). Segmentación producida utilizando la versión mono-contraste (T1w) del método propuesto.. No se encontraron diferencias significativas entre el uso de B-Splines y LASR, B-Splines y HR y LASR y HR. Mejores resultados en negrita.

Estructura	B-spline T1w	LASR T1w	HR T1w
<b>Promedio</b>	0.6082 ± 0.0986	<b>0.6176 ± 0.0953</b>	0.6222 ± 0.0946
CA1	0.6590 ± 0.0504	<b>0.6638 ± 0.0478</b>	0.6633 ± 0.0455
CA2\CA3	0.5011 ± 0.0823	<b>0.5154 ± 0.0787</b>	0.5186 ± 0.0788
CA4\DG	0.7139 ± 0.0278	<b>0.7166 ± 0.0236</b>	0.7242 ± 0.0254
SR\SL\SM	0.5046 ± 0.0531	<b>0.5154 ± 0.0521</b>	0.5245 ± 0.0566
Subiculum	0.6626 ± 0.0472	<b>0.6769 ± 0.0437</b>	0.6805 ± 0.0439
<b>Hipocampo</b>	0.8741 ± 0.0186	<b>0.8765 ± 0.0205</b>	0.8717 ± 0.0284

Tabla 5.3.7: Resultados sobre el dataset de Kulaga-Yoskovitz. DICE promedio en el espacio MNI y desviación estándar para cada estructura utilizando imágenes de alta resolución, imágenes interpoladas con B-Splines, imágenes superresueltas con LASR e imágenes de alta resolución (HR). Segmentación producida utilizando la versión multi-contraste del método propuesto. El test de Kurskal-Wallis reveló que existen diferencia en cuanto al DICE promedio, el CA1-3 y el hipocampo completo. Las diferencias significativas se han señalado con \* entre B-Spline y LASR, con † entre LASR y HR y con ‡ entre B-Spline y HR. Mejores resultados en negrita.

Estructura	B-spline T1w + T2w	LASR T1w + T2w	HR T1w + T2w
<b>Promedio</b>	0.8803 ± 0.0288	<b>0.8828 ± 0.0280*</b>	0.8879 ± 0.0271†‡
CA1-3	0.9100 ± 0.0146	<b>0.9120 ± 0.0137*</b>	0.9158 ± 0.0145†‡
CA4/DG	0.8525 ± 0.0331	<b>0.8563 ± 0.0325</b>	0.8616 ± 0.0339
Subiculum	0.8783 ± 0.0226	<b>0.8800 ± 0.0220</b>	0.8863 ± 0.0206
<b>Hipocampo</b>	0.9552 ± 0.0070	<b>0.9566 ± 0.0065*</b>	0.9595 ± 0.0064†‡

Tabla 5.3.8: Resultados sobre el dataset de Kulaga-Yoskovitz. DICE promedio en el espacio MNI y desviación estándar para cada estructura utilizando imágenes de alta resolución, imágenes interpoladas con B-Splines, imágenes superresueltas con LASR e imágenes de alta resolución (HR). Segmentación producida utilizando la versión multi-contraste del método propuesto. El test de Kurskal-Wallis reveló que existen diferencia en cuanto al DICE promedio y el CA4/DG. Las diferencias significativas se han señalado con \* entre B-Spline y LASR, con † entre LASR y HR y con ‡ entre B-Spline y HR. Mejores resultados en negrita.

Estructura	B-spline T1w	LASR T1w	HR T1w
<b>Promedio</b>	0.8709 ± 0.0314	<b>0.8732 ± 0.0307</b>	0.8797 ± 0.0265‡
CA1-3	0.9030 ± 0.0159	<b>0.9052 ± 0.0152</b>	0.9088 ± 0.0153
CA4/DG	0.8403 ± 0.0326	<b>0.8439 ± 0.0326</b>	0.8571 ± 0.0321‡
Subiculum	0.8693 ± 0.0218	<b>0.8704 ± 0.0214</b>	0.8733 ± 0.0209
<b>Hipocampo</b>	0.9546 ± 0.0080	<b>0.9566 ± 0.0077</b>	0.9583 ± 0.0073

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

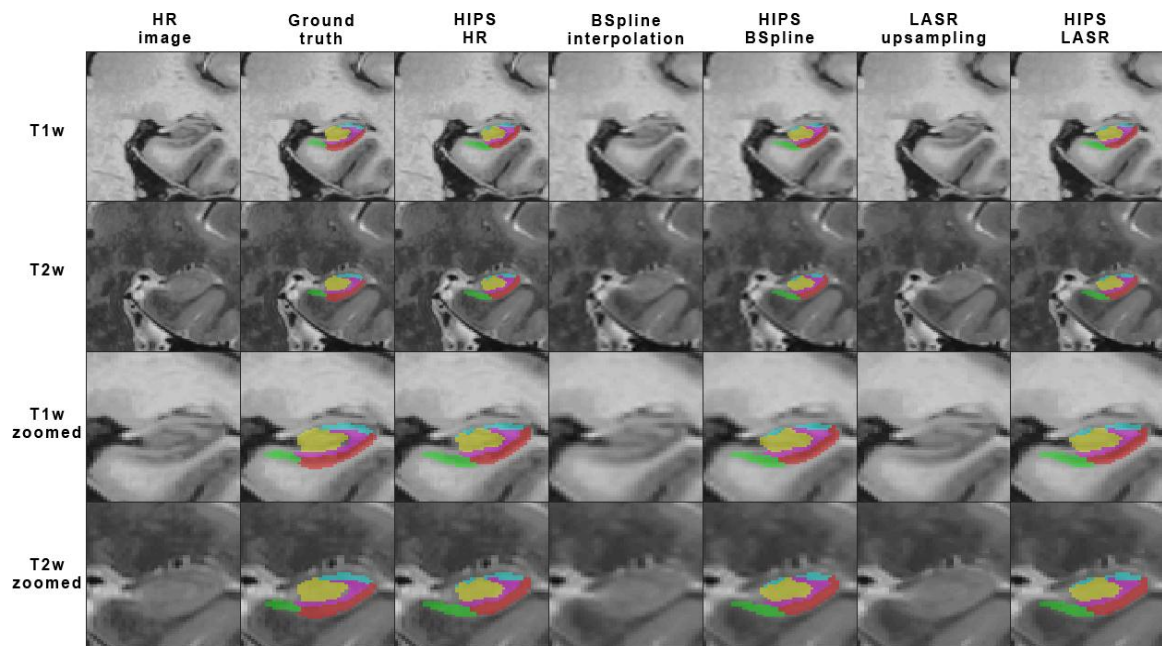


Figura 3.5.2: Ejemplo de segmentación utilizando la versión multi-contraste de HIPS sobre imágenes HR, interpoladas con B-Splines y superresueltas con LASR (dataset de Winterburn).

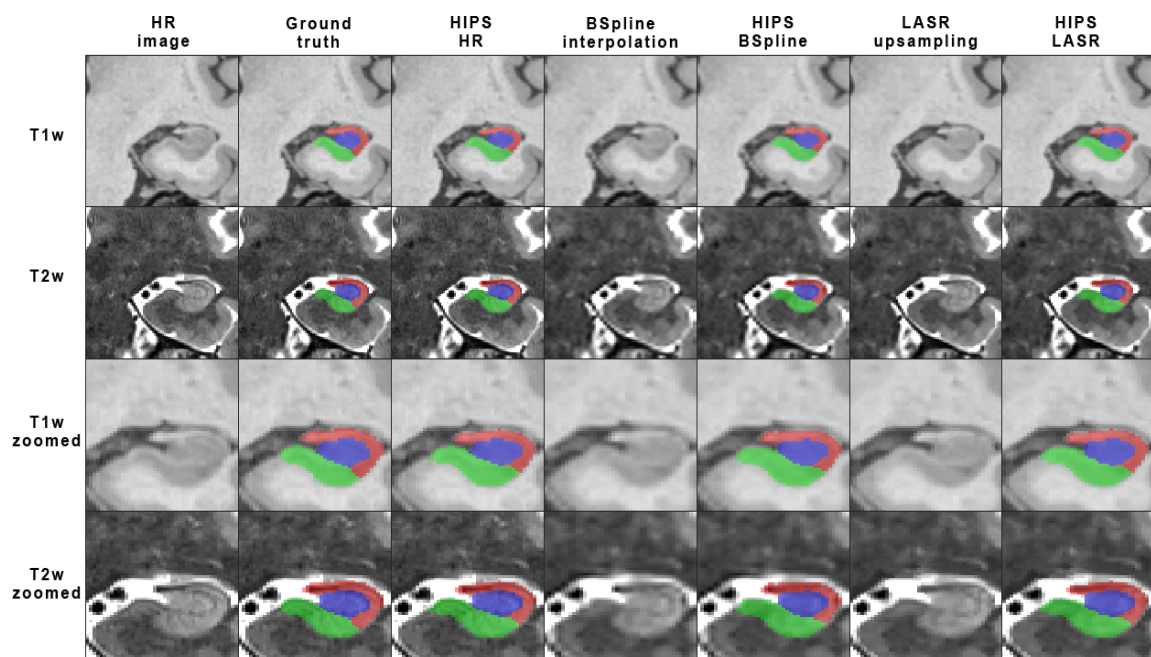


Figura 3.5.3: Ejemplo de segmentación utilizando la versión multi-contraste de HIPS sobre imágenes HR, interpoladas con B-Splines y superresueltas con LASR (dataset de Kulaga-Yoshcovitz).

### 5.3.4. Comparativa de métodos

En este experimento se ha comparado HIPS con otros métodos recientes aplicados a la segmentación de subregiones del hipocampo en ambos *datasets*. Se ha comparado HIPS sobre el *dataset* de Winterburn con los resultados de MAgE (Pipitone et al., 2014) proporcionados por los autores. También se ha comparado HIPS sobre el

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

*dataset* de Kulaga-Yoskovitz con ASHS (Yushkevich et al., 2015b) y SurfPatch (Caldairou et al., 2016). Se han tomado los resultados de ASHS y SurfPatch proporcionados en (Caldairou et al., 2016). Además, se ha incluido información sobre la precisión de la segmentación inter e intra-observador obtenida de los artículos originales (Winterburn et al., 2013) y (Kulaga-Yoskovitz et al., 2015) como referencia.

La tabla 5.3.9 muestra los resultados de MAGeT y HIPS sobre el *dataset* de Winterburn y la tabla 5.3.10 muestra los resultados de ASHS, SurfPatch y HIPS sobre el *dataset* de Kulaga-Yoskovitz. Para que la comparativa fuese justa los resultados de HIPS se calcularon en el espacio nativo de las imágenes utilizando la correspondiente transformación inversa ya que los resultados de los métodos comparados fueron calculados en ese espacio. Nótese que la utilización de esta transformación inversa reduce ligeramente la calidad de la segmentación del método propuesto debido a los artefactos introducidos por la interpolación usada.

En el caso de la comparativa con MAGeT, las imágenes de Winterburn tienen a una resolución nativa de  $0,3 \times 0,3 \times 0,3 \text{ mm}^3$  mientras que MAGeT produce segmentaciones a una resolución de  $0,9 \times 0,9 \times 0,9 \text{ mm}^3$  por eficiencia. Por esto, se ha bajado de resolución las segmentaciones producidas por HIPS de  $0,3 \times 0,3 \times 0,3 \text{ mm}^3$  a  $0,9 \times 0,9 \times 0,9 \text{ mm}^3$  en el espacio nativo para realizar una comparativa justa. En general HIPS mostró una mejora del 26% en comparación con MAGeT alcanzando un DICE promedio de 0,661.

Sobre el *dataset* de Kulaga-Yoskovitz nuestro método mejoró claramente tanto a ASHS como a SurfPatch en todas las estructuras llegando a superar la precisión inter-observador en un 3% para la estructura CA1-3 y permaneciendo solo un 1% por debajo del DICE inter-observador promedio con un DICE de 0,8744.

Tabla 5.3.9: DICE promedio en el espacio nativo para cada estructura. Segmentación producida por MAGeT y HIPS a  $0,9 \text{ mm}^3$  sobre el *dataset* de Winterburn. Mejores resultados en negrita. Segmentación del intra-observador a  $0,3 \text{ mm}$  incluida a modo de referencia.

<i>Estructura</i>	<i>MAGeT</i> (T1 0.9 mm)	<i>HIPS</i> (T1+T2 0.9 mm)	<i>Intra-observador</i> (T1 0.3 mm)
<b>Promedio</b>	0.526	<b>0.661</b>	0.742
<b>CA1</b>	0.563	<b>0.670</b>	0.780
<b>CA2 CA3</b>	0.412	<b>0.522</b>	0.640
<b>CA4 DG</b>	0.647	<b>0.763</b>	0.830
<b>SR SL SM</b>	0.428	<b>0.599</b>	0.710
<b>Subiculum</b>	0.580	<b>0.722</b>	0.750
<b>Hipocampo</b>	0.816	<b>0.876</b>	0.910

## Capítulo 5: Segmentación de subregiones del hipocampo (HIPS)

Tabla 5.3.10: DICE promedio en el espacio nativo para cada estructura. Segmentación producida por ASHS, SurfPatch HIPS y observador humano sobre el dataset de Kulaga-Yoskovitz. Mejores resultados en negrita.

Estructura	ASHS	SurfPatch	HIPS	Inter-observador	Intra-observador
<b>Promedio</b>	0.8513	0.8503	<b>0.8744</b>	0.8833	0.9113
<b>CA1-3</b>	0.8736±0.0197	0.8743±0.0247	<b>0.9030±0.0138</b>	0.8760 ± 0.048	0.9290 ± 0.010
<b>CA4DG</b>	0.8254±0.0345	0.8271±0.0285	<b>0.8497±0.0332</b>	0.9030 ± 0.036	0.9000 ± 0.019
<b>Subiculum</b>	0.8548±0.0243	0.8495±0.0245	<b>0.8705±0.0212</b>	0.8710 ± 0.053	0.9050 ± 0.016

En lo referente al tiempo de ejecución, el pipeline completo de HIPS necesita en promedio menos de 20 minutos mientras que los demás métodos tienen una carga computacional de varias horas por caso.

### 5.4. Conclusión y discusión

En este capítulo se ha presentado un método para la segmentación de las subregiones del hipocampo que produce segmentaciones al nivel del estado del arte en un tiempo de ejecución reducido.

Una de las contribuciones de este trabajo es una nueva métrica multispectral para calcular la similitud entre parches basada en una semi-norma SSD (MSN). Introduciendo esta métrica en OPAL (ahora MOPAL) se han conseguido buenos resultados en la segmentación utilizando información de las imágenes T1 y T2. Utilizando la semi-norma para combinar distancias se obtiene una métrica robusta y auto-balanceada que se beneficia de la información de ambos canales. Esto significa que puede lidiar con problemas tales como datos corruptos o de baja calidad en uno de los dos canales. Esta contribución hace el método más robusto especialmente cuando se aplica a conjuntos de datos diferentes.

Se ha demostrado que nuestro método puede producir segmentaciones competitivas utilizando imágenes de resolución estándar gracias al uso de un método de superresolución. Esta es una característica importante ya que hace del método propuesto una herramienta adecuada para el análisis de datos retrospectivos.

Desde el punto de vista de la robustez el registro es un paso clave. Por ello se ha utilizado un registro multi-contraste así como una segmentación multi-contraste. De este modo cubrimos una mayor variabilidad y reducimos la dispersión de los resultados en términos de precisión, especialmente en el *dataset* de Kulaga-Yoskovitz.

Hemos demostrado que HIPS supera a otros métodos del estado del arte en términos de calidad de segmentación alcanzado un DICE promedio de 0,661 para el dataset de Winterburn mientras que MAGeT obtuvo un 0,5260. En el *dataset* de Kulaga-Yoskovitz HIPS obtuvo un DICE promedio de 0,8744 mientras que ASHS y SurfPatch obtuvieron un 0,8513 y 0,8503 respectivamente. HIPS es además más rápido ya que solo necesita 20 minutos para completar la segmentación frente a sus competidores que

requieren varias horas por caso. Todos los experimentos incluida la medición del tiempo de cómputo fueron realizados en una maquina con un procesador Intel core i7 y 8 GB de memoria RAM ejecutando MATLAB 2017a en un sistema Windows 7 Profesional 64 bit.

Otra de las contribuciones del método propuesto es la utilización de un nuevo método de corrección de errores sistemáticos utilizando un conjunto de redes neuronales basadas en parche (PEC). El uso de este corrector mejoró significativamente los resultados en ambos *datasets* con una sobrecarga temporal de tan solo unos pocos segundos. Incluso siendo ambos (PEC y SegAdapter) métodos basados en conjuntos de clasificadores, PEC ha demostrado funcionar mejor que SegAdapter usando un número significativamente menor de clasificadores. Ambos métodos utilizan técnicas de *boosting* para aprender y corregir los patrones de error. La diferencia principal de PEC es el uso de clasificadores fuertes basados en parche en lugar de clasificadores débiles. Se escogió el uso de redes neuronales por su versatilidad y posibilidad de realizar predicciones estructuradas (corrección de parche frente a corrección de vóxel) mejorando la regularidad de las etiquetas.

Además, la elección de un clasificador fuerte permite que el método necesite tan solo 10 redes para alcanzar la máxima precisión. Se debe mencionar el hecho de que, en combinación con el corrector de errores, el método de segmentación propuesto se acerca a la calidad obtenida por el observador humano en el *dataset* de Kulaga-Yoskovitz donde obtuvo un DICE superior al del inter-observador para el CA1-3 (HIPS = 0,9030 frente a inter-observador = 0,8760), prácticamente la misma precisión para el *Subiculum* (HIPS = 0,8705 frente a inter-observador = 0,8710) y presenta un DICE promedio de 0,8744 que está considerablemente cerca del 0,8833 obtenido por el inter-observador.

Como puede deducirse de los resultados, la principal limitación del método propuesto es la utilización de una librería muy pequeña (especialmente en el *dataset* Winterbrun) lo cual limita las capacidades de generalización del método propuesto. Comparando los resultados obtenidos en ambos *datasets* es esperable que los resultados sobre Winterburn mejoren con la adición de más casos segmentados a mano. Incluso siendo esto altamente dependiente de la aplicación, basándonos en trabajos previos usando segmentación *multi-atlas*, pensamos que los resultados óptimos se pueden obtener utilizando al menos 20 casos de referencia. Planeamos añadir nuevos casos segmentados a mano a la librería para mejorar los resultados del método.

Finalmente, HIPS ha sido utilizado recientemente para demostrar como la esclerosis múltiple afecta al hipocampo específicamente desde el *dentate gyrus* hasta el CA1 (Planche et al., 2018) con un resultado que confirma esta hipótesis lo cual demuestra la validez de HIPS como herramienta para el estudio de enfermedades neurodegenerativas.

HIPS se encontrará en breve disponible para la comunidad científica como parte de nuestra plataforma online volBrain (<http://volbrain.upv.es>).

## Capítulo 6: Conclusiones generales y líneas de futuro

En esta tesis se han presentado tres métodos de segmentación de diferentes regiones y estructuras cerebrales basados en el uso de la tecnología de *multi-atlas label fusion*. Concretamente, se ha presentado NABS, que divide el cerebro en 5 regiones (hemisferios cerebrales izquierdo y derecho, hemisferios cerebelares izquierdo y derecho y bulbo raquídeo), CERES, que segmenta los lobulillos del cerebelo (I-II, III, IV, V, VI, Crus I, Crus II, VIIb, VIIa, VIIIa, VIIIb, IX y X) y HIPS que divide el hipocampo en subregiones de acuerdo con dos protocolos de segmentación distintos (Kulaga-Yoskovitz: CA1-3, CA4/DG y *Subiculum* y Winterburn: CA1, CA2-3, CA4/DG, SR/SL/SM y *Subiculum*).

Los objetivos principales en el desarrollo de estos tres métodos han sido mejorar la eficiencia reduciendo la carga computacional asociada al proceso de búsqueda de correspondencias entre parches y mejorar la calidad de las segmentaciones generadas.

En el caso de NABS, abordamos el problema de la eficiencia identificando el cuello de botella en el coste del método *Nonlocal label fusion* que realizaba una búsqueda exhaustiva de correspondencias entre parches. Al tener que aplicarse a todo el volumen cerebral suponía un coste temporal muy alto y para reducirlo implementamos una segmentación guiada mediante una máscara en la que gran parte del volumen cerebral se clasificaba solo utilizando información a priori. Esto fue posible gracias al registro de todos los casos al espacio MNI152 que nos permitió observar zonas que siempre se clasificaban igual en todos los sujetos de la librería. Esta implementación permitió a NABS obtener un DICE de 0,9962 frente al 0,9868 de su competidor ADisc en un tiempo muy reducido.

En el desarrollo de CERES no fue posible aplicar la misma optimización que se aplicó en NABS dada la naturaleza del cerebelo y su alta variabilidad. En su lugar, utilizamos OPAL, una optimización de *Nonlocal label fusion* que emplea el algoritmo PatchMatch reemplazando la búsqueda exhaustiva de correspondencias por una búsqueda probabilística basada en un principio de adyacencia de correspondencias (también llamado en inglés *Aproximated Nearest-Neighbor Field*, ANNF). Esto redujo drásticamente el tiempo de cómputo con lo que el cuello de botella pasó a ser el proceso de registro. Como ya hemos comentado, el cerebelo presenta una alta variabilidad entre sujetos. Por esta razón era necesario emplear un registro no lineal que tiene un coste computacional considerable. Además, nuestro objetivo era registrar no linealmente todos los sujetos de la librería al espacio MNI152 del caso a segmentar. Una implementación directa, habría tenido un coste excesivo por lo que se optó por una estrategia eficiente en la cual los registros de la librería se precomputan con lo que el coste computacional del método se mantuvo bajo límites aceptables teniendo, además, un impacto positivo en la calidad de las segmentaciones. Además, se mejoró la calidad de CERES introduciendo un mezclado multiescala adaptativo que utilizaba un coeficiente distinto para cada estructura, así como una extensión del número de casos de la librería incluyendo segmentaciones automáticas. Estas mejoras posicionaron a CERES por encima de sus competidores con un DICE de 0,7729 frente



al 0,7320 de MAGeT y al 0,6890 de RASCAL. Es importante mencionar que estos resultados se obtuvieron utilizando una librería generada a partir de tan solo 5 casos segmentados manualmente. Según trabajos previos, este es un número muy bajo de casos de referencia lo cual tiene un impacto negativo en la calidad de las segmentaciones. Es esperable que los resultados pudieran mejorar significativamente al incluir nuevos casos segmentados a mano en la librería.

Llegados a este punto, el tiempo de ejecución dejó de ser un problema y en el desarrollo de HIPS nos centramos en mantener la eficiencia conseguida en CERES mientras incorporábamos mejoras de calidad de la segmentación. Estas mejoras consistieron en el uso de una métrica multi-contraste autobalanceada que permitió utilizar imágenes T1 y T2 simultáneamente utilizando siempre la mejor información disponible de cada una de ellas. Además, se incorporó una etapa de postproceso para la corrección de errores sistemáticos basada en redes neuronales. Finalmente, HIPS obtuvo un DICE de 0,8744 superando así el 0,8513 de ASHS, y el 0,8503 de SurfPatch quedando considerablemente cerca de la precisión inter-observador (0,8833).

Además de las mejoras en la calidad, HIPS incorpora una característica muy importante y es su capacidad para trabajar con datos de resolución estándar. Como se ha comentado, para visualizar y segmentar las subregiones del hipocampo se requiere de datos de alta resolución. Desafortunadamente, estos no están siempre disponibles. Por esta razón, incorporamos en HIPS un método reciente de superresolución que permitió obtener resultados similares a los que produce con imágenes de alta resolución haciendo posible su aplicación a datos retrospectivos.

Debemos remarcar que el uso de corrección de errores sistemáticos supuso una gran mejora de los resultados de HIPS sobre el *dataset* de Winterburn de lo cual aprendimos que era un valioso recurso para mitigar el impacto negativo de contar con un número reducido de casos segmentados a mano. Siguiendo esta línea de trabajo, incluimos la corrección de errores sistemáticos en la segunda versión de CERES (CERES2) confirmando así su buen funcionamiento al resultar nuestro método vencedor del *challenge* ENIGMA de la conferencia MICCAI de 2017 ([Carass et al., 2018](#)).

En resumen, los métodos presentados producen segmentaciones con una calidad del estado del arte con un coste computacional reducido. No obstante, consideramos que una de sus limitaciones es la fuerte dependencia del número de casos segmentados a manos del que se dispone. Con pocos datos los resultados tienden a ser menos regulares. Por esta razón, se realizaron extensiones de librería que mitigan este efecto pero no resuelven el problema. Además, en nuestros experimentos con CERES, donde realizamos una extensión de librería considerable, pudimos observar que OPAL tiende a oscilar en torno al resultado óptimo cuando tiene que explorar una librería muy extensa. Pensamos que esto se debe a que OPAL no explora todo el espacio de los datos si no que busca aproximarse al máximo a la solución óptima de forma probabilística. Esto puede parecer que entra en contradicción con la necesidad de nuevos casos con lo que es un problema a tener en cuenta para futuros trabajos. Por ejemplo, realizando una selección de los atlas más similares cuando se disponga de librerías más grandes.

## Capítulo 6: Conclusiones generales y líneas de futuro

Finalmente, como propuesta para futuros proyectos nos basamos en la evidencia de que tanto los métodos MAS como las redes neuronales presentan un gran potencial. Esto junto con la importancia de la fusión de etiquetas nos hace pensar en las redes neuronales profundas como una buena tecnología de base para implementar dicha fusión de etiquetas. Combinando esta fusión de etiquetas basada en redes neuronales junto con la estrategia PatchMatch utilizada en OPAL obtendríamos un método con las ventajas de ambas tecnologías. Por un lado, la capacidad de establecer correspondencias entre imágenes de forma eficiente de OPAL y por otro lado, la capacidad de las redes neuronales para fusionar dicha información ya que se trata de un problema de integración mucho más sencillo.

Además, con esta combinación estamos abordando la problemática asociada a la poca disponibilidad de casos segmentados a mano ya que al entrenar una red basada en parches el número de muestras independientes de las que esta puede aprender pasa de unos pocos casos a cientos de miles de parches (es decir, cientos de miles de patrones con sus correspondiente clasificación).

Todas las herramientas desarrolladas en esta tesis se utilizarán en el futuro cercano en el estudio de diversas patologías como el AD o MS entre otras.

## Publicaciones realizadas

### Revistas

1. Carass A., Cuzzocreo J.L., Han S., Hernandez-Castillo C.R., Rasser P.E., Ganz M., Beliveau V., Price M., Cardenas V.A., Fein J., Dolz J., Ayed I.B., Desrosiers C., Thyreau B., Romero J.E., Coupé P., Manjón J.V., Fonov V.S., Collins D.L., Ying S.H., Crocetti D., Landman B.A., Mostofsky S.H., Thompson P.A., Prince J.L. 2018. **Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images.** *Submitted to Neuroimage journal.*
2. Planche V., Koubiyr I., Romero J.E., Manjón J.V., Coupé P., Deloire M., Dousset V., Brochet B., Ruet A., Tourdias T. 2018. **Regional hippocampal vulnerability in early multiple sclerosis: Dynamic pathological spreading from dentate gyrus to CA1.** *Human Brain Mapping* 00, 1–11.
3. Romero J.E., Coupe P., Manjón J.V. 2017. **HIPS: A new hippocampus subfield segmentation method.** *Neuroimage*, 163, 286-295.
4. Romero J.E., Coupe P., Giraud R., Ta VT., Fonov V., Park M.T.M., Chakravarty M.M., Voineskos A.N., Manjón J.V. 2017. **CERES: A new cerebellum lobule segmentation method.** *Neuroimage* 147, 916-924.
5. Romero J.E., Manjón J.V., Tohka J., Coupé P., Robles M. 2015. **Non-local Automatic Brain Hemisphere Segmentation.** *Magnetic Resonance Imaging* 33(4), 474-484.
6. Manjón J.V., Eskildsen S.F., Coupé P., Romero J.E., Collins D.L., Robles M. 2014. **Nonlocal intracranial cavity extraction.** *Journal of Biomedical Imaging* 10.

### Conferencias

1. Romero J.E., Coupe P., Manjón J.V. 2016. **High Resolution Hippocampus Subfield Segmentation Using Multispectral Multiatlas Patch-Based Label Fusion.** *International Workshop on Patch-based Techniques in Medical Imaging* 2, 117-124.
2. Romero J.E., Coupe P., Manjón J.V. 2016. **Non-local MRI Library-Based Super-Resolution: Application to Hippocampus Subfield Segmentation.** *International Workshop on Patch-based Techniques in Medical Imaging* 2, 68-75.

## Bibliografía

- Adams, R., Bischof, L., 1994. Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 641–647. <https://doi.org/10.1109/34.295913>
- Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage* 46, 726–738. <https://doi.org/10.1016/j.neuroimage.2009.02.018>
- Altshuler, L.L., Bartzokis, G., Grieder, T., Curran, J., Mintz, J., 1998. Amygdala enlargement in bipolar disorder and hippocampal reduction in schizophrenia: an MRI study demonstrating neuroanatomic specificity. *Arch. Gen. Psychiatry* 55, 663–664.
- Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de-Solorzano, C., 2009. Combination strategies in multi-atlas image segmentation: application to brain MR data. *IEEE Trans. Med. Imaging* 28, 1266–1277. <https://doi.org/10.1109/TMI.2009.2014372>
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *NeuroImage* 26, 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018>
- Asman, A.J., Landman, B.A., 2012. Muli-Atlas Segmentation using Non-Local STAPLE. *MICCAI Workshop Multi-Atlas Lebling* 87–90.
- Avants, B., Tustison, N.J., Song, G., 2009. Advanced Normalization Tools: V1.0. *Insight J.*
- B. Fisher, R., Oliver, P., 1998. MultiVariate Cross-Correlation and Image Matching. *Proc. 6th Br. Mach. Vis. Conf. 1995 Vol 1 2.* <https://doi.org/10.5244/C.9.62>
- Bargmann, V., Michel, L., Telegdi, V.L., 1959. Precession of the Polarization of Particles Moving in a Homogeneous Electromagnetic Field. *Phys. Rev. Lett.* 2, 435–436. <https://doi.org/10.1103/PhysRevLett.2.435>
- Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B., 2009. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing, in: *ACM SIGGRAPH 2009 Papers, SIGGRAPH '09.* ACM, New York, NY, USA, p. 24:1–24:11. <https://doi.org/10.1145/1576246.1531330>
- Barnes, J., Foster, J., Boyes, R.G., Pepple, T., Moore, E.K., Schott, J.M., Frost, C., Scahill, R.I., Fox, N.C., 2008. A comparison of methods for the automated calculation of volumes and atrophy rates in the hippocampus. *NeuroImage* 40, 1655–1671. <https://doi.org/10.1016/j.neuroimage.2008.01.012>

## Bibliografía

- Beauchemin, M., Thomson, K.P.B., Edwards, G., 1998. On the Hausdorff Distance Used for the Evaluation of Segmentation Results. *Can. J. Remote Sens.* 24, 3–8. <https://doi.org/10.1080/07038992.1998.10874685>
- Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., 2007. Greedy layer-wise training of deep networks, in: *In Nips*. MIT Press.
- Boccardi, M., Bocchetta, M., Apostolova, L.G., Barnes, J., Bartzokis, G., Corbetta, G., DeCarli, C., deToledo-Morrell, L., Firbank, M., Ganzola, R., Gerritsen, L., Henneman, W., Killiany, R.J., Malykhin, N., Pasqualetti, P., Pruessner, J.C., Redolfi, A., Robitaille, N., Soininen, H., Tolomeo, D., Wang, L., Watson, C., Wolf, H., Duvernoy, H., Duchesne, S., Jack, C.R., Frisoni, G.B., EADC-ADNI Working Group on the Harmonized Protocol for Manual Hippocampal Segmentation, 2015. Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance. *Alzheimers Dement. J. Alzheimers Assoc.* 11, 126–138. <https://doi.org/10.1016/j.jalz.2014.02.009>
- Bogovic, J.A., Bazin, P.-L., Ying, S.H., Prince, J.L., 2013. Automated segmentation of the cerebellar lobules using boundary specific classification and evolution. *Inf. Process. Med. Imaging Proc. Conf.* 23, 62–73.
- Braak, H., Braak, E., 1991. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol. (Berl.)* 82, 239–259.
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brummer, M.E., 1991. Hough transform detection of the longitudinal fissure in tomographic head images. *IEEE Trans. Med. Imaging* 10, 74–81. <https://doi.org/10.1109/42.75613>
- Buades, A., Coll, B., Morel, J.M., 2005. A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Presented at the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 60–65 vol. 2. <https://doi.org/10.1109/CVPR.2005.38>
- Cajal, S.R. y, 2012. *Histología del sistema nervioso del hombre y de los vertebrados. Tomo II - Segunda parte*, Edición: 1. ed. Agencia Estatal Boletín Oficial del Estado, Madrid.
- Caldairou, B., Bernhardt, B.C., Kulaga-Yoskovitz, J., Kim, H., Bernasconi, N., Bernasconi, A., 2016. A Surface Patch-Based Segmentation Method for Hippocampal Subfields, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Lecture Notes in Computer Science*. Presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, pp. 379–387. [https://doi.org/10.1007/978-3-319-46723-8\\_44](https://doi.org/10.1007/978-3-319-46723-8_44)
- Carass A., Cuzzocreo J.L., Han S., Hernandez-Castillo C.R., Rasser P.E., Ganz M., Beliveau V., Price M., Cardenas V.A., Fein J., Dolz J., Ayed IB., Desrosiers C., Thyreau B., Romero JE., Coupé P., Manjón JV., Fonov VS., Collins DL., Ying SH., Crocetti D., Landman BA., Mostofsky SH., Thompson PA., Prince JL. Comparing

## Bibliografía

fully automated state-of-the-art cerebellum parcellation from magnetic resonance images, 2018, Submitted to Neuroimage journal.

- Chakravarty, M.M., Steadman, P., van Eede, M.C., Calcott, R.D., Gu, V., Shaw, P., Raznahan, A., Collins, D.L., Lerch, J.P., 2013. Performing label-fusion-based segmentation using multiple automatically generated templates. *Hum. Brain Mapp.* 34, 2635–2654. <https://doi.org/10.1002/hbm.22092>
- Chupin, M., Gérardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., Benali, H., Garnero, L., Colliot, O., Alzheimer's Disease Neuroimaging Initiative, 2009. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus* 19, 579–587. <https://doi.org/10.1002/hipo.20626>
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* 20, 37–46. <https://doi.org/10.1177/001316446002000104>
- Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. *Hum. Brain Mapp.* 3, 190–208. <https://doi.org/10.1002/hbm.460030304>
- Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18, 192–205.
- Collins, D.L., Pruessner, J.C., 2010. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. *NeuroImage* 52, 1355–1366. <https://doi.org/10.1016/j.neuroimage.2010.04.193>
- Compston, A., Coles, A., 2008. Multiple sclerosis. *Lancet Lond. Engl.* 372, 1502–1517. [https://doi.org/10.1016/S0140-6736\(08\)61620-7](https://doi.org/10.1016/S0140-6736(08)61620-7)
- Coupé, P., Manjón, J.V., Chamberland, M., Descoteaux, M., Hiba, B., 2013. Collaborative patch-based super-resolution for diffusion-weighted images. *NeuroImage* 83, 245–261. <https://doi.org/10.1016/j.neuroimage.2013.06.030>
- Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L., 2011. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage* 54, 940–954. <https://doi.org/10.1016/j.neuroimage.2010.09.018>
- Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C., 2008. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans. Med. Imaging* 27, 425–441. <https://doi.org/10.1109/TMI.2007.906087>
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* 9, 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- Davie, C.A., Barker, G.J., Webb, S., Tofts, P.S., Thompson, A.J., Harding, A.E., McDonald, W.I., Miller, D.H., 1995. Persistent functional deficit in multiple

## Bibliografía

- sclerosis and autosomal dominant cerebellar ataxia is associated with axon loss. *Brain J. Neurol.* 118 ( Pt 6), 1583–1592.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* 39, 1–38.
- Derflinger, S., Sorg, C., Gaser, C., Myers, N., Arsic, M., Kurz, A., Zimmer, C., Wohlschläger, A., Mühlau, M., 2011. Grey-matter atrophy in Alzheimer's disease is asymmetric but not lateralized. *J. Alzheimers Dis. JAD* 25, 347–357. <https://doi.org/10.3233/JAD-2011-110041>
- Dickerson, B.C., Sperling, R.A., 2005. Neuroimaging biomarkers for clinical trials of disease-modifying therapies in Alzheimer's disease. *NeuroRx J. Am. Soc. Exp. Neurother.* 2, 348–360. <https://doi.org/10.1602/neurorx.2.2.348>
- Diedrichsen, J., 2006. A spatially unbiased atlas template of the human cerebellum. *NeuroImage* 33, 127–138. <https://doi.org/10.1016/j.neuroimage.2006.05.056>
- Duchesne, S., Pruessner, J., Collins, D.L., 2002. Appearance-based segmentation of medial temporal lobe structures. *NeuroImage* 17, 515–531.
- Duda, R.O., Hart, P.E., 1973. *Pattern Classification and Scene Analysis*. John Wiley & Sons Inc, New York.
- Dunn, J.C., 1973. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *J. Cybern.* 3, 32–57. <https://doi.org/10.1080/01969727308546046>
- Duvernoy, H.M., Cattin, F., Risold, P.-Y., 2013. *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI*, 4th ed. Springer-Verlag, Berlin Heidelberg.
- E. Hinton, G., 1995. What kind of a graphical model is the brain? *Computer*.
- Eskildsen, S.F., Coupé, P., Fonov, V., Manjón, J.V., Leung, K.K., Guizard, N., Wassef, S.N., Østergaard, L.R., Collins, D.L., Alzheimer's Disease Neuroimaging Initiative, 2012. BEaST: brain extraction based on nonlocal segmentation technique. *NeuroImage* 59, 2362–2373. <https://doi.org/10.1016/j.neuroimage.2011.09.012>
- Giraud, R., Ta, V.-T., Papadakis, N., Manjón, J.V., Collins, D.L., Coupé, P., Alzheimer's Disease Neuroimaging Initiative, 2016. An Optimized PatchMatch for multi-scale and multi-feature label fusion. *NeuroImage* 124, 770–782. <https://doi.org/10.1016/j.neuroimage.2015.07.076>
- Gonzalez, R.C., Woods, R.E., 2007. *Digital Image Processing*, 3 edition. ed. Pearson, Upper Saddle River, NJ.
- Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36. <https://doi.org/10.1148/radiology.143.1.7063747>
- Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A., 2006. Automatic anatomical brain MRI segmentation combining label propagation and

## Bibliografía

- decision fusion. *NeuroImage* 33, 115–126.  
<https://doi.org/10.1016/j.neuroimage.2006.05.061>
- Herrero, M.-T., Barcia, C., Navarro, J.M., 2002. Functional anatomy of thalamus and basal ganglia. *Childs Nerv. Syst. ChNS Off. J. Int. Soc. Pediatr. Neurosurg.* 18, 386–404. <https://doi.org/10.1007/s00381-002-0604-1>
- Hertz, J.A., 1991. *Introduction To The Theory Of Neural Computation*, 1 edition. ed. CRC Press, Cambridge, Mass.
- Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527–1554.  
<https://doi.org/10.1162/neco.2006.18.7.1527>
- Iglesias, J.E., Augustinack, J.C., Nguyen, K., Player, C.M., Player, A., Wright, M., Roy, N., Frosch, M.P., McKee, A.C., Wald, L.L., Fischl, B., Van Leemput, K., Alzheimer's Disease Neuroimaging Initiative, 2015. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *NeuroImage* 115, 117–137.  
<https://doi.org/10.1016/j.neuroimage.2015.04.042>
- Insausti, R., Amaral, D.G., 2012. Chapter 24 - Hippocampal Formation, in: Mai, J.K., Paxinos, G. (Eds.), *The Human Nervous System (Third Edition)*. Academic Press, San Diego, pp. 896–942. <https://doi.org/10.1016/B978-0-12-374236-0.10024-0>
- Jack, C.R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Tangalos, E.G., Kokmen, E., 2000. Rates of Hippocampal Atrophy Correlate with Change in Clinical Status in Aging and AD. *Neurology* 55, 484–489.
- Jack, C.R., Shiung, M.M., Weigand, S.D., O'Brien, P.C., Gunter, J.L., Boeve, B.F., Knopman, D.S., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Petersen, R.C., 2005. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnesic MCI. *Neurology* 65, 1227–1231.  
<https://doi.org/10.1212/01.wnl.0000180958.22678.91>
- James, A.C., James, S., Smith, D.M., Javaloyes, A., 2004. Cerebellar, Prefrontal Cortex, and Thalamic Volumes Over Two Time Points in Adolescent-Onset Schizophrenia. *Am. J. Psychiatry* 161, 1023–1029.  
<https://doi.org/10.1176/appi.ajp.161.6.1023>
- Jones, P.N., McLachlan, G.J., 1990. Algorithm AS 254: Maximum Likelihood Estimation from Grouped and Truncated Data with Finite Normal Mixture Models. *J. R. Stat. Soc. Ser. C Appl. Stat.* 39, 273–282.  
<https://doi.org/10.2307/2347776>
- Kase, C.S., Norrving, B., Levine, S.R., Babikian, V.L., Chodosh, E.H., Wolf, P.A., Welch, K.M., 1993. Cerebellar infarction. Clinical and anatomic observations in 66 cases. *Stroke* 24, 76–83.
- Klein, A., Mensh, B., Ghosh, S., Tourville, J., Hirsch, J., 2005. Mindboggle: Automated brain labeling with multiple atlases. *BMC Med. Imaging* 5, 7.  
<https://doi.org/10.1186/1471-2342-5-7>



## Bibliografía

- Klockgether, T., 2008. The clinical diagnosis of autosomal dominant spinocerebellar ataxias. *Cerebellum Lond. Engl.* 7, 101–105. <https://doi.org/10.1007/s12311-008-0023-2>
- Kriegeskorte, N., Goebel, R., 2001. An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage* 14, 329–346. <https://doi.org/10.1006/nimg.2001.0831>
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS'12*. Curran Associates Inc., USA, pp. 1097–1105.
- Kulaga-Yoskovitz, J., Bernhardt, B.C., Hong, S.-J., Mansi, T., Liang, K.E., van der Kouwe, A.J.W., Smallwood, J., Bernasconi, A., Bernasconi, N., 2015. Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. *Sci. Data* 2, 150059. <https://doi.org/10.1038/sdata.2015.59>
- LeCun, Y., 1985. Une procedure d'apprentissage ponr reseau a seuil asyemetrique. *Proc. Cogn.* 85 599–604.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep Learning. <https://doi.org/10.1038/nature14539>
- Liang, L., Rehm, K., Woods, R.P., Rottenberg, D.A., 2007. Automatic segmentation of left and right cerebral hemispheres from MRI brain volumes using the graph cuts algorithm. *NeuroImage* 34, 1160–1170. <https://doi.org/10.1016/j.neuroimage.2006.07.046>
- Liu, Y., Collins, R.T., Rothfus, W.E., 2001. Robust midsagittal plane extraction from normal and pathological 3-D neuroradiology images. *IEEE Trans. Med. Imaging* 20, 175–192. <https://doi.org/10.1109/42.918469>
- Lötjönen, J.M., Wolz, R., Koikkalainen, J.R., Thurfjell, L., Waldemar, G., Soininen, H., Rueckert, D., 2010. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *NeuroImage* 49, 2352–2365. <https://doi.org/10.1016/j.neuroimage.2009.10.026>
- MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. Presented at the Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, The Regents of the University of California.
- Mangin, J.F., Regis, J., Frouin, V., 1996. Shape bottlenecks and conservative flow systems [medical image analysis], in: *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*. Presented at the Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis, pp. 319–328. <https://doi.org/10.1109/MMBIA.1996.534084>
- Mangin, J.-F., Rivière, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Régis, J., 2004. A framework to study the cortical folding patterns. *NeuroImage* 23 Suppl 1, S129-138. <https://doi.org/10.1016/j.neuroimage.2004.07.019>

## Bibliografía

- Manjón, J.V., Coupé, P., Buades, A., Fonov, V., Louis Collins, D., Robles, M., 2010a. Non-local MRI upsampling. *Med. Image Anal.* 14, 784–792. <https://doi.org/10.1016/j.media.2010.05.010>
- Manjón, J.V., Coupé, P., Martí-Bonmatí, L., Collins, D.L., Robles, M., 2010b. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magn. Reson. Imaging JMRI* 31, 192–203. <https://doi.org/10.1002/jmri.22003>
- Manjón, J.V., Coupé, P., Raniga, P., Xia, Y., Fripp, J., Salvado, O., 2016. HIST: HyperIntensity Segmentation Tool, in: *Patch-Based Techniques in Medical Imaging, Lecture Notes in Computer Science*. Presented at the International Workshop on Patch-based Techniques in Medical Imaging, Springer, Cham, pp. 92–99. [https://doi.org/10.1007/978-3-319-47118-1\\_12](https://doi.org/10.1007/978-3-319-47118-1_12)
- Manjón, J.V., Eskildsen, S.F., Coupé, P., Romero, J.E., Collins, D.L., Robles, M., 2014. Nonlocal Intracranial Cavity Extraction [WWW Document]. *Int. J. Biomed. Imaging*. <https://doi.org/10.1155/2014/820205>
- Manjón, J.V., Tohka, J., Robles, M., 2010c. Improved estimates of partial volume coefficients from noisy brain MRI using spatial context. *NeuroImage* 53, 480–490. <https://doi.org/10.1016/j.neuroimage.2010.06.046>
- Manto, M., Gruol, D.L., Schmahmann, J., Koibuchi, N., Rossi, F. (Eds.), 2013. *Handbook of the Cerebellum and Cerebellar Disorders*. Springer Netherlands.
- Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L., 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.* 19, 1498–1507. <https://doi.org/10.1162/jocn.2007.19.9.1498>
- Milner, B., 1958. Psychological defects produced by temporal lobe excision. *Res. Publ. - Assoc. Res. Nerv. Ment. Dis.* 36, 244–257.
- Nenadic, I., Sauer, H., Gaser, C., 2010. Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. *NeuroImage* 49, 1153–1160. <https://doi.org/10.1016/j.neuroimage.2009.10.014>
- Nyúl, L.G., Udupa, J.K., 1999. On standardizing the MR image intensity scale. *Magn. Reson. Med.* 42, 1072–1081.
- Okugawa, G., Sedvall, G., Nordström, M., Andreasen, N., Pierson, R., Magnotta, V., Agartz, I., 2002. Selective reduction of the posterior superior vermis in men with chronic schizophrenia. *Schizophr. Res.* 55, 61–67.
- Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Parent, A., Carpenter, M.B., 1996. *Carpenter's Human Neuroanatomy*, Subsequent edition. ed. Williams & Wilkins, Baltimore.
- Park, M.T.M., Pipitone, J., Baer, L.H., Winterburn, J.L., Shah, Y., Chavez, S., Schira, M.M., Lobaugh, N.J., Lerch, J.P., Voineskos, A.N., Chakravarty, M.M., 2014. Derivation of high-resolution MRI atlases of the human cerebellum at 3T and

## Bibliografía

- segmentation using multiple automatically generated templates. *NeuroImage* 95, 217–231. <https://doi.org/10.1016/j.neuroimage.2014.03.037>
- Parker, D.B., 1985. *Learning Logic: Casting the Cortex of the Human Brain in Silicon*. Massachusetts Institute of Technology, Center for Computational Research in Economics and Management Science.
- Penny, W., KJ, F., Ashburner, J., J. Kiebel, S., E. Nichols, T., 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. <https://doi.org/10.1016/B978-0-12-372560-8.X5000-1>
- Pham, D.L., Xu, C., Prince, J.L., 2000. Current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–337. <https://doi.org/10.1146/annurev.bioeng.2.1.315>
- Pipitone, J., Park, M.T.M., Winterburn, J., Lett, T.A., Lerch, J.P., Pruessner, J.C., Lepage, M., Voineskos, A.N., Chakravarty, M.M., Alzheimer's Disease Neuroimaging Initiative, 2014. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. *NeuroImage* 101, 494–512. <https://doi.org/10.1016/j.neuroimage.2014.04.054>
- Planche, V., Koubiyr, I., Romero, J.E., Manjon, J.V., Coupé, P., Deloire, M., Dousset, V., Brochet, B., Ruet, A., Tourdias, T., 2018. Regional hippocampal vulnerability in early multiple sclerosis: Dynamic pathological spreading from dentate gyrus to CA1. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.23970>
- Pruessner, J.C., Li, L.M., Serles, W., Pruessner, M., Collins, D.L., Kabani, N., Lupien, S., Evans, A.C., 2000. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cereb. Cortex N. Y. N* 1991 10, 433–442.
- Ranzato, M., Poultney, C., Chopra, S., LeCun, Y., 2006. Efficient Learning of Sparse Representations with an Energy-based Model, in: *Proceedings of the 19th International Conference on Neural Information Processing Systems, NIPS'06*. MIT Press, Cambridge, MA, USA, pp. 1137–1144.
- Rohlfing, T., Brandt, R., Menzel, R., Maurer, C.R., Jr., 2003. Segmentation of three-dimensional images using non-rigid registration: Methods and validation with application to confocal microscopy images of bee brains, in: *Medical Imaging: Image Processing*. pp. 363–374.
- Romero, J.E., Coupé, P., Giraud, R., Ta, V.-T., Fonov, V., Park, M.T.M., Chakravarty, M.M., Voineskos, A.N., Manjón, J.V., 2017a. CERES: A new cerebellum lobule segmentation method. *NeuroImage* 147, 916–924. <https://doi.org/10.1016/j.neuroimage.2016.11.003>
- Romero, J.E., Coupé, P., Manjón, J.V., 2017b. HIPS: A new hippocampus subfield segmentation method. *NeuroImage* 163, 286–295. <https://doi.org/10.1016/j.neuroimage.2017.09.049>
- Romero, J.E., Manjón, J.V., Tohka, J., Coupé, P., Robles, M., 2015. NABS: non-local automatic brain hemisphere segmentation. *Magn. Reson. Imaging* 33, 474–484. <https://doi.org/10.1016/j.mri.2015.02.005>

## Bibliografía

- Rousseau, F., Habas, P.A., Studholme, C., 2011. A supervised patch-based approach for human brain labeling. *IEEE Trans. Med. Imaging* 30, 1852–1862. <https://doi.org/10.1109/TMI.2011.2156806>
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323, 533–536. <https://doi.org/10.1038/323533a0>
- Sakka, L., Coll, G., Chazal, J., 2011. Anatomy and physiology of cerebrospinal fluid. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* 128, 309–316. <https://doi.org/10.1016/j.anorl.2011.03.002>
- Sanroma, G., Benkarim, O.M., Piella, G., Camara, O., Wu, G., Shen, D., Gispert, J.D., Molinuevo, J.L., González Ballester, M.A., 2018. Learning non-linear patch embeddings with neural networks for label fusion. *Med. Image Anal.* 44, 143–155. <https://doi.org/10.1016/j.media.2017.11.013>
- Schapire, R.E., 1990. The strength of weak learnability. *Mach. Learn.* 5, 197–227. <https://doi.org/10.1007/BF00116037>
- Schmajuk, N.A., 1990. Role of the hippocampus in temporal and spatial navigation: an adaptive neural network. *Behav. Brain Res.* 39, 205–229.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17, 87–97. <https://doi.org/10.1109/42.668698>
- Snoek, C.G.M., Worring, M., Smeulders, A.W.M., 2005. Early Versus Late Fusion in Semantic Video Analysis, in: *Proceedings of the 13th Annual ACM International Conference on Multimedia, MULTIMEDIA '05*. ACM, New York, NY, USA, pp. 399–402. <https://doi.org/10.1145/1101149.1101236>
- Stoodley, C.J., 2012. The cerebellum and cognition: evidence from functional imaging studies. *Cerebellum Lond. Engl.* 11, 352–365. <https://doi.org/10.1007/s12311-011-0260-7>
- Ta, V.-T., Giraud, R., Collins, D.L., Coupé, P., 2014. Optimized patchMatch for near real time and accurate label fusion. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* 17, 105–112.
- Thomann, P.A., Schläfer, C., Seidl, U., Santos, V.D., Essig, M., Schröder, J., 2008. The cerebellum in mild cognitive impairment and Alzheimer's disease - a structural MRI study. *J. Psychiatr. Res.* 42, 1198–1202. <https://doi.org/10.1016/j.jpsychires.2007.12.002>
- Thompson, P.M., Moussai, J., Zohoori, S., Goldkorn, A., Khan, A.A., Mega, M.S., Small, G.W., Cummings, J.L., Toga, A.W., 1998. Cortical variability and asymmetry in normal aging and Alzheimer's disease. *Cereb. Cortex N. Y. N* 1991 8, 492–509.
- Timmann, D., Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., Kolb, F.P., 2010. The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex J. Devoted Study Nerv. Syst. Behav.* 46, 845–857. <https://doi.org/10.1016/j.cortex.2009.06.009>

## Bibliografía

- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L.L., Augustinack, J., Dickerson, B.C., Golland, P., Fischl, B., 2009. Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. *Hippocampus* 19, 549–557. <https://doi.org/10.1002/hipo.20615>
- Vergara, I.A., Norambuena, T., Ferrada, E., Slater, A.W., Melo, F., 2008. StAR: a simple tool for the statistical comparison of ROC curves. *BMC Bioinformatics* 9, 265. <https://doi.org/10.1186/1471-2105-9-265>
- Wang, H., Das, S.R., Suh, J.W., Altinay, M., Pluta, J., Craige, C., Avants, B., Yushkevich, P.A., Alzheimer's Disease Neuroimaging Initiative, 2011. A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. *NeuroImage* 55, 968–985. <https://doi.org/10.1016/j.neuroimage.2011.01.006>
- Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A., 2013. Multi-Atlas Segmentation with Joint Label Fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 611–623. <https://doi.org/10.1109/TPAMI.2012.143>
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612. <https://doi.org/10.1109/TIP.2003.819861>
- Weier, K., Beck, A., Magon, S., Amann, M., Naegelin, Y., Penner, I.K., Thürling, M., Aurich, V., Derfuss, T., Radue, E.-W., Stippich, C., Kappos, L., Timmann, D., Sprenger, T., 2012. Evaluation of a new approach for semi-automatic segmentation of the cerebellum in patients with multiple sclerosis. *J. Neurol.* 259, 2673–2680. <https://doi.org/10.1007/s00415-012-6569-4>
- Weier, K., Fonov, V., Lavoie, K., Doyon, J., Collins, D.L., 2014. Rapid automatic segmentation of the human cerebellum and its lobules (RASCAL)--implementation and application of the patch-based label-fusion technique with a template library to segment the human cerebellum. *Hum. Brain Mapp.* 35, 5026–5039. <https://doi.org/10.1002/hbm.22529>
- Werbos, P.J., 1975. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Harvard University.
- Winterburn, J.L., Pruessner, J.C., Chavez, S., Schira, M.M., Lobaugh, N.J., Voineskos, A.N., Chakravarty, M.M., 2013. A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. *NeuroImage* 74, 254–265. <https://doi.org/10.1016/j.neuroimage.2013.02.003>
- Xiao, Y., Fonov, V.S., Beriault, S., Gerard, I., Sadikot, A.F., Pike, G.B., Collins, D.L., 2015. Patch-based label fusion segmentation of brainstem structures with dual-contrast MRI for Parkinson's disease. *Int. J. Comput. Assist. Radiol. Surg.* 10, 1029–1041. <https://doi.org/10.1007/s11548-014-1119-4>
- Yang, Z., Ye, C., Bogovic, J.A., Carass, A., Jodynak, B.M., Ying, S.H., Prince, J.L., 2016. Automated cerebellar lobule segmentation with application to cerebellar

## Bibliografía

- structural analysis in cerebellar disease. *NeuroImage* 127, 435–444. <https://doi.org/10.1016/j.neuroimage.2015.09.032>
- Yushkevich, P.A., Amaral, R.S.C., Augustinack, J.C., Bender, A.R., Bernstein, J.D., Boccardi, M., Bocchetta, M., Burggren, A.C., Carr, V.A., Chakravarty, M.M., Chételat, G., Daugherty, A.M., Davachi, L., Ding, S.-L., Ekstrom, A., Geerlings, M.I., Hassan, A., Huang, Y., Iglesias, J.E., La Joie, R., Kerchner, G.A., LaRocque, K.F., Libby, L.A., Malykhin, N., Mueller, S.G., Olsen, R.K., Palombo, D.J., Parekh, M.B., Pluta, J.B., Preston, A.R., Pruessner, J.C., Ranganath, C., Raz, N., Schlichting, M.L., Schoemaker, D., Singh, S., Stark, C.E.L., Suthana, N., Tompary, A., Turowski, M.M., Van Leemput, K., Wagner, A.D., Wang, L., Winterburn, J.L., Wisse, L.E.M., Yassa, M.A., Zeineh, M.M., Hippocampal Subfields Group (HSG), 2015a. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: towards a harmonized segmentation protocol. *NeuroImage* 111, 526–541. <https://doi.org/10.1016/j.neuroimage.2015.01.004>
- Yushkevich, P.A., Avants, B.B., Pluta, J., Das, S., Minkoff, D., Mechanic-Hamilton, D., Glynn, S., Pickup, S., Liu, W., Gee, J.C., Grossman, M., Detre, J.A., 2009. A high-resolution computational atlas of the human hippocampus from postmortem magnetic resonance imaging at 9.4 T. *NeuroImage* 44, 385–398. <https://doi.org/10.1016/j.neuroimage.2008.08.042>
- Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage* 31, 1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015>
- Yushkevich, P.A., Pluta, J.B., Wang, H., Xie, L., Ding, S.-L., Gertje, E.C., Mancuso, L., Klot, D., Das, S.R., Wolk, D.A., 2015b. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Hum. Brain Mapp.* 36, 258–287. <https://doi.org/10.1002/hbm.22627>
- Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., Shen, D., 2015. Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation. *NeuroImage* 108, 214–224. <https://doi.org/10.1016/j.neuroimage.2014.12.061>
- Zhao, L., Ruotsalainen, U., Hirvonen, J., Hietala, J., Tohka, J., 2010. Automatic cerebral and cerebellar hemisphere segmentation in 3D MRI: adaptive disconnection algorithm. *Med. Image Anal.* 14, 360–372. <https://doi.org/10.1016/j.media.2010.02.001>
- Zijdenbos, A.P., Dawant, B.M., Margolin, R.A., Palmer, A.C., 1994. Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans. Med. Imaging* 13, 716–724. <https://doi.org/10.1109/42.363096>