



**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA**



ALGORITMO DE PARCELACIÓN CEREBRAL INTERSUJETO BASADO EN DATOS DEL PROYECTO DEL CONECTOMA HUMANO

POR

Martín Antonio López Aguilera

Informe Final de Memoria de Título presentada a la Facultad de Ingeniería de la Universidad de Concepción para optar al grado académico de Ingeniera Civil Biomédica

Profesoras guías
Pamela Guevara Alvez,
Cecilia Hernández Rivas

Comisión
Miguel Figueroa Toro

Abril, 2025.
Concepción
(Chile)

© 2025 Martín Antonio López Aguilera

© 2025 Nombre del alumno

Esta memoria puede ser reproducida, distribuida y comunicada públicamente, por cualquier medio o procedimiento, siempre que se cite adecuadamente al autor y no se haga un uso comercial de la misma.



Agradecimientos

Quiero expresar mi más profundo agradecimiento a todas las personas que hicieron posible la realización de este trabajo.

A **Joaquín Molina, Cristóbal Mendoza, Claudio Román, Martín Stockle y Yarelis Prieto**, por su valioso apoyo, colaboración y orientación a lo largo del desarrollo de este proyecto. Su experiencia y disposición fueron fundamentales para avanzar en cada etapa.

A la profesora **Pamela Guevara**, por su guía experta, constante apoyo académico y por liderar con excelencia el *Grupo de Análisis de Imágenes Médicas de la Universidad de Concepción*, espacio donde tuve la oportunidad de aprender y desarrollarme como profesional y futuro investigador.

Agradezco también al Proyecto ANID-FONDECYT Regular 1221665, titulado "*A method for multiscale parcellation of the cortex based on structural connectivity with application to functional variability and alteration in different clinical conditions*", por haber financiado y respaldado este trabajo de investigación.

Finalmente, mi más sincero agradecimiento a mi madre, **Susan Aguilera**, por su amor incondicional, esfuerzo incansable y por haber hecho posible toda mi formación académica. Este logro también es suyo.

Resumen

En ese estudio se desarrolla una metodología para la parcelación cortical de la corteza cerebral humana basada en Imágenes de Resonancia Magnética (MRI), con el objetivo de analizar el conectoma humano. El enfoque consiste en la segmentación de la corteza cerebral en regiones con patrones de conectividad estructural, derivados de tractogramas y mallados corticales. Los tractogramas son representaciones tridimensionales de las trayectorias de las fibras cerebrales, y los mallados representan la superficie de la corteza cerebral. Los tractogramas y mallados se calcularon a partir de imágenes de Resonancia Magnética Estructural (sMRI) y Resonancia Magnética por Difusión (dMRI) pertenecientes a la base de datos Human Connectome Project Young Adult (HCP-YA). Estos tienen 3 millones de fibras y 81920 triángulos por hemisferio en sus 100 sujetos, respectivamente.

Como preprocesamiento, se filtran las fibras pequeñas en base a un umbral, y se aplica un filtro media móvil para suavizar. Luego, se identifican los clusters representativos de la población de sujetos mediante dos etapas de clustering de fibras. Primero se aplica el algoritmo FFClust a los tractogramas de cada sujeto, obteniendo clusters de fibras intrajeto y su centroide respectivo. Luego, se migran estos centroides a un espacio común, donde se obtienen clusters de centroides al aplicar el algoritmo QuickBundles. Para obtener clusters representativos de la población de 100 sujetos, se seleccionaron aquellos clusters intersujeto que contienen centroides de al menos 88 sujetos. Finalmente, se almacena para cada sujeto los clusters de fibras (obtenidos de FFClust) que corresponden a los clusters de centroides (obtenidos de QuickBundles), obteniendo así sólo los clusters representativos de cada sujeto. Para cada sujeto, se obtuvieron 1025 clusters en promedio, en un rango de 664 a 1053 clusters por sujeto.

Para asignar estas fibras a la corteza, el algoritmo proyecta cada fibra sobre el mallado cortical, identificando el triángulo inicial y final de cada trayectoria. Esto permite etiquetar triángulos con sus respectivos clusters, formando subparcelas. Estas subparcelas están fuertemente superpuestas, lo que significa que tienen triángulos en común. Para solucionar esto, se determina la probabilidad de que un triángulo sea de cierta subparcela, contando cuántas veces una fibra de cierta subparcela cruza el triángulo y su vecindad. Luego, se definen los centros de densidad como aquellos triángulos que tienen mayor probabilidad en cada subparcela. Las subparcelas que contienen cierto traslape entre sus centros de densidad son combinadas como una sola. Finalmente, se calculan las componentes conexas principales de cada subparcela, y se realiza dilatación morfológica iterativa para obtener parcelas finales. La parcelación final incluye 189 parcelas en el hemisferio izquierdo y 139 en el derecho.

Abstract

This study develops a methodology for cortical parcellation of the human cerebral cortex based on Magnetic Resonance Imaging (MRI) to analyze the human connectome. The approach involves segmenting the cerebral cortex into regions with structural connectivity patterns derived from tractograms and cortical meshes. Tractograms are three-dimensional representations of brain fiber trajectories, while meshes represent the cortical surface. Both were computed from Structural MRI (sMRI) and Diffusion MRI (dMRI) data from the Human Connectome Project Young Adult (HCP-YA) database, containing 3 million fibers and 81,920 triangles per hemisphere across 100 subjects.

As a preprocessing step, small fibers were filtered based on a threshold, and a moving average filter was applied for smoothing. Then, representative clusters of the subject population were identified through two clustering stages. First, the FFClust algorithm was applied to each subject's tractograms, obtaining intra-subject fiber clusters and their respective centroids. These centroids were then migrated to a common space, where QuickBundles clustering was applied to group them. To select population-representative clusters, only those containing centroids from at least 88 subjects were retained. Finally, for each subject, the fiber clusters (from FFClust) corresponding to these centroid clusters (from QuickBundles) were stored, resulting in an average of 1,025 clusters per subject, ranging from 664 to 1,053.

To assign these fibers to the cortex, the algorithm projects each fiber onto the cortical mesh, identifying the initial and final triangle of each trajectory. This allows labeling triangles with their respective clusters, forming subparcels. These subparcels exhibit significant overlap, meaning they share common triangles. To address this, the probability of a triangle belonging to a subparcel is determined by counting how often a fiber from a given subparcel crosses the triangle and its neighborhood. Density centers are then defined as the triangles with the highest probability within each subparcel. Subparcels with overlapping density centers are merged. Finally, the largest connected components of each subparcel are computed, and an iterative morphological dilation process is applied to obtain the final parcels. The final parcellation includes 189 parcels in the left hemisphere and 139 in the right.

Tabla de contenidos

CAPITULO 1. INTRODUCCIÓN.....	9
1.1 INTRODUCCIÓN GENERAL	9
1.1.1 <i>Objetivo General</i>	11
1.1.2 <i>Objetivos Específicos</i>	11
1.2 ALCANCES Y LIMITACIONES	12
CAPITULO 2. REVISIÓN BIBLIOGRÁFICA.....	13
2.1 INTRODUCCIÓN	13
2.2 BASES BIOLÓGICAS	13
2.2.1 <i>Cerebro Humano</i>	13
2.2.2 <i>Tejido Nervioso</i>	14
2.2.3 <i>Fibras Cerebrales</i>	15
2.2.4 <i>Corteza Cerebral</i>	16
2.3 CONCEPTOS DE RESONANCIA MAGNÉTICA	17
2.3.1 <i>Resonancia Magnética Estructural (sMRI)</i>	17
2.3.2 <i>Generación de Mallados Corticales de la GM</i>	18
2.3.3 <i>Resonancia Magnética de Difusión (dMRI)</i>	19
2.4 ALGORITMOS DE TRACTOGRAFÍA	23
2.5 ALGORITMOS DE CLUSTERING DE FIBRAS	25
2.5.1 <i>Algoritmos de Agrupamiento o Clustering</i>	25
2.5.2 <i>FFClust</i>	26
2.5.3 <i>QuickBundles</i>	29
2.6 PARCELACIÓN CEREBRAL.....	30
2.6.1 <i>Criterios de Parcelación basado en MRI</i>	30
2.6.2 <i>Análisis del conectoma en poblaciones de sujetos basados en tractografía</i>	31
2.6.3 <i>La necesidad de parcelaciones con distintos niveles de granularidad</i>	31
2.6.4 <i>Estado del Arte</i>	32
CAPITULO 3. MATERIALES Y MÉTODOS	37
3.1 INTRODUCCIÓN	37

3.2	MATERIALES	37
3.2.1	<i>Datos de Tractografías</i>	37
3.2.2	<i>Obtención de Mallados Corticales</i>	38
3.3	MÉTODOS	38
3.3.1	<i>Obtención de fascículos representativos de la población de sujetos</i>	38
3.3.2	<i>Aplicación de Algoritmo de Intersección entre Mallados y Clusters de Fibras</i>	42
3.3.3	<i>Formación de subparcelas preliminares</i>	45
3.3.4	<i>Alternativa al traslape</i>	46
3.3.5	<i>Comparación con Parcelaciones del Estado del Arte</i>	49
CAPITULO 4. RESULTADOS		51
4.1	INTRODUCCIÓN	51
4.2	CREACIÓN DEL ATLAS REPRESENTATIVO DE FIBRAS	51
4.2.1	<i>Preprocesamientos y Clustering Intrasujeto</i>	51
4.2.2	<i>Clustering Intersujeto</i>	54
4.2.3	<i>Creación de Parcelación Cortical</i>	56
4.2.4	<i>Comparación de Parcelaciones con el Estado del Arte</i>	61
CAPITULO 5. DISCUSIÓN, CONCLUSIONES Y TRABAJO FUTURO		68
5.1	INTRODUCCIÓN	68
5.2	DISCUSIÓN	68
5.2.1	<i>Preprocesamientos</i>	68
5.2.2	<i>Respecto a la Creación del Atlas Representativo de Fibras</i>	68
5.2.3	<i>Respecto al Criterio de Manejo del Traslape</i>	69
5.2.4	<i>Respecto a la Dilatación Morfológica</i>	70
5.2.5	<i>Respecto a la Comparación con las Parcelaciones del Estado del Arte</i>	70
5.3	CONCLUSIÓN	71
5.4	TRABAJO FUTURO	72
CAPITULO 5. GLOSARIO		73
CAPITULO 6. REFERENCIAS		74

Índice de Figuras

2.1	Tejidos de WM y GM	14
2.2	Disposición del tejido neuronal y estructura neuronal	15
2.3	Tipos de fibras de WM.	16
2.4	Imagen T1 superpuesta con mallado cortical de la WM	19
2.5	Relaciones anatómicas entre fibras cerebrales y mapas de difusión	20
2.6	DTI en patologías	21
2.7	Anisotropía Fraccional	22
2.8	Tractografía determinística y probabilística	24
2.9	Algoritmo FFClust.	28
2.10	Similitudes fuertes entre matrices SC y rsFC en los SFMs	33
2.11	Esquema del proceso de agrupación jerárquica de Moreno-Dominguez.	35
2.12	Pasos algoritmo de Joaquín Molina	36
3.13	Suavizado de Fibras.	39
3.14	Diagrama del Proceso de Obtención de Clusters Representativos	41
3.15	Esquema del proceso de intersección entre fibras y mallado cortical.	44
3.16	Resumen del procesamiento de subparcelas preliminares para obtención de parcelaciones finales.	49
4.17	Método de codo.	52
4.18	Clusters intrasujeto.	53
4.19	Distribución del tamaño de los clusters intrasujeto.	54
4.20	Atlas de clusters representativos	55
4.21	Distribución del tamaño de los clusters intersujeto.	56
4.22	Distribuciones del tamaño de las subparcelas en cada etapa de procesamiento.	57
4.23	Matrices de traslape entre centros de densidad de un sujeto.	58
4.24	Pasos de la solución al traslape. Parcelas finales.	59
4.25	Parcelas Finales Obtenidas	60
4.26	Evolución de la dilatación.	60

4.27	Comparativa de Parcelas Generadas con Parcelación de Brainnetome del lado izquierdo.	64
4.28	Comparativa de Parcelas Generadas con Parcelación de Brainnetome del lado derecho. . .	64
4.29	Comparativa de Parcelas Generadas con Parcelación de Lefranc del lado izquierdo. . . .	65
4.30	Comparativa de Parcelas Generadas con Parcelación de Lefranc del lado derecho	65
4.31	Comparativa de Parcelas Generadas con Parcelación de Narciso del lado izquierdo. . . .	66
4.32	Comparativa de Parcelas Generadas con Parcelación de Narciso del lado derecho. . . .	66
4.33	Comparativa de Parcelas Generadas con Parcelación de Richards del lado izquierdo. . . .	67
4.34	Comparativa de Parcelas Generadas con Parcelación de Richards del lado derecho. . . .	67

Capítulo 1. Introducción

1.1 Introducción General

Se ha demostrado de manera empírica que la *materia gris* (GM) de la corteza cerebral humana se divide de forma modular, donde se pueden identificar zonas con funcionalidades específicas. Estas zonas se interconectan e interactúan mediante señales de impulso nervioso. El mapa estructural y funcional de estas zonas se conoce como conectoma humano, representando las interacciones entre regiones a distintas escalas.

Existe una fuerte necesidad de crear modelos del conectoma humano con tal de comprender la red de conexiones neuronales y cómo logran establecer las funciones cognitivas, emocionales, motoras y sensitivas del ser humano. La parcelación es el proceso de subdividir el cerebro en un número dado de unidades cerebrales.

La parcelación cerebral es la división del cerebro en regiones distintas, basadas en criterios anatómicos, funcionales o citoarquitectónicos. Es una técnica fundamental para estudiar el conectoma, ya que proporciona una estructura organizada y simplificada del cerebro sobre la cual se pueden analizar las conexiones neuronales, permitiendo crear modelos detallados de cómo se interconectan las distintas parcelas.

Para hacer este proceso más detallado, se han aplicado diversas técnicas de procesamiento de MRI para parcelar el cerebro, las cuales realizan un mapeo completo de las conexiones neuronales entre las zonas de materia gris de la corteza cerebral, ofreciendo una visión integral de cómo estas interactúan tanto estructural como funcionalmente [1].

Los primeros criterios establecidos fueron los citoarquitectónicos, los cuales dividen el cerebro en regiones o módulos según la morfología del tejido neuronal. Esta clasificación se basa en características como los tipos celulares, la densidad celular, la distribución y la organización de las neuronas en cada región [2].

Con el desarrollo de las Imágenes de Resonancia Magnética (MRI), surgieron criterios de parcelación basados en patrones de conectividad entre regiones de la corteza cerebral. Gracias a la Resonancia Magnética por Difusión (dMRI), es posible mapear la materia blanca (WM) al analizar la dirección del

movimiento del agua en sus fibras neuronales [3]. Por otro lado, la resonancia magnética funcional (fMRI) mide la actividad neuronal a través del mapeo de la perfusión en el tejido cerebral, basándose en el contraste dependiente del nivel de oxígeno en la sangre y el spin, parámetros que reflejan la actividad en función del consumo de oxígeno [4]. Finalmente, la resonancia magnética estructural (sMRI) permite medir el contraste entre las respuestas de los distintos tejidos cerebrales a la resonancia magnética. A partir de estos datos, es posible realizar una representación topográfica de la corteza mediante la generación de *mallados corticales*, calculados a partir del contraste observado en las imágenes sMRI.

Las imágenes de sMRI, al contener información morfológica de los tejidos dadas por sus diferencias de contraste, nos permiten segmentar máscaras de la materia gris (GM) y la materia blanca (WM) a partir de segmentación plano por plano de las imágenes que contienen el voxel. Estas máscaras, al muestrearlas en puntos, nos permiten generar mallados corticales con la forma de la superficie de estos tejidos de WM y GM [5].

Por otro lado, las imágenes dMRI almacenan datos sobre la magnitud y la dirección de la difusión del agua en cada voxel del cerebro, a partir de la cual se puede obtener una estimación de la morfología de las neuronas largas que componen el tejido de WM. Esta información se recopila utilizando secuencias que aplican gradientes de campo magnético en diferentes direcciones y niveles de intensidad, lo que permite medir cómo se difunde el agua en las diferentes partes del tejido cerebral [6]. Para caracterizar y modelar la difusión de agua en el cerebro, se utilizan modelos de difusión para poder obtener modelos 3D llamados *tractogramas*. A este proceso de pasar de imagen dMRI a tractograma se le conoce como *tractografía*. Los tractogramas usualmente consisten en conjuntos de puntos de 3 dimensiones que representarían una fibra de WM mediante una línea en el espacio.

En base a estos criterios, es posible encontrar distintas formas de realizar parcelaciones. Los *algoritmos de parcelación cerebral* (ADP) permiten diferenciar y definir las regiones de interés de la corteza cerebral en función de los criterios ya mencionados mediante el procesamiento automático de los distintos tipos de MRI.

El *proyecto del conectoma humano* (HCP) es una iniciativa que busca crear mapas detallados de la conectividad estructural y funcional del cerebro humano en poblaciones de sujetos. Para ello, se sometieron varios sujetos a técnicas de neuroimagen distintas con el fin de recopilar información sobre la anatomía y la función cerebral a diferentes escalas espaciales y temporales. Dentro del proyecto, se tienen 20 estudios distintos del conectoma humano. Uno de estos estudios consiste en el de una población de 1200 sujetos sanos, de los cuales se tienen imágenes sMRI y dMRI de alta calidad. Este estudio lleva por nombre *Human Connectome Project Young Adult HCP-YA* [7], y es de acceso público.

En el presente proyecto, se proponen mejoras al algoritmo propuesto trabajado en la tesis de Joaquín Molina [8]. Se realiza la aplicación del ADP que se propone en esta tesis a 100 sujetos del HCP-YA [7], logrando obtener clusters de fibras intersujeto y parcelaciones basadas en dichos clusters. El algoritmo aplicado se basa en la topografía de la GM y la conectividad entre las regiones de la corteza. Se parte con 100 mallados corticales calculados de imágenes sMRI del HCP-YA [7], calculados por Cristóbal Mendoza y 100 tractogramas generados a partir de imágenes dMRI de múltiples capas (multi-shell) con modelo de difusión de alta resolución (HARDI), utilizando algoritmos de tractografía probabilística para obtener el tractograma. Estos tractogramas se tienen gracias al trabajo de Claudio Román [6].

Para generar la parcelación, se aplican algoritmos de clustering de fibras al tractograma obteniendo clusters intra e inter sujeto. Se realiza la intersección de los clusters representativos de la población de sujetos con los triángulos del mallado, generando así subparcelas preliminares y perfiles de conectividad, los cuales representan la conexión de una parcela con todas las demás. Para el análisis intra-sujeto se utiliza FFClust [9] y para el análisis inter-sujeto se utiliza [10]. Finalmente, se aplican postprocesamientos para obtener parcelas homogéneas con el fin de solucionar el traslape entre parcelas. Esta parte se inspiró en el trabajo de Narciso López [11].

1.1.1 Objetivo General

Desarrollar un algoritmo de parcelación cortical basado en conectividad estructural de las fibras de WM, utilizando datos de tractografía probabilística y mallados de la superficie cortical provenientes de 100 sujetos de la base de datos HCP-YA.

1.1.2 Objetivos Específicos

- Objetivo 1.

Aplicar algoritmo de clustering intra-sujeto (FFClust) a los 100 sujetos de estudio, seguido de un clustering inter-sujeto (QuickBundles) aplicado al conjunto de centroides de los clusters intra-sujeto, para obtener un atlas de fibras representativo de la población de sujetos.

- Objetivo 2.

Implementar e identificar mejoras en la robustez de la solución al traslape y dilatación morfológica, propuesta por Narciso López en [11].

- Objetivo 3.

Obtener parcelaciones representativas de la población de sujetos, analizando cuantitativamente cada paso para la futura creación de un nuevo algoritmo de parcelación jerárquica multiescala.

1.2 Alcances y Limitaciones

En tanto a los alcances, los códigos serán desarrollados en Python, C y C++, según los requerimientos que se tengan en el momento de la implementación. Además, se cuenta con los scripts de trabajos previos de Joaquín Molina y Yarelis Prieto [12], [13], que ayudarán a montar el experimento descrito, correspondientes al análisis intra-sujeto e inter-sujeto. Estos scripts serán probados y ejecutados en sistema operativo Ubuntu 22. Para la visualización de los datos se usará el software Anatomist, BrainVisa [14] y Visualization Toolkit (VTK) [15].

En tanto a las limitaciones, la utilización de datos de tractografía probabilística multishell en vez de tractografías determinísticas single shell significa un aumento de los posibles recorridos de las fibras de WM al considerar un mayor número de trayectorias desde la semilla de seguimiento, pero también un aumento significativo del ruido en los datos. El uso de algoritmos de tractografía probabilística al considerar un mayor rango de posibilidades genera una mayor cantidad de fibras, lo cual es bastante útil para poder tener fascículos más prominentes y cercanos a estructuras definidas, aunque también nos puede llevar a generar falsos positivos (fibras que no existen) o falsos negativos (fibras reales que no se detectan).

Otra limitante a tomar en cuenta es la variabilidad anatómica entre individuos, ya que puede afectar la precisión y la generalización del algoritmo. Es importante considerar cómo manejar esta variabilidad para garantizar la robustez y la aplicabilidad del método en diferentes poblaciones. Además, la tractografía en sí es una técnica que es bastante sensible al movimiento durante la toma de datos, lo que disminuye la confiabilidad de sus resultados.

Por último, se puede considerar como una limitante intrínseca a los datos de tractografía al hecho de que no se pueda interpretar a partir de esta misma si corresponden a fibras aferentes o eferentes. Este análisis sólo puede llegar a concluirse si hacemos relación con datos funcionales.

Capítulo 2. Revisión Bibliográfica.

2.1 Introducción

En esta sección, se verán los conceptos básicos para poder entender el proyecto realizado y se revisará el estado del arte en el área a investigar, correspondiente a los ADP. Primero, se muestran las bases biológicas básicas a comprender. Luego, se explican los conceptos de MRI necesarios para poder entender los protocolos de adquisición de la base de datos inicial. Seguido a ello, se explican los algoritmos de tractografía determinística y probabilística. Posteriormente, se describen los algoritmos de agrupamiento de fibras utilizados. Finalmente, define una de parcelación, las necesidades de la comunidad científica en este ámbito y un análisis del estado del arte.

2.2 Bases Biológicas

2.2.1 Cerebro Humano

El cerebro humano, el órgano central del sistema nervioso, es una estructura altamente compleja y multifuncional que gobierna la mayoría de las actividades del organismo. Este está compuesto por dos hemisferios, derecho e izquierdo. Estos se conectan a partir de una masa de WM conocida como el cuerpo calloso. La corteza cerebral que rodea los hemisferios está hecha de GM. Por su parte, la GM está formada principalmente por cuerpos celulares neuronales y dendritas, así como por capas externas de la corteza cerebral. Esta región es responsable de procesar y coordinar la información sensorial, controlar los movimientos musculares voluntarios y regular las funciones cognitivas como el pensamiento, la memoria y el aprendizaje. Por otro lado, la WM está compuesta principalmente por axones mielinizados, que son las prolongaciones de las neuronas responsables de transmitir señales eléctricas entre diferentes regiones del cerebro y el sistema nervioso periférico. En este contexto, interpretamos la WM como el "cableado" del cerebro, facilitando la comunicación y la coordinación entre distintas áreas cerebrales. Podemos apreciar los relieves de WM y GM en una MRI en la figura 2.1.

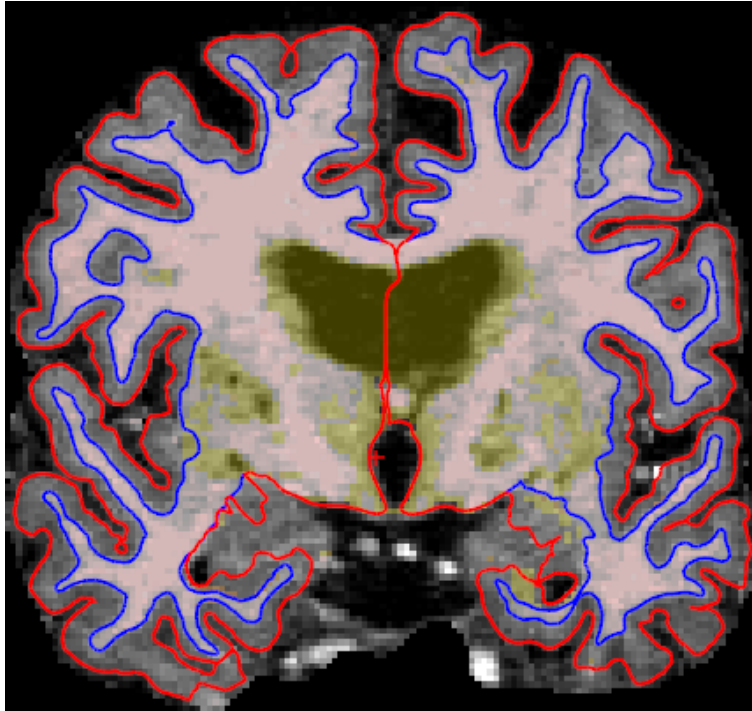


Fig. 2.1: Podemos apreciar una imagen de sMRI con los bordes de los tejidos de WM y GM identificados. En rojo, se puede apreciar la parte externa de la GM, y en azul la capa que separa la GM de la WM. Este proceso se realiza durante el cálculo de mallados corticales cerebrales de la GM mediante el muestreo de puntos de estas máscaras a lo largo de todos los planos de vóxeles de la imagen sMRI en el pipeline de conversión de FreeSurfer. Imagen obtenida de [5].

2.2.2 Tejido Nervioso

El tejido neuronal está compuesto principalmente por neuronas, que son las unidades fundamentales del sistema nervioso. Las neuronas son células especializadas en transmitir información en forma de gradientes electroquímicos. Cada neurona consta de un cuerpo celular que contiene el núcleo y numerosas dendritas que reciben señales de otras neuronas (ver figura 2.2), así como un axón largo que transmite señales a través de sinapsis, concepto que se refiere al medio de comunicación de una neurona con otra, o bien, con otras células, como células neurogliales (que le dan soporte sistémico y estructural a las neuronas) células musculares, endocrinas, y así un largo etcétera. Estas sinapsis permiten la comunicación entre neuronas y son esenciales para el procesamiento de información y la transmisión de señales en el sistema nervioso. El tejido neuronal, formado por la interconexión de miles de millones de estas neuronas, constituye la base de la función cerebral y del sistema nervioso en su conjunto, desempeñando un papel crucial en el control de procesos cognitivos, emocionales y motores en el organismo [16].

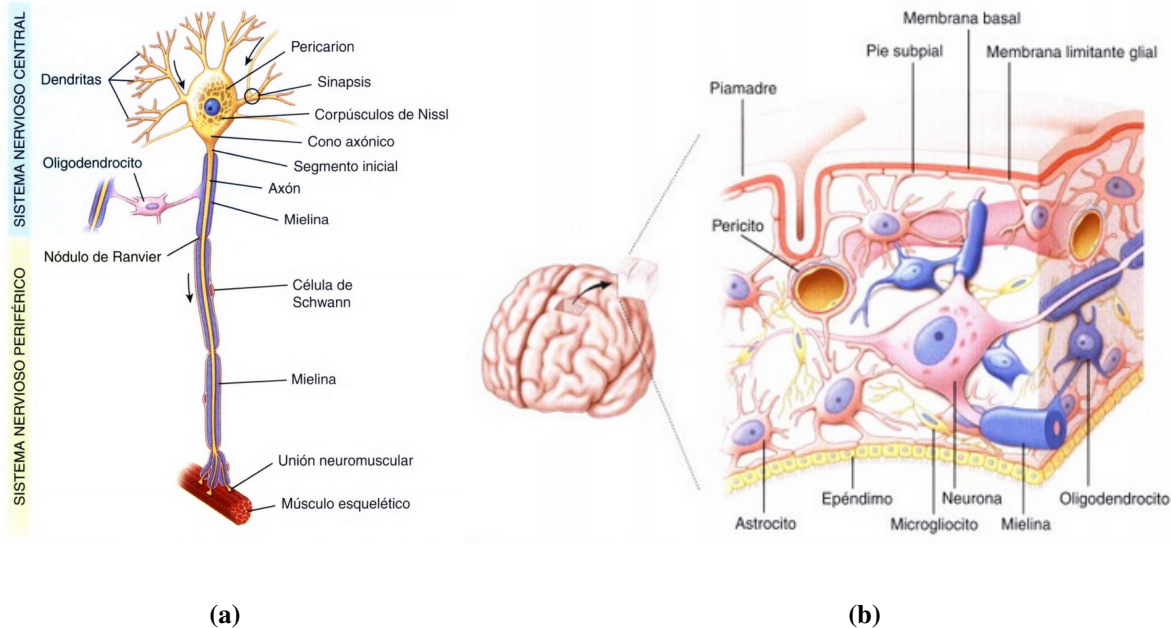


Fig. 2.2: En (a) podemos ver la estructura básica de una neurona motora y sus elementos acompañantes que hacen que esta cumpla su función correctamente. En (b), se aprecian los tres tipos de células gliales: los astrocitos, oligodendrocitos y microglíocitos. Imagen adaptada de [16]

2.2.3 Fibras Cerebrales

El término "fibra cerebral" se refiere a las estructuras anatómicas que componen las conexiones de proyección entre diferentes regiones del cerebro. Estas fibras son axones de neuronas que forman haces o tractos, y su función principal es transmitir señales eléctricas y químicas entre áreas cerebrales distantes. Las *Fibras de Asociación* interconectan áreas corticales en cada hemisferio. Estas se encargan de la comunicación entre regiones cerebrales adyacentes o distantes dentro de un hemisferio y juegan un papel crucial en funciones cognitivas complejas. Un ejemplo de fibras de asociación es el fascículo uncinado, que conecta áreas del lóbulo temporal con áreas del lóbulo frontal. Las *Fibras de Proyección* interconectan áreas corticales con núcleos profundos como el tronco cerebral, cerebelo y médula espinal. Ejemplos de fibras de proyección incluyen las vías corticoespinales, que transmiten señales motoras desde la corteza cerebral hasta la médula espinal, y las vías corticotálamicas, que conectan la corteza cerebral con el tálamo. Por último, las *Fibras Comisurales* son fibras que conectan regiones homólogas entre los dos hemisferios cerebrales, permitiendo la comunicación entre ambos lados del cerebro. Las fibras comisurales cruzan la línea media del cerebro a través de estructuras como el cuerpo calloso, la comisura anterior y la comisura posterior. Ejemplos de fibras comisurales incluyen las fibras del cuerpo

caloso, que conectan áreas corticales similares en los hemisferios derecho e izquierdo del cerebro [17]. A continuación, podemos apreciar de forma gráfica esta clasificación en la figura 2.3.

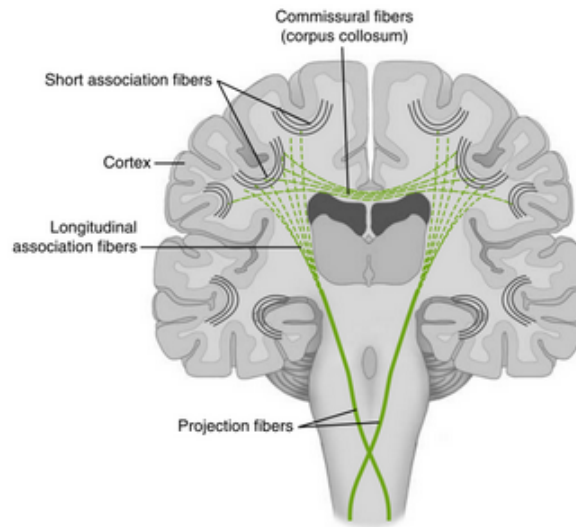


Fig. 2.3: Se puede apreciar las fibras de asociación, las fibras comisurales y las fibras de proyección ya mencionadas. Imagen adaptada de [18]

2.2.4 Corteza Cerebral

La corteza cerebral es la capa más externa del cerebro y una de las estructuras más importantes del sistema nervioso central, ya que se encarga de funciones complejas como la percepción, el pensamiento, el lenguaje y el control voluntario de los movimientos. Tiene un grosor aproximado de 2 a 4 mm y su superficie está repleta de pliegues o circunvoluciones, lo que permite una mayor área de superficie en un espacio reducido, optimizando así la capacidad del cerebro.

La corteza está compuesta principalmente por materia gris, la cual está formada por cuerpos neuronales (somatos), dendritas y terminales axonales, así como células gliales. El tejido está compuesto de neuronas que tienen un tamaño mucho menor en comparación a las de WM. Las neuronas de la corteza cerebral están organizadas en capas horizontales y se agrupan en áreas funcionales específicas que se especializan en diversas funciones cognitivas y motoras. Hay seis capas principales en la corteza, aunque su espesor y composición celular pueden variar dependiendo de la región cortical y su función particular.

Una de las primeras parcelaciones la realizó Korbinian Brodmann [2] parceló el cerebro en base a esta citoarquitectura usando como criterio las características celulares de las capas mencionadas en distintas

zonas del cerebro humano, guiándose también a través de la morfología cerebral.

2.3 Conceptos de Resonancia Magnética

En el ámbito de la neurociencia y la parcelación cerebral, es fundamental comprender ciertos conceptos clave que, tras su aparición, han revolucionado la manera en que exploramos el cerebro humano. En esta sección, se abordarán los conceptos básicos necesarios para entender la metodología empleada en este Proyecto de Investigación.

La resonancia magnética es una técnica de imagen no invasiva que utiliza un campo magnético fuerte, generalmente de 1.5 a 3 Tesla, para alinear los protones en las moléculas de agua del cuerpo. Un pulso de radiofrecuencia específico interrumpe esta alineación, haciendo que los protones giren fuera de la dirección del campo magnético principal, cambiando el spin de sus átomos. Al cesar el pulso de radiofrecuencia, los protones regresan a su alineación original, emitiendo señales de radiofrecuencia. Estas señales son captadas por bobinas del escáner y utilizadas para crear imágenes detalladas de los tejidos blandos del cuerpo mediante algoritmos de reconstrucción, permitiendo así una visualización precisa y en alta resolución de su estructura interna.

2.3.1 Resonancia Magnética Estructural (sMRI)

La sMRI es una técnica de imagen médica que utiliza campos magnéticos y ondas de radio para generar imágenes detalladas de los tejidos blandos y las estructuras internas del cuerpo humano. Esta técnica es especialmente útil para visualizar el cerebro y otras partes del sistema nervioso central, así como también para detectar anomalías en articulaciones, músculos, órganos internos y otros tejidos. La sMRI proporciona imágenes tridimensionales de alta resolución que pueden mostrar con precisión la anatomía interna del cuerpo, lo que la convierte en una herramienta valiosa en el diagnóstico y seguimiento de una amplia gama de condiciones médicas, como tumores cerebrales, lesiones en la médula espinal, enfermedades neurodegenerativas, problemas cardíacos, lesiones musculoesqueléticas, entre otros [19].

Las imágenes sMRI se pueden tomar en varias secuencias distintas, dentro de las cuales destacan T1 (Tiempo de Relajación Longitudinal) y T2 (Tiempo de Relajación Transversal). Estas secuencias se diferencian en cómo responden los tejidos a los campos magnéticos durante el proceso de imagenología. En una imagen T1, los tejidos con mayor contenido de protones de hidrógeno, como la WM y la GM en el cerebro, tienen una señal de intensidad relativamente alta, lo que los hace parecer más brillantes

en la imagen. Esta secuencia es útil para visualizar la anatomía estructural y proporciona buen contraste entre diferentes tipos de tejidos. En una imagen T2, los tejidos con mayor contenido de agua, como el líquido cefalorraquídeo y el edema, tienden a tener una señal de intensidad alta, lo que los hace parecer más brillantes en la imagen. La imagen T2 es útil para detectar anomalías como inflamación, lesiones y líquidos, y también puede proporcionar información sobre la estructura de los tejidos.

2.3.2 Generación de Mallados Corticales de la GM

Un mallado cortical de la superficie cerebral se refiere a la creación de una representación tridimensional de la superficie externa de la corteza cerebral, lo que permite un análisis detallado de su forma y estructura.

Para crear un mallado cerebral, se utilizan secuencias T1 ponderadas, que son particularmente útiles para distinguir entre la GM y WM en el cerebro. La secuencia T1 proporciona un buen contraste entre estos dos tipos de tejidos, lo que facilita la segmentación y reconstrucción de la superficie cortical. Una vez que se ha adquirido la imagen de resonancia magnética T1, se lleva a cabo un procesamiento de imágenes para corregir artefactos, mejorar la calidad de la imagen y alinear adecuadamente las imágenes. Luego, se realiza una segmentación automática de los tejidos cerebrales para distinguir entre GM, WM y líquido cefalorraquídeo, proceso que se refiere al reconocimiento de objetos mediante la diferencia de contraste entre vóxeles vecinos. El siguiente paso es la construcción del modelo de superficie cortical, donde se utilizan algoritmos para reconstruir la superficie cortical a partir de los datos segmentados. Esto implica la creación de un modelo de superficie que representa la frontera entre la GM y WM en la corteza cerebral. Una vez construido el modelo de superficie inicial, se pueden realizar procesos adicionales, como inflado y suavizado, para mejorar la precisión y la calidad del mallado cerebral.

En cuanto al almacenamiento, el mallado cerebral se guarda como un conjunto de datos tridimensionales que contienen información sobre la topología y la geometría de la superficie cortical. Estos datos se pueden almacenar en diferentes formatos, como archivos de malla tridimensional (por ejemplo, formatos OBJ o GII) que contienen la información necesaria para visualizar y analizar el mallado cerebral utilizando software especializado en neuroimagen.

Una vez completado el proceso, se obtiene un mallado cortical detallado y preciso que puede utilizarse para estudiar diversas características de la corteza cerebral, como el grosor cortical, la morfología de las circunvoluciones y la topología de las regiones corticales. Esta información es invaluable para comprender la organización estructural del cerebro humano y su relación con funciones cognitivas y

trastornos neurológicos. Podemos apreciar estos mallados en la figura 2.4.

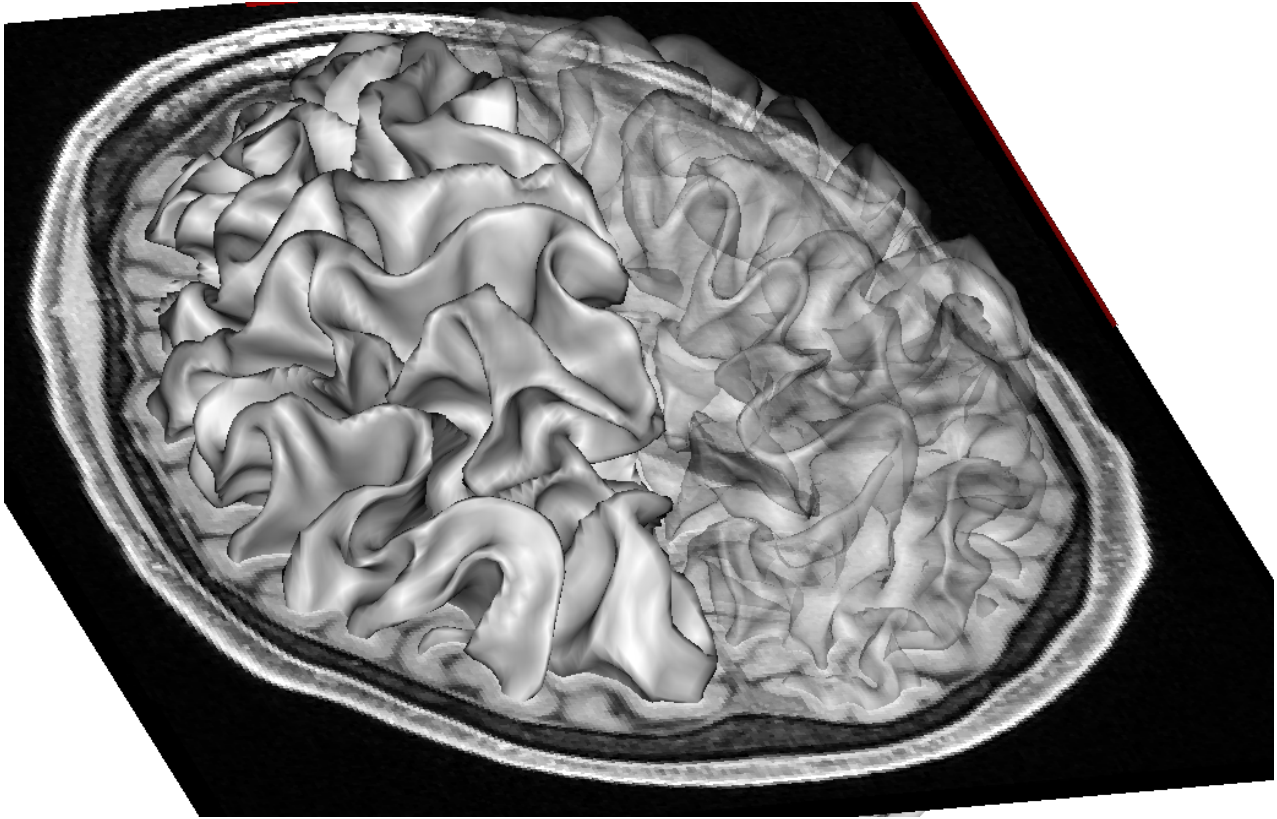


Fig. 2.4: Imagen T1 superpuesta con Mallado Cortical de la WM. Es posible notar que la WM se encuentra completamente alineada con el mallado cortical. El sujeto pertenece a la base de datos HCP-YA y sus datos fueron visualizados con el software Anatomist (Elaboración propia)

2.3.3 Resonancia Magnética de Difusión (dMRI)

La dMRI es una técnica avanzada de imagenología cerebral que se utiliza para visualizar y cuantificar la dirección y la magnitud de la difusión del agua en los tejidos in-vivo de forma no invasiva. Si la aplicamos al cerebro, esta técnica proporciona información detallada sobre la microestructura de la WM y las vías de fibras nerviosas a partir del gradiente de intensidad de campo magnético inducido en las moléculas de agua durante el efecto de resonancia.

En términos de la técnica de adquisición, la dMRI se diferencia de la sMRI en que los gradientes de campo magnético en dMRI se toman en varias direcciones e intensidades, lo que permite medir la magnitud y dirección de la difusión del agua en los tejidos. Este proceso genera imágenes con contraste

de difusión que reflejan la anisotropía de la difusión, es decir, la direccionalidad preferida del movimiento del agua. Las moléculas de agua tienden a moverse con mayor facilidad a lo largo de las fibras nerviosas que a través de ellas, lo que permite mapear la forma de las fibras de WM [20]. La dMRI permite estudiar la organización de las fibras cerebrales, identificar patologías y comprender la conectividad neuronal en el cerebro humano [21]. En la figura 2.5, se pueden apreciar las fibras identificables en un mapa de difusión generado a partir de un *Modelo de Tensor de Difusión* (DTI).

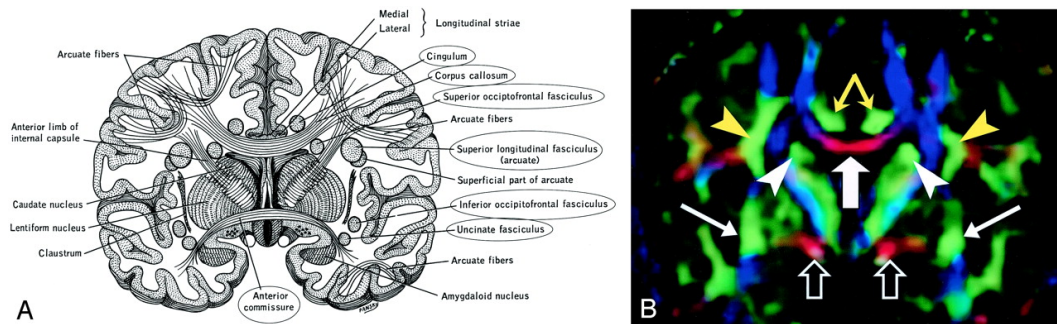


Fig. 2.5: **A**, La ilustración muestra las relaciones anatómicas de varios haces de fibras de WM en el plano coronal. Los haces circulados son aquellos que se ilustran con más detalle en esta revisión. El cuerpo calloso está entre el cíngulo en la posición superomedial y el fascículo occipitofrontal superior en la posición inferolateral. El fascículo longitudinal superior se extiende a lo largo del margen superior del claustrum en un gran arco. El fascículo occipitofrontal inferior se encuentra a lo largo del borde inferolateral del claustrum. **B**, Mapa direccional correspondiente a la imagen A. Los cíngulos emparejados se identifican fácilmente en verde (flechas amarillas) justo cefálicos al cuerpo calloso en rojo (flecha gruesa blanca). Las puntas de flecha blanca indican el fascículo occipitofrontal superior; flechas delgadas blancas, fascículo occipitofrontal inferior; puntas de flecha amarillas, fascículo longitudinal superior. Al igual que el cuerpo calloso, las fibras comisurales de la comisura anterior están orientadas de izquierda a derecha hacia la línea media, lo que da como resultado el característico color rojo (flechas abiertas). Más lateralmente, las fibras divergen y se mezclan con otros haces donde dejan de ser identificables. Imagen disponible en [22]

■ Modelos de Difusión

Los modelos de difusión son métodos que se utilizan para calcular la trayectoria de las fibras a partir de dMRI. Estos modelos calculan la distribución de la dirección o direcciones principales de las fibras contenidas en dicho vóxel. La dirección final estimada por el modelo dependerá de la dirección de las fibras más predominantes de dicho vóxel. La calidad de la reconstrucción dependerá entonces del proceso del método utilizado, habiendo muchos que escoger. Uno de los más utilizados, sobre todo en el rubro clínicos, es el DTI [23], debido a su baja complejidad y velocidad, ya que asume una función de distribución particular con los parámetros calculados, ajustando la señal de difusión con el modelo, asumiendo una distribución gaussiana 3D, calculando el tensor del

modelo como una matriz de covarianza de gaussiana. En la figura 2.6 podemos ver cómo podemos relacionar la DTI con patologías [6].

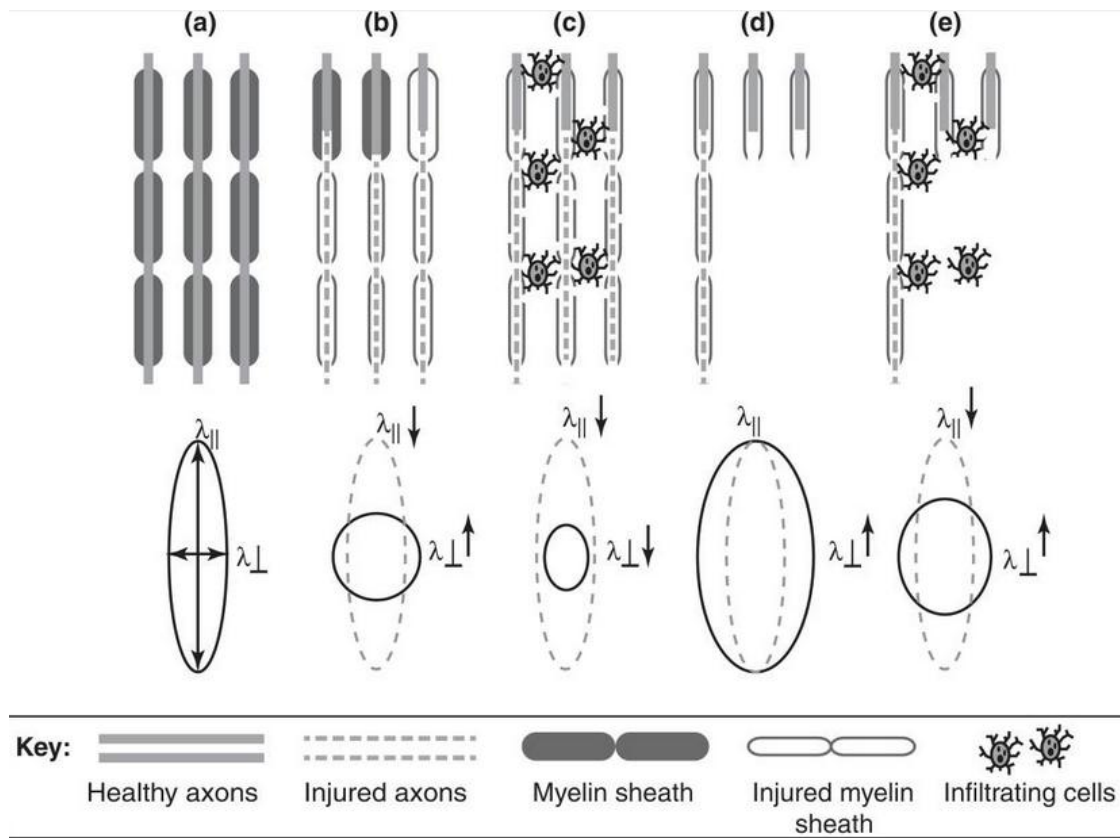


Fig. 2.6: Este diagrama ilustra cómo las neuropatologías afectan los resultados de la DTI. Podemos encontrar en (a) axones normalmente mielinizados, en (b,c) axones y mielina dañados, y en (d,e) axones y mielina dañados con pérdida del axón como tal. (Imagen disponible en [6], [24]).

• Modelo de Tensor de Difusión (DTI)

Propuesto por Basser [23], DTI es un modelo que cuantifica el grado de *Anisotropía Fraccional* (FA), el cual es un índice que mediante los valores propios $\lambda_1, \lambda_2, \lambda_3$ asociados a tres vectores propios ortogonales que representan distribución gaussiana en 3D, como se muestra en la ecuación 2.3.1, donde $\lambda = (\lambda_1 + \lambda_2 + \lambda_3)/3$. λ_1 representa la difusión de agua a lo largo de la longitud de la fibra, siendo este valor siempre el mayor. λ_2 y λ_3 representan la difusión perpendicular de las fibras. Podemos encontrar una representación gráfica de cómo varía este parámetro en función a los tensores dominantes en la figura 2.7.

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \lambda)^2 + (\lambda_2 - \lambda)^2 + (\lambda_3 - \lambda)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (2.3.1)$$

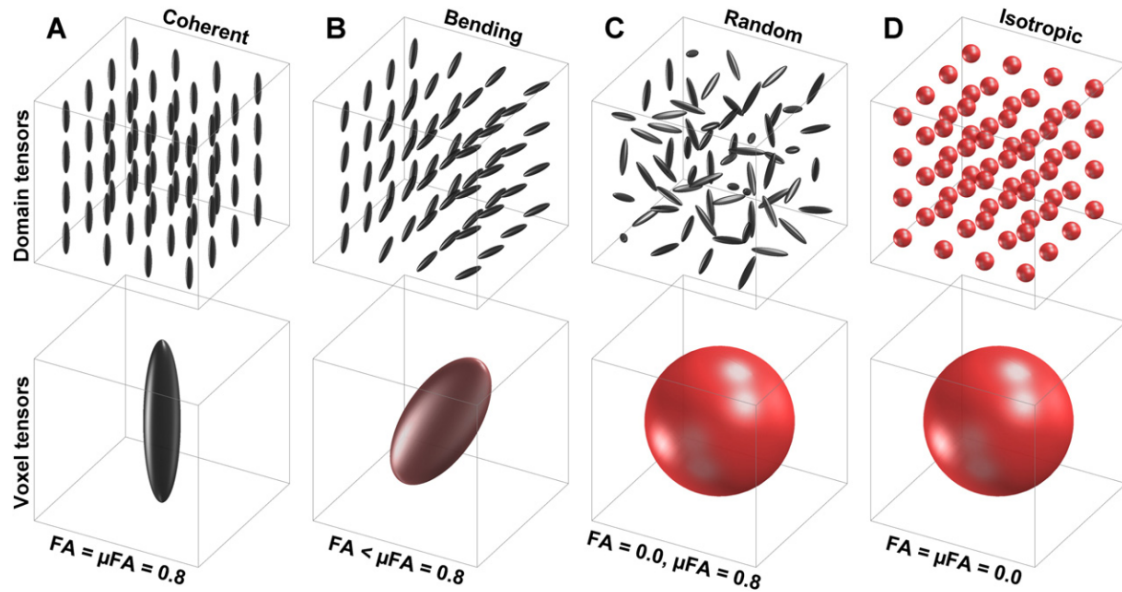


Fig. 2.7: Podemos apreciar como varía la FA en función a los tensores dominantes que representan la dirección de los torrentes de difusión de un vóxel. En algunos tejidos, como el que muestra la imagen A, el agua se mueve principalmente en una dirección, lo que significa que el tejido es bastante ordenado. En otros, como el que muestra la imagen C, el agua se mueve en todas las direcciones, lo que significa que el tejido es más desordenado. La FA es una medida que usamos para describir cuánto orden hay en el movimiento del agua. En la imagen C, donde el tejido es muy desordenado, la FA es igual a cero, lo que significa que no hay orden en absoluto. Pero en la imagen A, donde el tejido es muy ordenado, la FA es alta, lo que significa que hay mucho orden. Sin embargo, la FA no puede decir la diferencia entre tejido desordenado (imagen C) y tejido completamente ordenado en todas las direcciones (imagen D), porque en ambos casos, la FA es cero. (Imagen disponible en [25]).

- **Espectro de Difusión (DSI)**

El *Espectro de Difusión* (DSI) es una representación gráfica que muestra la distribución de la difusión del agua en diferentes direcciones y magnitudes en los tejidos biológicos. En este caso, la difusión a la magnitud del vector de difusión y la dirección del gradiente de campo magnético aplicado durante la adquisición de datos de dMRI [26].

- **Q-ball imaging**

Esta técnica utiliza la transformada de Funk-Randon para calcular la ODF, esto requiere un esquema de muestreo de difusión con forma de Shell, como una adquisición HARDI con un b-value y múltiples direcciones [27].

- **Deconvolución esférica restringida (CSD)**

En la *Deconvolución esférica restringida* (CSD), la señal de difusión se asume como la convolución entre la función de distribución de orientación de las fibras (fODF) y la función de

respuesta de una sola fibra. Entonces, el método comienza calculando la función de respuesta de la fibra y posteriormente se realiza una deconvolución entre la señal medida y la señal de respuesta de la fibra para obtener la fODF, logrando una mejor resolución angular [6].

■ **Imágenes de alta resolución angular (HARDI)**

Las *imágenes de alta resolución angular* (HARDI) son un protocolo de imágenes en el cual el espacio de difusión en 3D o q-space (que representa la magnitud y la dirección de los gradientes de campo magnético) es muestreado a lo largo de una gran cantidad de direcciones para poder aumentar las direcciones que puede llegar a tomar la difusión y así aumentar el rango de posibilidades de la distribución [28]. A diferencia de los otros protocolos de dMRI que se suelen usar con DTI, este toma más muestras en distintas direcciones y, por tanto, su adquisición es más lenta.

Podemos clasificar las imágenes HARDI dependiendo de los esquemas de muestreo que presenten en el q-space, como sigue [6].

1. *Single-Shell HARDI*: En este enfoque, se adquieren datos de difusión utilizando gradientes de difusión en una única dirección o cascarón esférico. Aunque permite una adquisición más rápida y simple de datos, tiene limitaciones en la capacidad para modelar la complejidad de la difusión en tejidos con múltiples orientaciones de fibra, esto debido a que las muestras se sacan a partir de una esfera única en el q-space.
2. *Multi-Shell HARDI*: En este enfoque, se utilizan gradientes de difusión en múltiples direcciones o cascarones esféricos con diferentes b-values (valores de ponderación de difusión). Esto permite una mayor sensibilidad para detectar y caracterizar la difusión en tejidos con diferentes orientaciones de fibra y propiedades microestructurales. La adquisición de datos en múltiples cascarones también facilita la estimación de modelos de difusión más avanzados, como el modelo de difusión de múltiples compartimentos, que pueden proporcionar información adicional sobre la organización microestructural de los tejidos, sin embargo esto requiere una mayor cantidad de imágenes.

2.4 Algoritmos de Tractografía

Los algoritmos de tractografía estiman la trayectoria de las fibras de WM con la información de difusión de agua en cada voxel, tomando la componente principal de los tractos de las fibras. Los modelos de difusión (DTI, DSI, Q-ball, CSD y otros) proporcionan la información sobre la función de distribución

de orientación de la difusión del agua en cada voxel, que luego se utiliza en la tractografía para estimar las trayectorias de las fibras neuronales en el cerebro.

Existen dos enfoques de tractografía, el enfoque probabilístico y el determinista. Estos se diferencian en la manera en que cada enfoque aborda la incertidumbre inherente en la estimación de las trayectorias de las fibras neuronales en la tractografía. En la tractografía determinista, se sigue una única dirección de difusión predominante en cada paso de la trayectoria, lo que puede conducir a la detección de trayectorias de fibras más claras y reproducibles, pero que también pueden pasar por alto áreas de cruce de fibras o zonas con dispersión de la difusión. Por otro lado, la tractografía probabilística modela esta incertidumbre al permitir que múltiples direcciones de difusión sean exploradas en cada paso a partir del muestreo de las funciones de distribución de probabilidad que modelan las trayectorias, generando así un conjunto de posibles trayectorias que reflejan la variabilidad, teniendo en cuenta la incertidumbre del proceso. Este enfoque probabilístico proporciona una representación más completa de las vías de las fibras en áreas de cruce y dispersión, aunque puede requerir más tiempo computacional y puede generar mayor variabilidad y ruido.

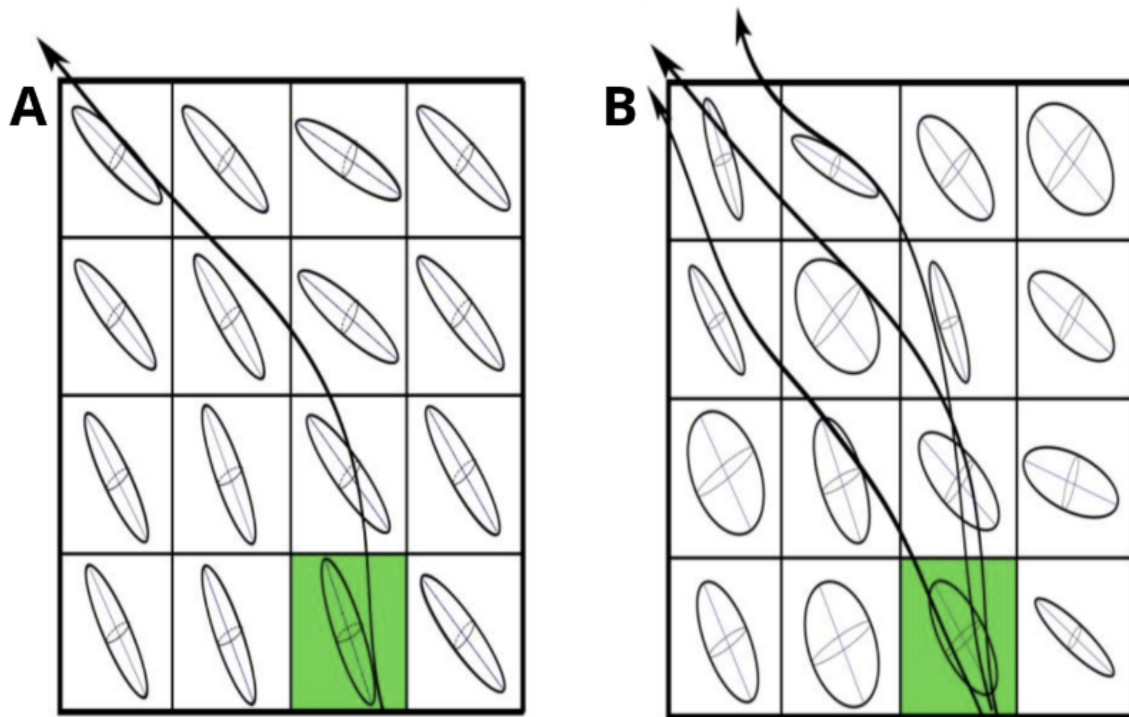


Fig. 2.8: Representación visual de las diferencias entre determinista (A) y probabilístico; (B) aproximaciones a la tractografía. Los enfoques probabilísticos son considerablemente más extensos y exigentes desde el punto de vista computacional, ya que rastrean todas las orientaciones en todos los vóxeles adyacentes al punto inicial (verde). (Imagen disponible en [29]).

2.5 Algoritmos de Clustering de Fibras

Antes de explicar el funcionamiento de los algoritmos de clustering de fibras, analizaremos los algoritmos de agrupamiento como tal.

2.5.1 Algoritmos de Agrupamiento o Clustering

El clustering es una técnica de aprendizaje automático no supervisado que tiene como objetivo agrupar un conjunto de datos sin etiquetar en grupos homogéneos llamados clusters. Cada cluster agrupa datos que son más similares entre sí en comparación con los datos de otros clusters. Este método se utiliza para identificar patrones o estructuras subyacentes en los datos sin conocer de antemano las categorías a las que pertenecen. Los algoritmos de clustering de fibras tienen por función agrupar conjuntos de fibras obtenidas por tractografía en función de la semejanza que existe entre sus trayectorias, lo que permite identificar regiones de materia blanca que son más prominentes de manera no supervisada. Intentamos que estas regiones tengan el mayor parecido posible con los fascículos de materia blanca conocidos disponibles en los distintos atlas anatómicos, con el fin de intentar determinar automáticamente estos mismos y qué tanto se diferencian entre sí para una población de sujetos.

El **algoritmo K-Means** es uno de los métodos de clustering más utilizados y se basa en la partición de los datos en k clusters. Este algoritmo se basa en la minimización de la inercia, que corresponde a la suma de las distancias cuadradas entre los puntos y el centroide del cluster al que pertenecen, lo que resulta en la asignación de grupo a aquellos puntos que estén más cercanos entre sí. Matemáticamente, el objetivo del algoritmo K-Means se expresa como:

$$\arg \min_C \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (2.5.1)$$

donde C_i es el conjunto de puntos que pertenecen al cluster i , μ_i es el centroide de C_i , y $\|\cdot\|$ denota la norma euclidiana.

El **MiniBatch K-Means** es una variante del K-Means que utiliza pequeños lotes de datos aleatorios para actualizar los centroides, lo que reduce significativamente el tiempo de cálculo y permite que el algoritmo sea escalable a conjuntos de datos más grandes.

El **Método del Codo** es una técnica gráfica utilizada para determinar el número óptimo de clusters

en un conjunto de datos. Consiste en ejecutar el algoritmo de clustering (por ejemplo, K-Means) para diferentes valores de k y calcular la inercia (suma de las distancias cuadradas dentro de los clusters) para cada valor. La inercia se gráfica en el eje y y el número de clusters k en el eje x . Intuitivamente, la inercia mide cuán dispersos están los puntos dentro de cada cluster. Al graficar la inercia en función del número de clusters, el punto donde la tasa de disminución de la inercia se reduce drásticamente se considera el "codo", indicando el número óptimo de clusters. El punto óptimo, o "codo", se encuentra donde la gráfica de la inercia cambia de una pendiente pronunciada a una más suave, indicando una mejora marginal al aumentar el número de clusters.

2.5.2 FFClust

FFClust es un algoritmo de agrupamiento de fibras diseñado para manejar grandes conjuntos de datos de tractografía que pueden contener millones de fibras. Su objetivo principal es identificar clústeres compactos y homogéneos que representen toda la estructura de la materia blanca del cerebro. FFClust es especialmente útil para el estudio de la conectividad local para el análisis a nivel poblacional de fibras, haciendo posible la creación de nuevos atlas de fibras de materia blanca de manera no supervisada. A continuación, revisaremos su funcionamiento. Podemos ver la representación gráfica de sus pasos en la figura (2.9).

1. **STEP 1:** Con la idea de reducir la dimensionalidad del problema, se construyen clusters preliminares de los puntos de los extremos, puntos centrales y puntos intermedios de las fibras mediante el algoritmo Minibatch K-Means [30]. La cantidad de clusters deseados para los puntos iniciales y finales se denota K_{po} , y la cantidad de clusters deseados para los puntos intermedios y centrales K_{pc} . El ejecutable de FFClust recibe tanto la posición de estos puntos (como un arreglo de posiciones de los 21 puntos) como el número de clusters a formar con dichos puntos. Denotaremos al vector de posición `points` y al vector de cantidad de clusters `ks` (ver figura 2.9.a).

Para encontrar la cantidad óptima de clusters es posible aplicar el método de codo [31]. El número óptimo de clusters será la cantidad que disminuya la bajada brusca de la inercia, que representa el error cuadrático medio entre los clusters. Si escogemos una cantidad de clusters en un punto donde la inercia ya no baja de manera significativa, veremos que los clusters serán menos compactos y con menos fibras (ver figura 2.9.b).

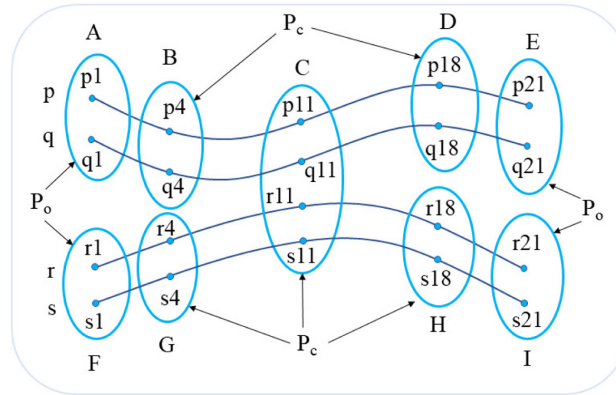
2. **STEP 2:** Este paso construye clusters preliminares de fibras agrupándolas en función de las etiquetas de pertenencia a los clusters de puntos obtenidos en el paso anterior. Un cluster preliminar

de fibras contendrá todas las fibras cuyos puntos comparten las mismas etiquetas de cluster de puntos. El método utiliza una estructura de datos de diccionario en la cual la clave es el conjunto de etiquetas de pertenencia a un cluster y el valor corresponde a las IDs de las fibras que comparten el mismo conjunto de etiquetas de pertenencia a un cluster de puntos. De esta manera, se pueden encontrar las fibras que contienen los puntos que pertenecen a los mismos clusters preliminares de puntos.

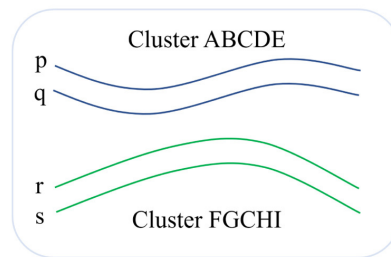
3. **STEP 3:** Este paso reasigna clusters preliminares de fibras pequeños que podrían haberse separado de los clusters grandes en los pasos anteriores. Un cluster pequeño se reasigna al cluster grande más cercano si la distancia entre sus centroides es menor a la distancia máxima d_{Rmax} , la cual es un hiperparámetro. Además, con este procesamiento, se identifican y descartan pequeños clusters ruidosos, lo cual implica que tendremos una pérdida de fibras con este paso (ver figura 2.9.c).
4. **STEP 4:** Este es un paso de refinamiento global que tiene como objetivo fusionar clusters candidatos que aún podrían estar cerca basándose en un hiperparámetro de distancia máxima d_{Mmax} . Este paso primero crea grupos de clusters candidatos, donde cada grupo consiste en clusters que comparten la misma etiqueta de pertenencia del punto central obtenida durante el Paso 1. Luego, los centroides de los clusters candidatos en cada grupo se fusionan basándose en el parámetro de distancia máxima d_{ME} , correspondiente a la siguiente ecuación (ver figura 2.9.d).

$$\begin{aligned}
 d_E(a_i, b_i) &= \|a_i - b_i\| = \sqrt{(a_i^x - b_i^x)^2 + (a_i^y - b_i^y)^2 + (a_i^z - b_i^z)^2} \\
 d_E(a, b) &= \max_{i \in [1, 21]} (d_E(a_i, b_i)) \\
 d_{EF} &= d_E(a, b^F) = d_E(a^F, b) \\
 d_{ME} &= \min(d_E(a, b), d_{EF}(a, b))
 \end{aligned} \tag{2.5.2}$$

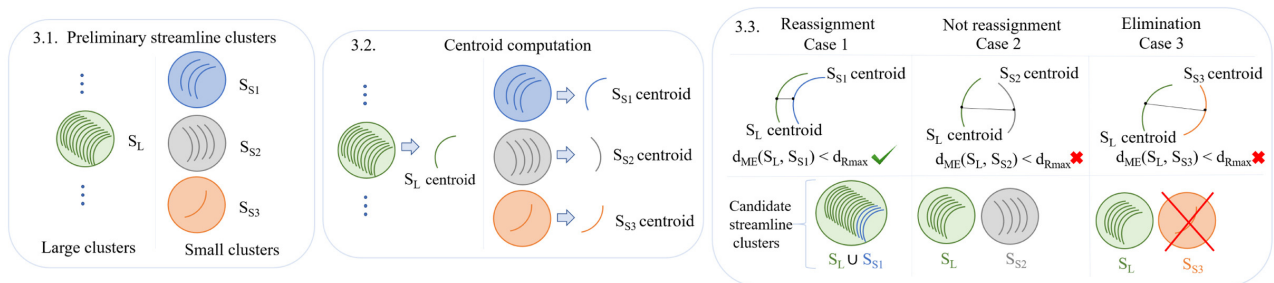
donde d_E es la distancia eucladiana, d_{EF} es la distancia eucladiana en el sentido opuesto, y d_{ME} es la distancia máxima eucladiana entre las dos direcciones de una fibra. En este paso, si dos centroides tienen una $d_{ME} < d_{Mmax}$, entonces fusionamos los clusters.



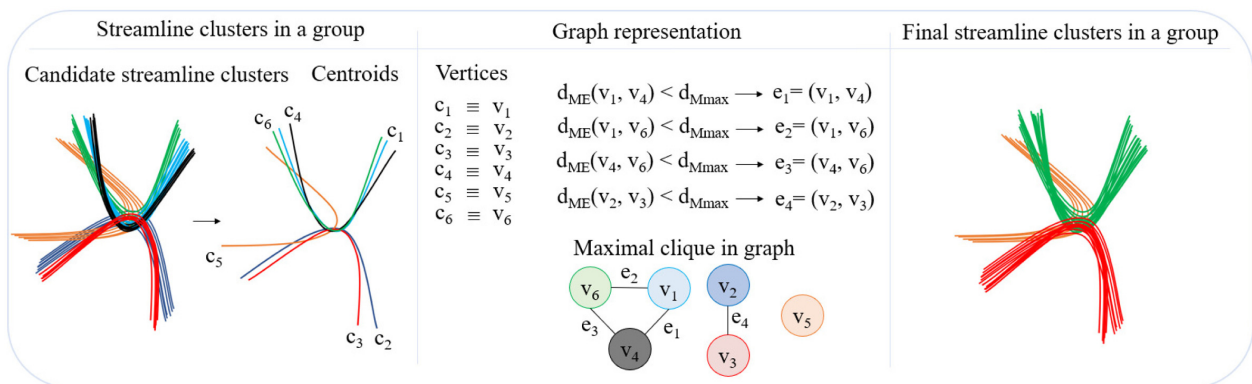
(a) STEP 1: Building point clusters.



(b) STEP 2: Generating preliminary streamline clusters.



(c) STEP 3: Reassigning small preliminary streamline clusters.



(d) STEP 4: Merging candidate streamline clusters.

Fig. 2.9: Descripción del algoritmo FFClust. (Imagen disponible en [9]).

2.5.3 QuickBundles

El algoritmo QuickBundles (QB) [10] se utiliza para agrupar fibras de tractografía cerebral en clusters de manera eficiente. El proceso comienza creando el primer cluster con la primera fibra s_1 . Luego, para cada fibra subsiguiente s_i (donde i varía de 2 a N), el algoritmo calcula la distancia mínima de la fibra actual a los centroides de todos los clusters existentes, considerando también la posibilidad de invertir la fibra para minimizar esta distancia. Para cada fibra, se mantiene un buffer de distancias y un buffer de verificación de inversión. Si la distancia de la fibra invertida es menor que la distancia directa, la fibra se marca como invertida. Una vez calculadas todas las distancias, el algoritmo identifica el cluster con la distancia mínima m . Si m es menor que un umbral predefinido θ , la fibra se agrega al cluster más cercano y se actualiza el centroide del cluster sumando la fibra (o su versión invertida). Si m es mayor o igual a θ , se crea un nuevo cluster con la fibra actual. Este procedimiento se repite para cada fibra hasta que todas las fibras se hayan procesado, resultando en una lista de clusters C . El algoritmo es eficiente debido a que cada fibra se procesa solo una vez y las sumas de las fibras en cada cluster se actualizan de forma acumulativa, sin necesidad de recalcular los clusters repetidamente.

La distancia MDF (minimum average direct-flip) se utiliza para medir la similitud entre dos fibras, considerando ambas orientaciones posibles de las fibras (normal y volteada). Esta distancia se define como el mínimo entre la distancia directa y la distancia invertida:

$$\begin{aligned}
 d_{\text{direct}}(s, t) &= \frac{1}{K} \sum_{i=1}^K \|s_i - t_i\|, \\
 d_{\text{flipped}}(s, t) &= d_{\text{direct}}(s^F, t), \\
 \text{MDF}(s, t) &= \min(d_{\text{direct}}(s, t), d_{\text{flipped}}(s, t)),
 \end{aligned}
 \tag{2.5.3}$$

donde s y t son las dos fibras a comparar, y s^F es la versión invertida de la fibra s . Aquí, d_{direct} es la distancia promedio entre puntos correspondientes de las dos fibras, mientras que d_{flipped} es la distancia promedio considerando una de las fibras invertida. MDF toma el mínimo de estas dos distancias para asegurar que se considera la orientación que minimiza la distancia entre las fibras. Se calcula encontrando la distancia Euclidiana entre puntos correspondientes en dos fibras y tomando el máximo de estas distancias.

Cabe mencionar que existe QuickBundles X [32], pero no presenta mejor desempeño que el original en términos de calidad de clusters, aunque es más eficiente en términos de tiempo y memoria.

2.6 Parcelación Cerebral

La parcelación cerebral consiste en dividir el cerebro en regiones distintas, basadas en ciertos criterios. Estos criterios pueden incluir características funcionales, estructurales o citoarquitectónicas, con el objetivo de identificar regiones con comportamiento homogéneo o interconectado. Dependiendo de los datos y métodos empleados, los algoritmos de parcelación pueden clasificarse en diferentes tipos, entre los que destacan la parcelación basada en conectividad funcional y la parcelación basada en conectividad estructural.

2.6.1 Criterios de Parcelación basado en MRI

La **parcelación basada en conectividad funcional** utiliza datos obtenidos de técnicas como la *resonancia magnética funcional* (fMRI), la *magnetoencefalografía* (MEG) o la *electroencefalografía* (EEG), las cuales miden la activación de las distintas zonas del cerebro en el tiempo. A partir de ello, es posible identificar regiones cerebrales con actividad similar o patrones de comportamiento repetitivos, lo que nos puede llevar a una parcelación utilizando parámetros como la correlación funcional para agrupar regiones que están relacionadas entre sí. Por ejemplo, en [33], se describe una parcelación basada en datos de MEG, y en [34], se emplean datos de fMRI para calcular la conectividad funcional de distintas zonas y utilizar algoritmos de aprendizaje profundo para clasificar síndromes de dolor crónico.

La **parcelación basada en conectividad estructural** se fundamenta en la idea de que la organización funcional del cerebro está determinada por su anatomía conectiva. Es posible segmentar regiones corticales funcionalmente distintas identificando patrones de conectividad neural, mediante el uso de tractografía. Estos métodos suelen seguir pasos similares: primero, se establece un perfil de conectividad para cada subunidad en la región a parcelar, utilizando datos de tractografía para identificar las conexiones que tienen las distintas zonas de materia gris, a partir de la estimación de las trayectorias de materia blanca. Estos perfiles se comparan y se agrupan en regiones con patrones de conectividad similares, permitiendo parcelar el cerebro según estas.

Ambos enfoques, ya sea basados en conectividad funcional o estructural, tienen como premisa fundamental que las regiones cerebrales que interactúan estrechamente, ya sea a través de patrones funcionales o conexiones anatómicas, pueden agruparse en unidades coherentes que representan áreas funcionales o estructurales del cerebro.

2.6.2 Análisis del conectoma en poblaciones de sujetos basados en tractografía

Al comparar tractogramas entre diferentes individuos, es posible identificar variaciones y patrones comunes en la conectividad cerebral que pueden estar asociados con diversas condiciones neurológicas o psicológicas, como enfermedades neurodegenerativas, trastornos del desarrollo o diferencias en el rendimiento cognitivo.

Es posible realizar análisis de poblaciones con tractografía al migrar los datos de tractografía en un espacio común y la posterior comparación agrupación de fibras dentro de esta población de sujetos. En este ámbito, la tractografía probabilística ofrece mejoras significativas en el análisis de poblaciones en comparación con la tractografía determinista. Mientras que la tractografía determinista sigue un único camino basado en la dirección de mayor difusión en cada voxel, la tractografía probabilística explora múltiples trayectorias posibles. Esto se traduce en una representación más completa de la conectividad cerebral, especialmente en regiones complejas donde las trayectorias en un sector del espacio pueden variar más, permitiendo la identificación de patrones de conectividad más sutiles y la detección de diferencias significativas entre grupos, lo que es crucial para avanzar en la comprensión de la organización cerebral y su variabilidad entre individuos.

2.6.3 La necesidad de parcelaciones con distintos niveles de granularidad

El uso de parcelaciones cerebrales con diferentes niveles de granularidad podría ser clave para capturar la complejidad de la organización cerebral y sus variaciones entre individuos. Cuando hablamos de distintos niveles de granularidad en el contexto de la parcelación cortical, nos referimos a la capacidad de dividir la corteza cerebral en áreas o subparcelas de diferentes tamaños y especificidades. Es decir, dependiendo del nivel de granularidad, las parcelas pueden ser más grandes y abarcar regiones más generales del cerebro (baja granularidad), o ser más pequeñas y específicas, identificando subregiones más detalladas (alta granularidad).

En el proceso de creación de un dendrograma, se agrupan subparcelas en diferentes niveles jerárquicos. En los niveles superiores del dendrograma, las subparcelas se combinan para formar regiones más grandes, representando una visión de bajo detalle del cerebro. En niveles inferiores, las subparcelas permanecen más separadas, manteniendo un alto nivel de detalle y especificidad. Por lo tanto, los "distintos niveles de granularidad" permiten que los investigadores analicen la conectividad cerebral y la organización cortical en múltiples escalas, desde una vista global hasta un análisis más fino y detallado de las subparcelas.

2.6.4 Estado del Arte

En [35], se presenta un enfoque de parcelación que consiste en abordar la relación entre la conectividad estructural (SC) y la conectividad funcional en estado de reposo (rsFC) en el cerebro humano. Para ello, los autores proponen un método que se basa en la organización modular jerárquica de las redes cerebrales. Este enfoque busca identificar la partición común óptima compartida por la estructura y la función, maximizando una medida llamada *cross-modularity*, para evaluar cómo las estructuras modulares de estas se superponen. La hipótesis subyacente es que al asumir que funciones segregadas están asociadas con módulos estructurales distintos, la visualización de los datos rsFC en términos de la SC ayudaría a revelar cómo la estructura influye en la función, y viceversa. Este enfoque permite extraer una partición óptima que ilustra una correlación más estrecha entre estructura y función de lo que se había considerado previamente. Se teoriza en este que la nueva partición divide el cerebro en regiones disjuntas denominadas "módulos estructura-función" (SFMs), representando un esqueleto a nivel grueso del cerebro, que está en gran medida compartido por la estructura y la función. Este estudio es de vital importancia, pues muestra la necesidad del estudio de la jerarquía de la SC, pues esta puede utilizarse para comprender a nivel más detallado cómo se relaciona esta misma con parámetros funcionales, específicamente con la rsFC. Se puede ver el detalle del algoritmo en la figura 2.10.

La investigación centrada en los módulos de conectividad funcional (SC-FC) [35] destaca la presencia de correlaciones significativas entre diferentes módulos cerebrales, revelando una estructura modular compleja. Este enfoque destaca la importancia de comprender las correlaciones, reconociéndolas como inherentes a las redes funcionales en estado de reposo.

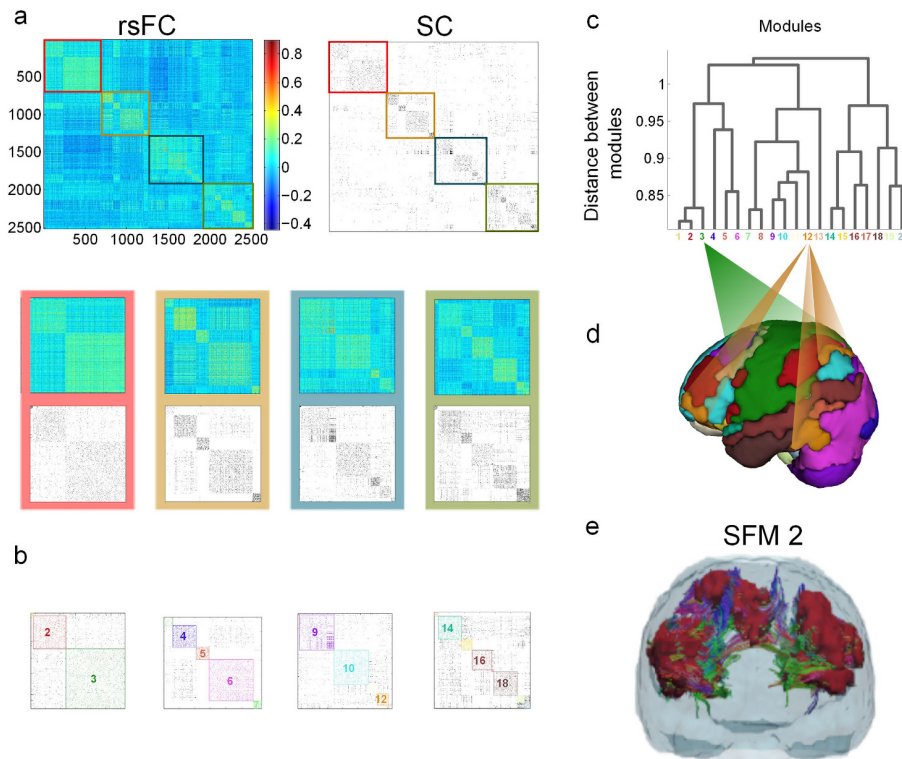


Fig. 2.10: Similaridades fuertes entre matrices SC y rsFC en los SFMs. *A:* La fila superior de imágenes muestra las matrices rsFC (izquierda) y SC (derecha) promediadas sobre 12 pacientes sanos con 2514×2514 ROIs (el 47 % de ellos son ROIs corticales, el 53 % subcorticales, incluyendo el cerebelo). Se identificaron $M = 20$ módulos cerebrales aplicando clustering jerárquico aglomerativo a la matriz rsFC, representando los elementos de la matriz SC (binarizados para mayor claridad de visualización) en el mismo orden que en la rsFC. La amplificación de la diagonal de la rsFC y SC se presenta en rectángulos de colores diferentes (rojo, ocre, azul y verde). La matriz SC muestra similitudes sólidas con la organización modular de la rsFC, a pesar de que la SC es más dispersa que la rsFC. *B:* Los $M = 20$ módulos se colorearon sobre la SC. *C:* El dendrograma aplicado a rsFCs identificando los $M = 20$ módulos. *D:* Una representación tridimensional del cerebro de los 20 módulos. *E:* Como ejemplo, se muestra que la partición cerebral satisfacen que las diferentes áreas cerebrales correlacionadas funcionalmente (coloreadas en granate) que pertenecen al mismo módulo están conectadas estructuralmente.

Por su parte, el algoritmo de Lefranc [36] parte con la segmentación de las circunvoluciones cerebrales usando FreeSurfer. Utilizando la segmentación del giro postcentral, encargado de la actividad motora, calcularon una matriz de conectividad estructural usando la correspondencia de cada fibra con el vértice más cercano entre cada fibra que pasa por el giro postcentral y termina en otra parte del mallado cerebral. Seguido de esto, computa una matriz de conectividad entre los vértices del mallado y el giro postcentral. Esto lo hace para cada sujeto. Luego, aplica una máscara intersujeto, conservando sólo los vértices donde al menos la mitad de los sujetos posea un valor no nulo en su perfil de conectividad. El

perfil de conectividad corresponde a una arista de la matriz de conectividad. Estos se suman y normalizan en función del número de conexiones. Al analizar estos perfiles, Lefranc se percató de que espacialmente las regiones fuertemente conectadas varían en cada sujeto. Para solucionar este problema, aplicó un algoritmo watershed para poder hacer que todas esas regiones sean una sola, concatenando las regiones con cuencas de baja profundidad con las de alta profundidad. Luego, se computó la conectividad entre las zonas segmentadas y los vértices del giro postcentral. Finalmente, se aplicó clustering k-means a la matriz, parcelando en función a esta las zonas del giro postcentral. Este procesamiento se hizo para todo el cerebro.

Por otro lado, Moreno-Dominguez [37] propone un ADP que realiza su parcelación inicial a partir de un criterio de distancia basado en la correlación de coeficientes de conectividad estructural entre 2 regiones, las cuales son las regiones de interés seleccionadas como punto de partida para el análisis de conectividad. Una vez calculadas las distancias, se construye un dendrograma a partir de estas, que representa la organización jerárquica de la estructura de conectividad cerebral. La partición de este mismo obtiene la conectividad entre parcelas a distintos niveles de granularidad. Según esta investigación, el método jerárquico para la parcelación basada en la conectividad de todo el cerebro supera las limitaciones de los métodos existentes al codificar la información de la estructura subyacente en todos los niveles de granularidad en un árbol jerárquico o dendrograma. Esto permite una representación más exhaustiva de la verdadera estructura subyacente del cerebro. Los métodos existentes suelen buscar un número predefinido de áreas y se limitan a pequeñas regiones de materia gris. Sin embargo, es poco probable que una sola parcelación que divida el cerebro en un número finito de áreas sea una representación adecuada de la organización funcional-anatómica del cerebro. El método jerárquico aborda esta limitación al permitir la granularidad anidada múltiple y la posibilidad de una comparación de la estructura completa. Podemos ver el proceso de agrupación en 2.11.

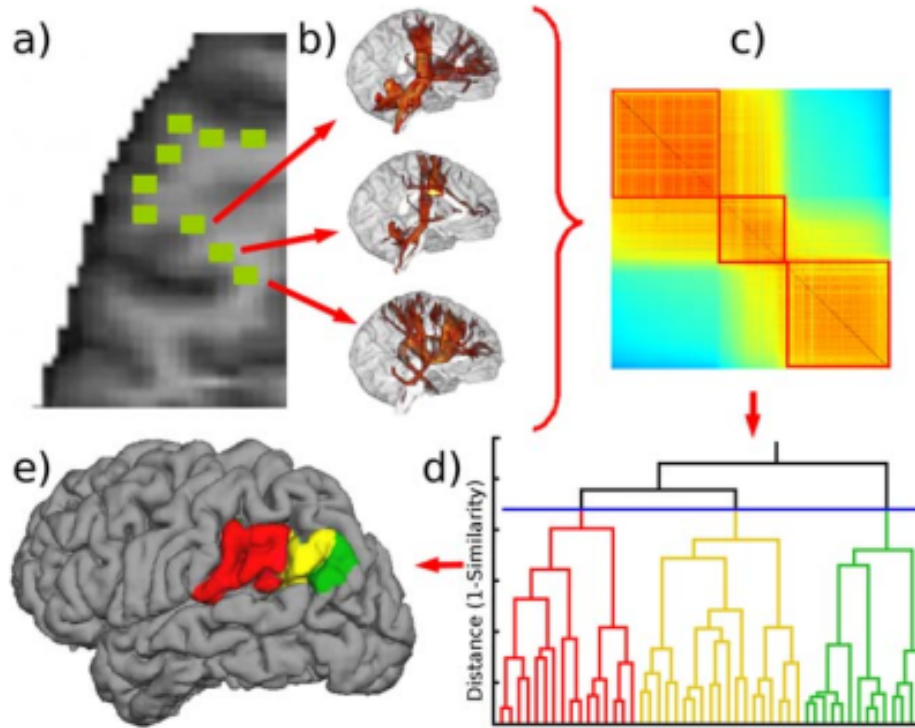


Fig. 2.11: a) Seleccionar vóxeles en la interfaz de materia gris/materia blanca; b) generar tractogramas probabilísticos de vóxeles semilla; c) calcular similitudes entre los tractogramas; d) construir un árbol de conectividad; e) seleccionar particiones dentro del árbol y mapearlas de nuevo en la corteza cerebral.

La metodología propuesta por Joaquín Molina [12] se basa en la integración de información espacial y de conectividad estructural para generar una parcelación cortical a nivel de grupo. Inicialmente, se realiza una sub-parcelación preliminar mediante la intersección de clusters de fibras representativos con la malla cortical de cada sujeto. Dado que múltiples clusters pueden intersectar la misma región cortical, se genera una superposición entre subparcelas, lo que se cuantifica mediante una matriz de solapamiento. Para cada par de subparcelas sp_i y sp_j , el índice de solapamiento se define en la ecuación 2.6.1.

$$I_{\text{overlap}}(sp_i, sp_j) = \frac{|T_{sp_i} \cap T_{sp_j}|}{|T_{sp_i}| + |T_{sp_j}|} \quad (2.6.1)$$

donde $|T_{sp_i}|$ y $|T_{sp_j}|$ representan el número de triángulos en cada sub-parcela, y $|T_{sp_i} \cap T_{sp_j}|$ es la cantidad de triángulos compartidos entre ambas. Adicionalmente, se computa una matriz de conectividad que almacena la fuerza de conexión entre las sub-parcelas basada en la cantidad de fibras que las enlazan. Ambas matrices se combinan ponderadamente mediante coeficientes α (para conectividad) y β (para solapamiento) para construir un grafo de afinidad, sobre el cual se aplica un clustering jerárquico.

El dendrograma resultante permite la generación de parcelaciones a distintas escalas de granularidad, fusionando sub-parcelas según su jerarquía. Finalmente, se aplican operaciones morfológicas de erosión y dilatación para eliminar ruido y homogenizar las regiones, garantizando una parcelación cortical coherente y representativa de la conectividad estructural a nivel de grupo.

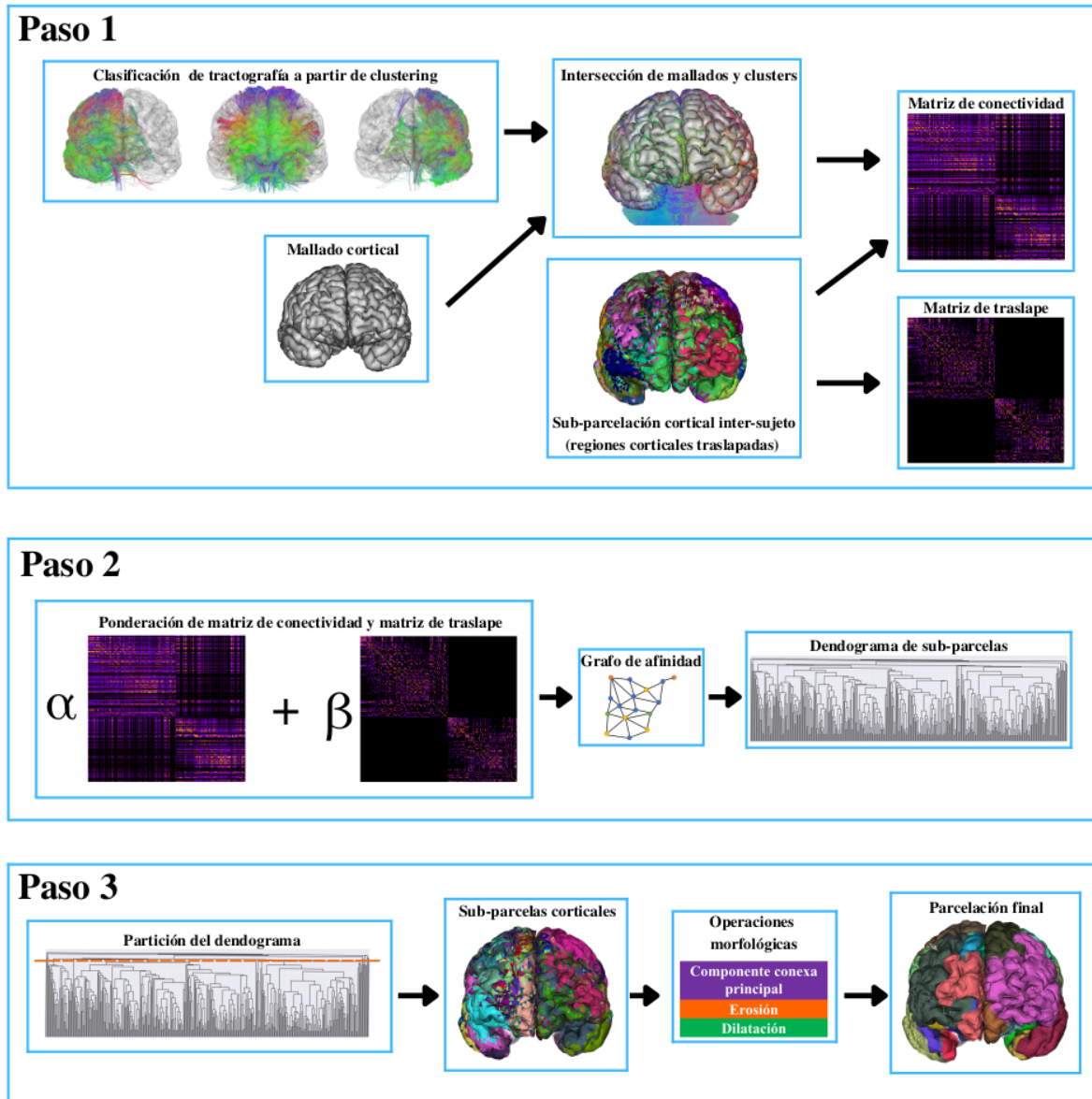


Fig. 2.12: En primer lugar (a), se calcula la intersección de la tractografía con los mallados corticales utilizando clusters de fibras cerebrales. Luego (b), se generan matrices de conectividad y de traslape entre las sub-parcelas resultantes, las cuales son ponderadas adecuadamente. A continuación (c), se obtiene un grafo de afinidad a partir de la matriz ponderada, y posteriormente (d), se aplica el clustering jerárquico para construir un dendrograma. Luego (e), se procede al particionamiento del dendrograma, y finalmente (f), se obtienen las parcelas corticales finales mediante operaciones morfológicas sobre la sub-parcelación.

Capítulo 3. Materiales y Métodos

3.1 Introducción

En esta sección se describen detalladamente los procedimientos y técnicas empleados para llevar a cabo el estudio. La metodología se estructura en diferentes fases. Primero, en la sección materiales se describe la base de datos y sus detalles de obtención. Luego, en la sección métodos se describe todo el proceso del algoritmo detalladamente. Primero, se describen los pasos de obtención de los clusters intersujeto representativos de la población de sujetos. Luego, se explica el algoritmo de intersección de fibras utilizado. Posteriormente, se habla de cómo se obtuvieron subparcelas a partir de los clusters de fibras. Finalmente, se explica la problemática del traslape y del algoritmo de dilatación, proponiendo soluciones a estas mismas.

3.2 Materiales

Inicialmente, se cuenta con una base de datos compuesta de mallados corticales y tractogramas de 100 sujetos del HCP-YA. A continuación, se revisará la metodología de obtención de estas.

3.2.1 Datos de Tractografías

Las imágenes dMRI que dispone HCP-YA se obtuvieron a partir de imágenes dMRI multi-shell adquiridas con 3 b-values, $(1000, 2000, 3000s/mm^2)$, con un tamaño de voxel isotrópico de $1.25mm$. Estas imágenes fueron reconstruidas utilizando el algoritmo SENSE [38], corregidas de distorsiones de susceptibilidad, movimiento del sujeto y corrientes de eddy usando los protocolos disponibles en [39], [40] y [41].

Los tractogramas utilizados en este trabajo fueron calculados en la investigación de Claudio Román [6]. Utilizando el software MRtrix3 [42], se calculó tractografía probabilística a las imágenes dMRI multi-shell de los 100 sujetos mencionados usando un modelo multi-capa [43], basado en un modelo de deconvolución esférica restringida [44] y usando los siguientes parámetros: tamaño de paso = $0.1mm$, umbral angular = 90° , longitud mínima = $30mm$, longitud máxima = $250mm$ y valor de

corte = $0.06mm$. Finalmente, los tractogramas fueron segmentados a 21 puntos equidistantes y fueron trasladadas a espacio MNI. Los tractogramas finales contienen 3 millones de fibras.

3.2.2 Obtención de Mallados Corticales

Inicialmente, se dispone de 100 imágenes estructurales T1w segmentadas con el software FreeSurfer [45], disponibles en la página oficial del HCP-YA [7]. La salida de FreeSurfer incluye una variedad de datos de interés, como superficies corticales de distintos tejidos, volúmenes segmentados, parcelaciones corticales basadas en atlas (como Desikan-Killiany y Destrieux), mapas de espesor cortical y curvas que describen la geometría de la corteza cerebral [45]. Estos datos son fundamentales para estudios de morfometría cortical y conectividad estructural, permitiendo un análisis detallado de la anatomía cerebral.

Para facilitar el manejo de las superficies corticales, estos archivos fueron transformados al formato GIFTI (`.gii`). La conversión se realizó utilizando el *pipeline* de conversión de FreeSurfer a BrainVISA, un conjunto de herramientas que permite la interoperabilidad entre distintos formatos neurocientíficos [46]. Posteriormente, estos archivos fueron transformados a formato OBJ mediante un script en Python desarrollado para este propósito, siguiendo la especificación del formato Wavefront OBJ [47].

El código utilizado para la conversión implementa dos métodos principales: (i) una conversión basada en la herramienta AimsFileConvert de BrainVISA, que permite una transformación directa entre formatos, y (ii) una conversión manual utilizando la biblioteca NiBabel, que extrae los vértices y las caras de la malla cortical y las exporta en el formato OBJ. En este último método, los vértices se escriben en el archivo OBJ precedidos por la etiqueta ‘v’, mientras que las caras se representan mediante la etiqueta ‘f’, asegurando que los índices de los triángulos sigan la convención de Wavefront OBJ. Para procesar múltiples sujetos, se implementó una función que itera sobre los directorios de entrada, generando una carpeta de salida organizada por sujeto.

3.3 Métodos

3.3.1 Obtención de fascículos representativos de la población de sujetos

1. **Preprocesamientos:** En primer lugar, filtramos aquellas fibras menores a $40mm$, debido a que considerar 3 millones de fibras tiene un alto coste computacional y temporal al realizar clustering de fibras de cerebro completo. Luego, aplicamos un suavizado a cada fibra del tractograma, debido

a que estas fueron tomadas con tractografía probabilística, lo que implica que se consideran más trayectorias posibles, y por tanto, sus fibras son más irregulares. El método utilizado consta de un suavizado mediante un promedio ponderado sobre los puntos de la fibra en cada una de sus dimensiones (x , y , z). La fórmula aplicada corresponde a un filtro de media ponderada, en el que se asigna un peso mayor al punto central, y menores a los puntos vecinos, lo que se traduce en la siguiente expresión:

$$x_{ff}(i) = \frac{(1-p)}{2}x(i-1) + px(i) + \frac{(1-p)}{2}x(i+1) \quad (3.3.1)$$

donde para cada punto de la fibra, la nueva posición $x_{ff}(i)$ es una combinación lineal del punto original $x(i)$ y sus vecinos $x(i-1)$ y $x(i+1)$, donde los vecinos tienen un peso de $\frac{(1-p)}{2}$, distribuyendo así la contribución del suavizado de manera simétrica. La misma operación se aplica de manera independiente a las coordenadas y y z , asegurando que la fibra sea suavizada en todo el espacio tridimensional. El resultado de este proceso es una representación más estable y continua de las fibras, lo que facilita su posterior agrupamiento en el proceso de *clustering*, evitando la presencia de artefactos producidos por trayectorias irregulares en la tractografía probabilística. Podemos observar un ejemplo de este suavizado en la Figura 3.15, donde se aprecia la reducción de irregularidades en comparación con las fibras originales.

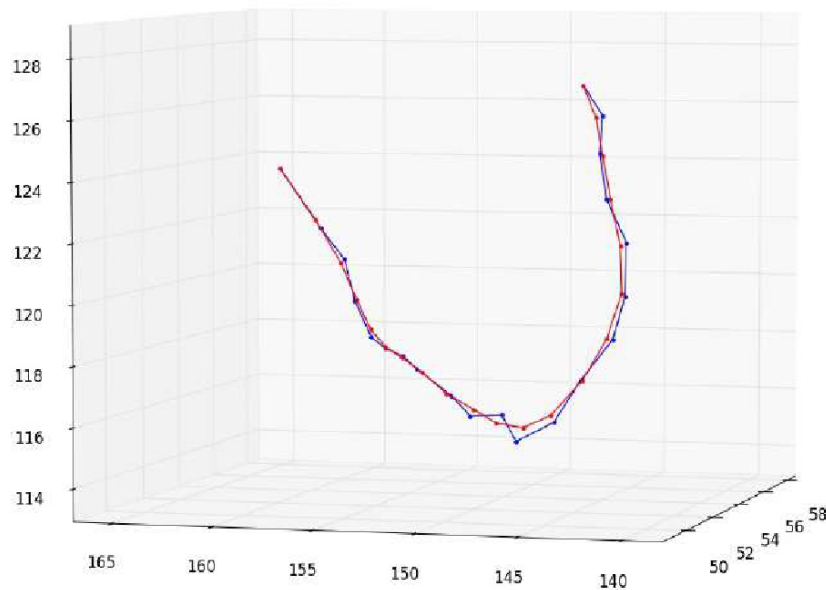


Fig. 3.13: Podemos apreciar en azul una fibra sin suavizar y en rojo una fibra suavizada. Imagen extraída de [6]

2. **Cálculo de clusters intra-sujeto:** En esta etapa, se realizó clustering de fibras FFClust individualmente a los tractogramas suavizados y filtrados de cada sujeto [9]. Para la selección del número

óptimo de clusters de puntos iniciales, se utilizó el método del codo. Para entender cómo funciona este algoritmo, el método de codo y sus parámetros, revisar la sección 2.5.2. Finalmente, se eliminan los clusters con menos de 10 fibras. Para comprobar la calidad de los clusters, se visualizan los 50 clusters más grandes para ver si se están encontrando patrones coherentes. El algoritmo FF-Clust tiene como salida los archivos `.bundles` para los clusters y centroides de los clusters, y un archivo `.hie` que almacena los colores aleatorios que se asignan a cada cluster.

3. **Migración de los centroides a un espacio común:** Una vez realizada la etapa anterior, se migraron los centroides de los clusters de todos los sujetos a un espacio único y se guardaron en un archivo `.npy`. También se almacenó una lista que indexa cada centroide con su respectivo sujeto para hacer posible la siguiente etapa. Esta migración es fundamental, ya que permite que todos los centroides estén alineados en un mismo sistema de referencia, facilitando la comparación y el posterior clustering intersujeto. De esta manera, se garantiza que la similitud entre los centroides se evalúe correctamente en la siguiente etapa, donde se aplicará el algoritmo QuickBundles para identificar clusters representativos a nivel grupal.
4. **Cálculo de clusters representativos inter-sujeto:** En esta etapa, se aplicó el algoritmo QuickBundles sobre el conjunto de centroides previamente migrados al espacio común, con el objetivo de agrupar aquellos que presenten trayectorias similares entre los distintos sujetos.

Para asegurar que los clusters obtenidos sean representativos a nivel grupal, se estableció un umbral de representatividad, denominado *rep_thresh*. Un cluster se considera representativo si contiene fibras provenientes de al menos *rep_thresh* por ciento del total de sujetos. Los clusters que no cumplen con este criterio fueron descartados, ya que podrían corresponder a fibras atípicas o con poca consistencia intersujeto.

Una vez seleccionados los clusters representativos, se procedió a identificar, para cada sujeto, las fibras originales que pertenecen a estos clusters. De esta manera, se obtuvo un conjunto de clusters representativos específicos para cada sujeto. Además, se generó un conjunto global que agrupa todos los clusters representativos de los distintos sujetos en un solo archivo. Ambos conjuntos fueron almacenados en formato `.bundles`.

En la figura 3.14, podemos obtener un diagrama con los procesamientos realizados resumidos.

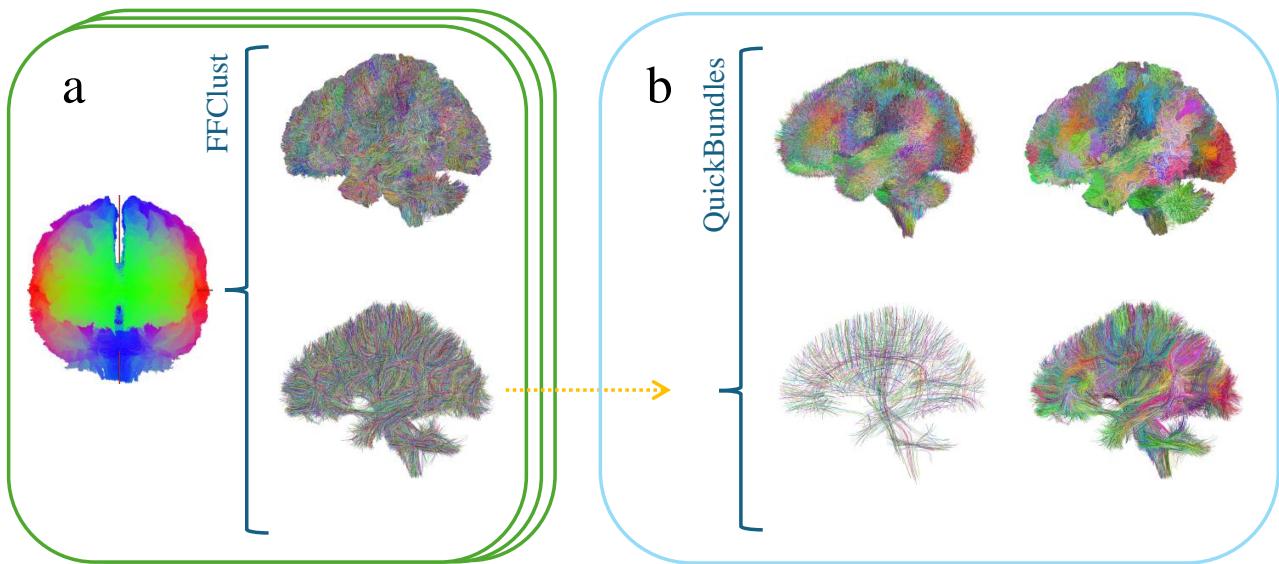


Fig. 3.14: Se ilustra el proceso de agrupación de fibras en dos etapas principales: el clustering intra-sujeto utilizando el algoritmo FFClust (panel a) y el clustering inter-sujeto con QuickBundles (panel b). En la primera etapa, cada sujeto es procesado de manera independiente, aplicando FFClust sobre su tractograma para obtener agrupaciones locales de fibras con trayectorias similares. A continuación, los centroides de estos clusters individuales son migrados a un espacio común, representado por la línea discontinua amarilla, lo que permite alinear la información de todos los sujetos en un mismo marco de referencia. Posteriormente, en la segunda etapa, se emplea el algoritmo QuickBundles para identificar clusters inter-sujeto, es decir, agrupaciones de fibras que presentan trayectorias consistentes a nivel poblacional. Solo aquellos clusters con una representación suficiente en la población (según un umbral de representatividad) son considerados en la etapa final, donde se extraen las fibras originales de cada sujeto que pertenecen a estos grupos. Así, se obtiene un conjunto de clusters inter-sujeto representativos, los cuales reflejan la organización común de las fibras cerebrales en la población analizada.

3.3.2 Aplicación de Algoritmo de Intersección entre Mallados y Clusters de Fibras

A continuación, se explica el algoritmo de intersección utilizado, creado por Felipe Silva [48]. Este fue luego modificado por Joaquín Molina, con el fin de que también pueda considerar la intersección de fibras interhemisféricas. Se tiene acceso a los ejecutables de estos y sus scripts, en C++.

Primero, se procede a realizar el alineamiento de los clusters dentro del espacio común. Un alineamiento de clusters se refiere al proceso de llevar todos los clusters de diferentes sujetos a un mismo espacio de referencia o sistema de coordenadas común. Se utilizó un enfoque basado en la MDF (ver ecuación) entre las fibras. Para cada par de fibras en un cluster, se calculó la MDF. Se seleccionó la orientación que minimizara esta distancia con respecto a la referencia.

Una vez que se tienen los clusters representativos alineados para cada sujeto, se calcula la intersección de estos con el mallado cortical para encontrar los triángulos en los cuales el conjunto de fibras se cruza con la corteza cerebral. Para reducir el espacio de búsqueda de intersecciones, se divide el espacio 3D en celdas cúbicas, almacenando en cada cubo los centroides de los triángulos que contiene. Mientras más pequeño es el tamaño del cubo, menor es el coste de intersección. Por esta razón, cada triángulo se subdividió N veces en triángulos equiláteros, obteniendo así 4^N centroides de triángulos, reduciendo el tamaño del cubo 2^N veces. Se establece el tamaño de la celda en el valor más pequeño posible sin arriesgarse a perder una intersección, que es de aproximadamente $1.5mm$.

Posterior a esto, cada fibra se proyecta un punto hacia atrás y dos puntos adelante en ambos extremos. Las celdas que contienen estos puntos se proyectan hacia sus celdas vecinas, debido a que la intersección puede ocurrir fuera de los límites de la celda principal. A continuación, se seleccionan las celdas no vacías para calcular la distancia de cada triángulo con el punto final de la fibra, se ordenan las distancias de menor a mayor y se asigna el triángulo que contiene el centroide con menor distancia, pudiendo asignar así el triángulo objetivo.

El algoritmo de Möller-Trumbore es un método eficiente para determinar si un rayo (definido por un origen O y una dirección D) intersecta un triángulo en el espacio 3D, cuyos vértices son V_0 , V_1 y V_2 . Este algoritmo calcula la intersección resolviendo la ecuación paramétrica del rayo y el triángulo usando geometría de vectores. Podemos apreciar la ecuación en 3.3.2

$$O + tD = (1 - u - v)V_0 + uV_1 + vV_2 \quad (3.3.2)$$

donde t es el parámetro que indica la distancia de intersección a lo largo del rayo, y u y v son coordenadas baricéntricas que indican si la intersección ocurre dentro del triángulo. Si se encuentran t , u y v con valores positivos y $u + v \leq 1$, se confirma la intersección.

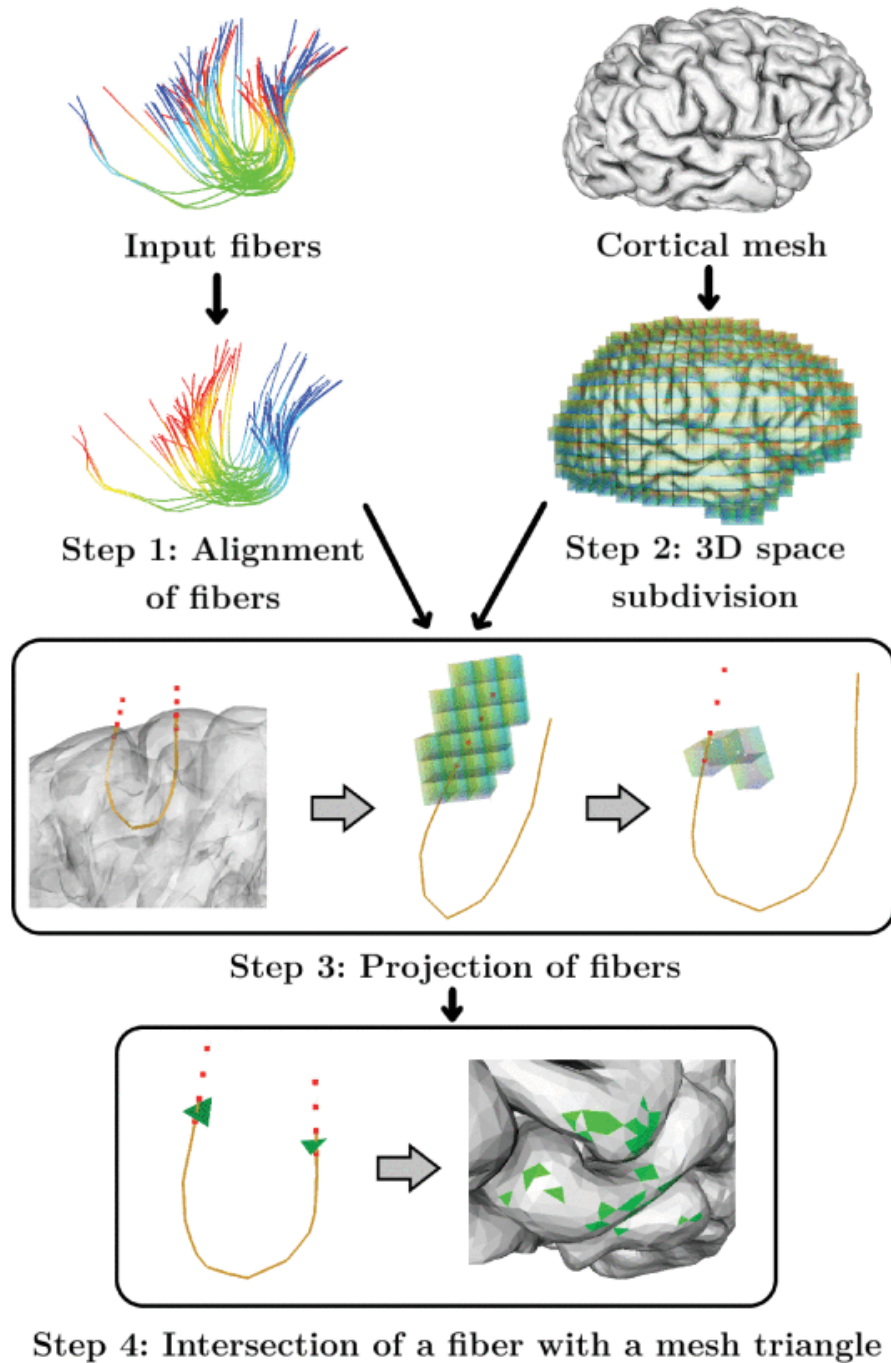


Fig. 3.15: La imagen ilustra los pasos principales del algoritmo de intersección. En el paso 1, las fibras de entrada son alineadas en un espacio común para garantizar la coherencia espacial entre sujetos. En el paso 2, el mallado cortical es subdividido en una estructura de celdas cúbicas que permite reducir la complejidad computacional de la intersección. En el paso 3, se proyectan los extremos de cada fibra hacia atrás y adelante para asegurar que la intersección sea detectada correctamente, incluyendo los posibles puntos de contacto fuera de la celda principal. Finalmente, en el paso 4, se aplica el algoritmo de Möller-Trumbore para determinar la intersección exacta entre las fibras y los triángulos del mallado cortical, permitiendo asignar cada fibra a una región específica de la corteza cerebral. Imágen extraída de [49]

En este trabajo, para aplicar el algoritmo de intersección, se creó una función que lee iterativamente el mallado de cada sujeto (`lh.obj` y `rh.obj`) y sus clusters representativos alineados `cluster_<num_cluster>.bundles` separados por archivos individuales de cada cluster y los pasa como entrada al ejecutable del algoritmo de intersección `interx`. Como salida, se obtiene por cada cluster un archivo de nombre `cluster_<num_cluster>.intersectiondata`. Este archivo de salida corresponde a un diccionario, que por cada cluster almacena las siguientes variables:

- *InTri*: triángulos que cruzan el extremo inicial de una fibra.
- *FnTri*: triángulos que cruzan el extremo final de una fibra.
- *fib_index*: Índice que indica las fibras correspondientes en el diccionario de intersección
- *fiber_class*: Indica la ubicación de las fibras en relación a los hemisferios cerebrales. Las fibras que pertenecen al hemisferio izquierdo (LH) se etiquetan con un valor de 1, mientras que las fibras del hemisferio derecho (RH) se asignan con un valor de 2. Aquellas que cruzan de un hemisferio a otro se categorizan como fibras de izquierda a derecha (LR) con un valor de 3, y las que cruzan de derecha a izquierda (RL) se marcan con un valor de 4.

Finalmente, se clasificaron y escribieron las fibras de acuerdo a su ubicación espacial respecto a los hemisferios cerebrales leyendo los archivos de intersección de cada cluster e iterando sobre las fibras de estos. Para cada cluster, se escribieron archivos `cluster_<num_cluster>_lh.bundles`, `cluster_<num_cluster>_rh.bundles`, para las intrahemisféricas de LH y RH, y `cluster_<num_cluster>_l_r.bundles`, `cluster_<num_cluster>_r_l.bundles` para las inter fibras interhemisféricas de cada clase, respectivamente. Este proceso se realizó para verificar la calidad de la intersección y las subparcelas generadas.

3.3.3 Formación de subparcelas preliminares

En esta etapa, se aplicaron los pasos para solución del traslape propuestos por Narciso López [11]. Se generaron subparcelas preliminares a partir de la intersección de fibras con el mallado cerebral. Cada cluster genera dos subparcelas en el mallado cortical dependiendo de su orientación. Estas se denotan por las etiquetas `cluster<num_cluster>_A` y `cluster<num_cluster>_B` para la porción inicial y final del cluster, respectivamente.

Para cada sujeto, se iteró sobre cada clúster y luego sobre cada fibra del clúster, accediendo a los datos de intersección almacenados en los archivos `.intersectiondata` correspondientes. Este proceso permitió identificar los triángulos iniciales y finales donde las fibras intersectan el mallado, lo que nos permitió definir las subparcelas de acuerdo con estas intersecciones y la localización de las fibras en relación con los hemisferios cerebrales.

Para almacenar la información de parcelación, se crearon dos tipos de diccionarios en formato `.pkl`. El primero, `subparcels_per_triangle`, registra las subparcelas asignadas a cada triángulo, y el segundo, `triangles_per_subparcel`, guarda los triángulos correspondientes a cada subparcela. Estos diccionarios se generan por hemisferio cerebral, resultando en un total de 4 archivos por sujeto para cada paso. Estos llevan una etiqueta, correspondiente al proceso del cual se obtuvieron.

A continuación, se filtraron las subparcelas que contengan menos de una cantidad mínima de triángulos (`min_triangles`), fijado en 82 triángulos como mínimo para ser considerada. Este valor es arbitrario y está sujeto a una futura validación. Los diccionarios de salida se etiquetaron como `filtered`.

Para la visualización de estos archivos, se creó una clase con la biblioteca VTK que muestra la subparcela que más se repite en cada triángulo si es que hay traslape. Esta se utilizó para cada plot a mostrar.

Dado que un triángulo puede ser intersectado por múltiples fibras de diferentes clústeres, se guardaron tantas etiquetas de subparcelas como veces que una fibra del clúster cruzó ese triángulo. Este fenómeno es conocido como *traslape*, y representa una problemática al analizar las subparcelas, ya que cada triángulo debe tener una etiqueta única. Esto crea la necesidad de buscar una solución para asignar subparcelas únicas a cada triángulo.

3.3.4 Alternativa al traslape

A continuación, se aplica una solución al traslape cuyos pasos 1 y 2 se inspiran en el algoritmo de Narciso López [11], y el paso 3 del algoritmo de Joaquín Molina [12].

1. **Cálculo de Mapas de Probabilidad:** Los diccionarios con la etiqueta `filtered` contienen tantas etiquetas como fibras cruzaron dicho triángulo. Se obtuvieron mapas de probabilidad para cada hemisferio, guardados en diccionarios con la etiqueta `probabilities`. Estos mapean la probabilidad de que cada triángulo pertenezca a cierta subparcela, contando las etiquetas de sus vecindades. Por cada triángulo T_i se tiene una vecindad V_i , correspondiente a los triángulos que comparten vértices con T_i . Se contó cuántas veces se repite cada etiqueta en la vecindad, y se

dividió por el número de etiquetas totales presentes en esta. Los mapas de probabilidad almacenan por cada triángulo la subparcela y su probabilidad de que sea de dicha subparcela.

2. **Obtención de Centros de Densidad (DC):** Definimos los DC como los triángulos que tienen para alguna de sus subparcelas asignadas una probabilidad mayor de ser de esa subparcela, de acuerdo a cierto umbral `thresh_DC`. Por consiguiente, los DC serán los sectores de una subparcela que presentan una mayor concentración de fibras del cluster del que nace dicha subparcela. Un DC puede serlo para más de una subparcela. En tal caso, decimos que sus DC están traslapados. Los diccionarios generados llevan la etiqueta DC.
3. **Cálculo de una Matriz de Traslape de Conjuntos de DC entre Subparcelas:** Una vez obtenidos los DC para cada subparcela, se obtienen los triángulos que son DC de dos o más subparcelas, y se calcula su grado de traslape. El grado de traslape entre ellas se calcula mediante la ecuación 3.3.3. Esta métrica normalizada oscila entre 0 (sin traslape) y 1 (traslape completo).

$$idc(SP_i, SP_j) = \frac{dc(SP_i) \cap dc(SP_j)}{\min(dc(SP_i), dc(SP_j))} \quad (3.3.3)$$

Donde SP_i y SP_j son el par de subparcelas a comparar, $idc()$ es el índice de traslape entre DC y $dc()$ representa el conjunto de DC de una subparcela. El resultado de estos cálculos se almacena en una matriz de traslape, donde cada entrada (i, j) representa el valor de traslape entre las subparcelas i y j . Finalmente, esta matriz puede ser visualizada mediante un mapa de calor, lo que permite identificar fácilmente las subparcelas que tienen un alto grado de traslape entre sus DC.

4. **Grafo de traslape para determinación de subparcelas candidatas a ser fusionadas:** Se genera un grafo de traslapes partiendo de la matriz de traslape entre los DC de las subparcelas. La *matriz de traslape* es un mapa de calor, que nos permite verificar que todas las zonas generadas son conectadas entre sí por más de un cluster. La presencia de traslape implica la tarea de utilizar un enfoque para poder asignar estos triángulos con asignación a más de una parcela. Cada nodo en el grafo representa una subparcela, y una arista conecta a otro nodo si el valor de traslape entre esas dos subparcelas (extraído de la matriz de traslape) supera un umbral predefinido `overlap_thresh`. Al construir este grafo, las subparcelas con traslapes significativos quedan conectadas por aristas, formando componentes conexas, que son grupos de subparcelas conectadas entre sí directamente o a través de otras subparcelas. Cada una de estas componentes conexas se considera un grupo de subparcelas candidatas a ser fusionadas. Para extraer estos candidatos, el código utiliza la estructura de grafos de la biblioteca `NetworkX` para identificar las componentes conexas, donde cada componente representa un conjunto de subparcelas que se traslapan suficientemente para ser

fusionadas en una sola entidad, almacenándose en grupos. Luego, para cada grupo, se selecciona la subparcela que tiene más triángulos como la etiqueta final de la subparcela. Los triángulos de todas las subparcelas del grupo se fusionan en esta subparcela principal, eliminando duplicados. Finalmente, se actualiza un mapa que registra qué subparcelas han sido absorbidas, y se generan los diccionarios de salida con la etiqueta `merge` para almacenar los datos de las subparcelas fusionadas, y `subparcel_merge_map_LH.pkl` y `subparcel_merge_map_RH.pkl` para almacenar la composición de las parcelas que se juntaron.

5. **Obtención de Componentes Conexas Principales:** Con los procesamientos anteriores, se logró eliminar parte del traslape, pero no en su totalidad. Además, se observó que las zonas cerebrales identificadas en algunos casos tenían sus triángulos separados y dispersos en ciertos sectores, y en otros, bastante concentrados. Para obtener parcelas definidas, se calculan las componentes conexas principales de las subparcelas cerebrales a partir de los mallados tridimensionales de la superficie cortical. Estas componentes corresponderán al sector más densamente conectado de cierta subparcela.

Para cada sujeto, se cargan los mallados correspondientes a los hemisferios izquierdo y derecho. Posteriormente, se obtiene un mapa que almacena los triángulos conectados a cada arista. A partir de este, podemos identificar los triángulos adyacentes evaluando que las aristas se comparten entre triángulos. Luego, para cada subparcela, se construye un grafo en el que los nodos representan triángulos y las aristas indican que dos triángulos son vecinos. A partir de este grafo, se identifican las componentes conexas de los triángulos, y solo se conserva la componente conexa más grande para cada subparcela, eliminando las componentes pequeñas. Finalmente, se almacenan los resultados para ambos hemisferios en diccionarios con etiqueta `CC`. Finalmente, se itera en los diccionarios y se asigna a cada triángulo sólo la parcela que más se repite en su diccionario, con el fin de que las parcelas finales sean *hard parcels*, es decir, que sólo tengan asignada una parcela por triángulo.

6. **Dilatación Morfológica:** En esta etapa del procesamiento, se implementa una dilatación morfológica iterativa sobre las subparcelas corticales previamente identificadas. La dilatación tiene como objetivo expandir cada subparcela agregando triángulos vecinos.

Primero, se eliminan las parcelas más pequeñas (3 o menos triángulos). Se calcula el mapeo de aristas a triángulos en el mallado de la superficie cortical para ambos hemisferios. Luego, se identifican los bordes de cada subparcela, es decir, los triángulos que están en el límite de la subparcela. A partir de esos bordes, se localizan los triángulos vecinos que aún no han sido asignados a ninguna subparcela. Estos triángulos vecinos se añaden a las subparcelas existentes en cada iteración del

proceso de dilatación. El algoritmo está diseñado para dilatar primero las subparcelas más pequeñas, lo que permite una expansión controlada y evita que subparcelas más grandes invadan áreas que deberían ser asignadas a subparcelas más pequeñas. Si en una iteración un triángulo ya asignado es candidato a agregarse por ser vecino de una subparcela, el algoritmo asigna ese triángulo a la parcela que más se repita en los datos crudos iniciales. Si no está presente en los datos crudos, este triángulo no se asigna. Así, nos aseguramos de que la dilatación no se extienda hacia zonas donde no existe conectividad con el cluster de donde nació dicha subparcela.

A continuación, se puede apreciar un resumen de los procedimientos aplicados como alternativa al traslape.

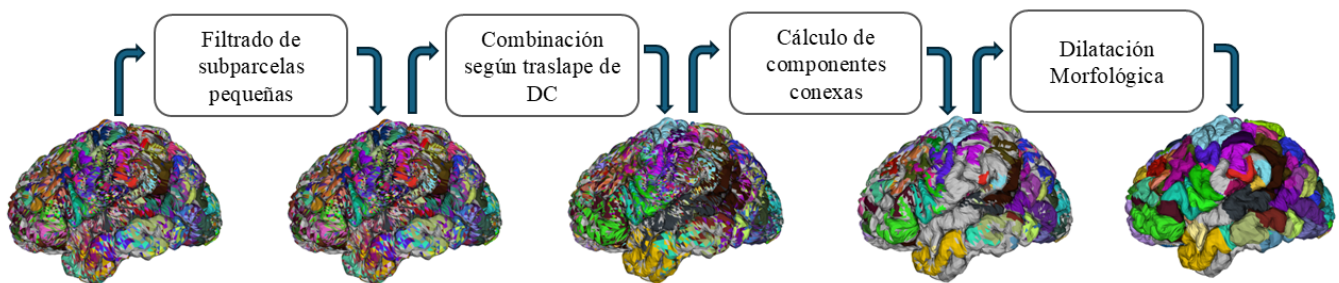


Fig. 3.16: Se muestra un resumen del procesamiento de subparcelas preliminares para obtención de parcelaciones final. Podemos apreciar ejemplos de la salida de cada paso. (Elaboración Propia).

Por último, se hace un análisis del progreso de dilatación para un hemisferio, con el fin de mostrar qué ocurre a medida que van avanzando las iteraciones.

3.3.5 Comparación con Parcelaciones del Estado del Arte

Para validar el método, se calculó el coeficiente de DICE entre la parcelación generada de un sujeto y varias parcelaciones del estado del arte. Estas son las de Brainnetome, Richards, Narciso, y Lefranc ([50], [51], [11] y [36]). Este coeficiente permite cuantificar la similitud entre las regiones de los distintos atlas. Se computaron matrices para almacenar el coeficiente de DICE entre cada par de subparcelas de las dos parcelaciones. El coeficiente de DICE se define en la ecuación 3.3.5.

$$DICE(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (3.3.4)$$

donde A y B son los conjuntos de triángulos que componen cada parcela, $|A \cap B|$ es el tamaño de la intersección entre ambos conjuntos, y $|A|$ y $|B|$ representan el tamaño de los conjuntos A y B respectivamente. Este coeficiente cuantifica el solapamiento entre las parcelas, proporcionando un valor de similitud entre 0 y 1 para cada par de regiones comparadas. Decimos que la similitud es significativa cuando DICE es > 0.5 . Se identifican la cantidad de pares de parcelas con DICE superior a 0.5, y los pares de parcelas que contienen los DICE más altos para cada comparación con el estado del arte. Estos pares de parcelas fueron visualizados para comparar.

Luego de la comparación de parcelas usando el coeficiente de DICE, se calcularon índices de similitud normalizados, basados en la proporción de parcelas con similitud significativa (DICE > 0.5) respecto al número total de parcelas entre la parcelación generada y las parcelaciones del estado del arte. Para cada hemisferio $h \in \{lh, rh\}$, se definió un índice de similitud S_h , dado por la Ecuación 3.3.5:

$$S_h = \frac{n_{\text{simil},h}}{n_{a,h} + n_{b,h}} \quad (3.3.5)$$

donde $n_{\text{simil},h}$ representa el número de pares de parcelas con DICE significativo en el hemisferio h , y $n_{a,h}, n_{b,h}$ son el número total de parcelas en el hemisferio correspondiente para cada parcelación. Además, se definió un índice de similitud global S_{global} , que considera ambos hemisferios simultáneamente, como se muestra en la Ecuación 3.3.6:

$$S_{\text{global}} = \frac{n_{\text{simil},lh} + n_{\text{simil},rh}}{n_{a,lh} + n_{a,rh} + n_{b,lh} + n_{b,rh}} \quad (3.3.6)$$

Estos índices permiten cuantificar de manera sintética la similitud entre parcelaciones, tanto a nivel hemisférico como global, y fueron utilizados para generar las tablas de comparación entre la parcelación propuesta y las parcelaciones de Brainnetome, Lefranc, Narciso y Richards.

Capítulo 4. Resultados

4.1 Introducción

En este capítulo, se presentan los resultados obtenidos a partir de la aplicación del algoritmo. El hardware utilizado está compuesto por una placa madre Intel DX79TO, una tarjeta gráfica NVIDIA GF100GL Quadro 6000, y un procesador Intel Core i7-3820, junto con un entorno operativo Ubuntu 22.04 y Python 3.11.7.

Los resultados se estructuran en varias secciones, comenzando con los preprocesamientos y el clustering intrasujeto, donde se filtran las fibras de menor longitud y se aplican algoritmos específicos para obtener grupos significativos de fibras. A continuación, se detalla el proceso de clustering intersujeto, que permite fusionar la información de varios sujetos en un único espacio representativo, y se finaliza con la creación de una parcelación cortical que integra las características anatómicas de las fibras. Finalmente, se realiza una comparación de los resultados con atlas existentes en la literatura, incluyendo Brainnetome, Richards, Lefranc y Narciso, utilizando el coeficiente de Dice para evaluar la similitud entre las parcelaciones.

4.2 Creación del Atlas Representativo de Fibras

4.2.1 Preprocesamientos y Clustering Intrasujeto

Primero, se eliminaron todas las fibras menores a $40mm$. De los 3 millones de fibras por sujeto que habían inicialmente, quedaron en promedio 834057, en un rango de 694475 a 1062873 fibras por sujeto. Luego, se realizó la primera etapa de clustering haciendo uso del algoritmo FFClust con parámetros $d_{Rmax} = 6mm$, $d_{Mmax} = 6mm$, $K_s = [150, 125, 125, 125, 150]$ y $points = [0, 3, 10, 17, 20]$. Los parámetros d_{Mmax} y d_{Rmax} se establecieron en $6mm$ debido al análisis realizado en [9], donde se encontró que con esta distancia se pueden lograr clusters más compactos. La cantidad de clusters preliminares K_s se estableció utilizando el método de codo. Podemos visualizar su gráfica en la figura 4.17.

En la figura 4.18 podemos apreciar un ejemplo de la salida. Finalmente, se eliminaron los clusters con menos de 10 fibras, quedando un promedio de 13713 fibras, en un rango de 11389 a 16521 clusters

por sujeto. También se contaron las fibras por cluster de cada sujeto. Podemos apreciar su distribución en la figura 4.19.

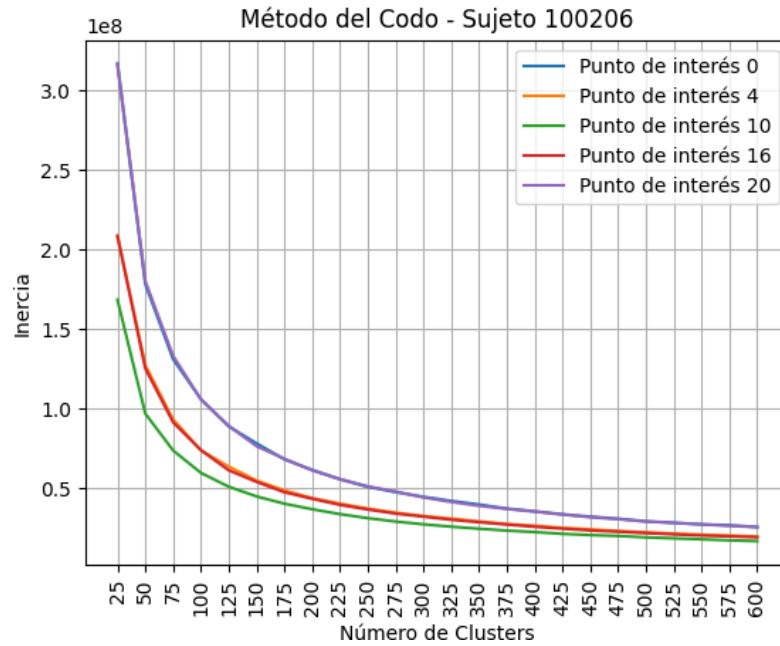


Fig. 4.17: Se puede apreciar el codo para el sujeto 100206. Podemos apreciar que el codo de la línea morada y la azul tienen un codo en los 150 clusters, y las líneas verde, roja y amarilla cerca de los 125 clusters. (Elaboración propia.)

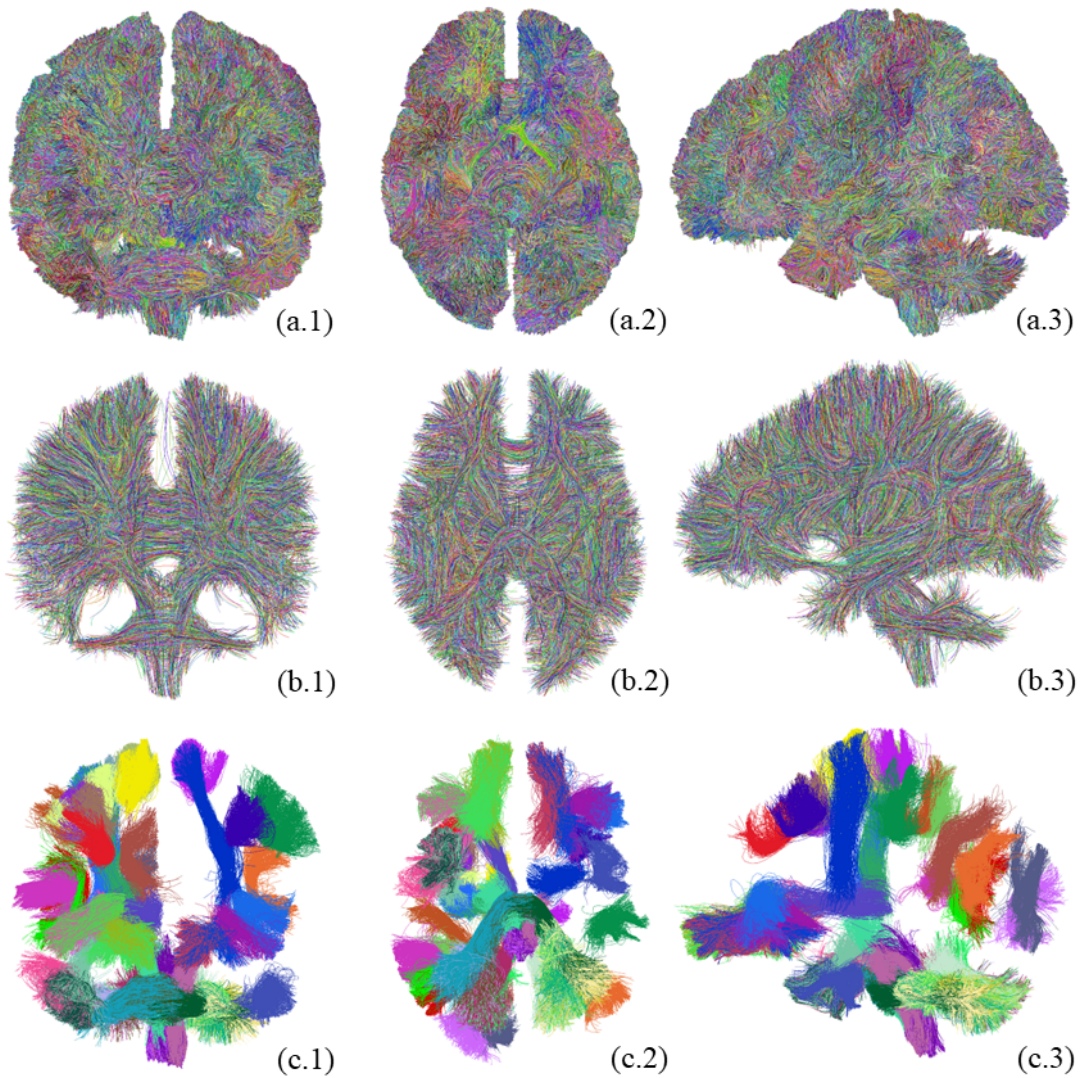


Fig. 4.18: Podemos apreciar todos los clusters intrasujeto en la fila *a*. En la fila *b*, se aprecian sus centroides y en la *c*, se pueden apreciar los 50 clusters más grandes, con el fin de mostrar su compacticidad. Este comportamiento se mostró similar en todos los sujetos. (Elaboración propia).

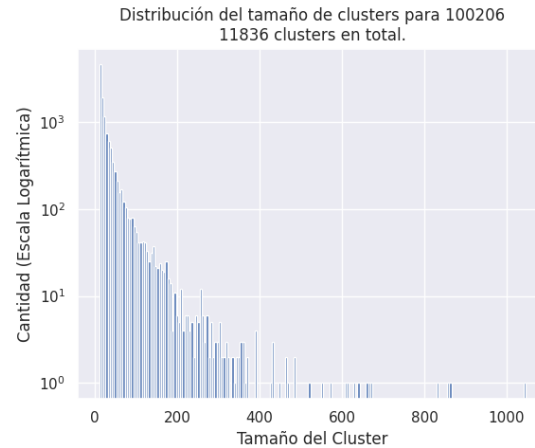


Fig. 4.19: Se aprecia el tamaño de los clusters del sujeto 100206. Podemos ver que existe un gran número de clusters pequeños. (Elaboración propia).

4.2.2 Clustering Intersujeto

Recordemos que en esta etapa, se migraron todos los centroides de los clusters del paso anterior en un espacio común, y se tiene una lista que indexa cada centroide al sujeto respectivo al que pertenece. Se aplicó al archivo con todos los centroides el algoritmo QuickBundles con $\theta = 10mm$ (ver sección 2.5.3), logrando obtener clusters representativos de todos los sujetos en un mismo espacio, y los clusters representativos que corresponden a cada sujeto en su propio espacio.

Fijando un umbral de representatividad $rep_{thresh} = 88$, se obtuvieron clusters intersujeto (de centroides de clusters intersujeto) de 1071868 centroides ordenados en 1059 clusters. Los clusters intersujeto de cada sujeto obtenidos contienen, en promedio, 398655 fibras en un rango de 294844 a 623839 fibras. Estas fibras se agrupan en un promedio de 1025 clusters en un rango de 664 a 1053 clusters. En la figura 4.20, podemos apreciar los atlas descritos y en la figura 4.21 podemos ver gráficos con la distribución del tamaño de los clusters descritos.

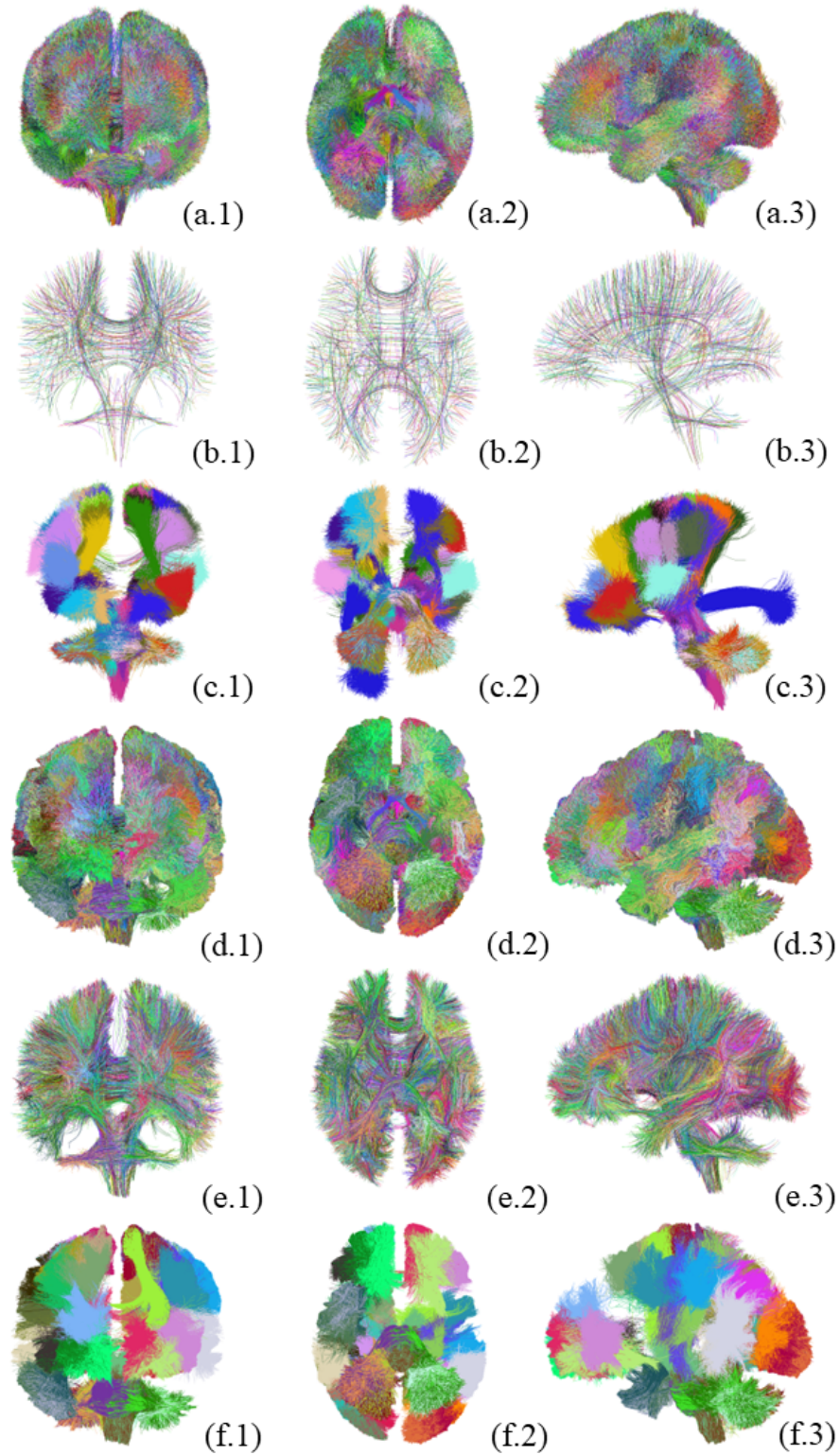


Fig. 4.20: En las filas (a,b,c), se muestran los clusters de centroides representativos, los centroides de los clusters de centroides y los 50 clusters globales más grandes, respectivamente. Así mismo, en las filas (d,e,f) se muestran los clusters representativos del sujeto 100206, los centroides de los clusters representativos y los 50 clusters representativos más grandes de ese sujeto.

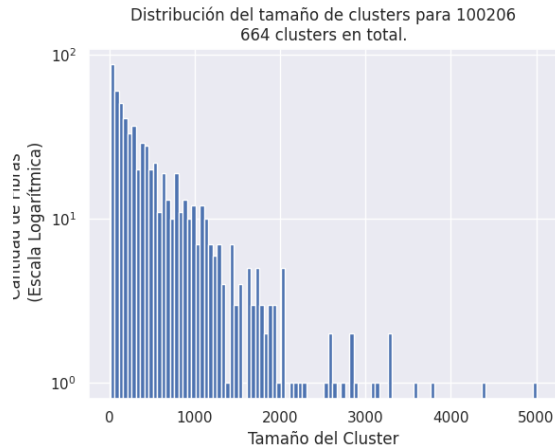


Fig. 4.21: Se aprecia el tamaño de los clusters representativos del sujeto 100206. Podemos ver que existe un gran número de clusters pequeños. (Elaboración propia).

4.2.3 Creación de Parcelación Cortical

En esta etapa, se obtuvieron subparcelas preliminares al calcular la intersección de los clusters representativos de cada sujeto con su mallado cerebral, generando 2 subparcelas por cada cluster. Luego, se filtraron las subparcelas que tienen menos de $min_triangles = 82$ triángulos. Se calculó la probabilidad de que un triángulo pertenezca a una subparcela (como se mencionó en la sección 3.3.4) para determinar los *centros de densidad* de cada subparcela, usando un umbral del $thresh_DC = 20\%$.

En las tablas 4.1 y 4.22 podemos observar la cantidad de subparcelas y la distribución del tamaño de las subparcelas por cada paso. Las matrices de traslape están disponibles en la figura 4.23. Finalmente, se puede apreciar la salida completa de cada etapa en la figura 4.25. Las parcelas finales resultantes fueron de 189 parcelas para el hemisferio izquierdo y 139 parcelas para el hemisferio derecho.

	Hemisferio Izquierdo			Hemisferio Derecho		
	Promedio	Mínimo	Máximo	Promedio	Mínimo	Máximo
Subparcelas Preliminares	1194	806	1275	1123	774	1207
Subparcelas Filtradas	519	372	618	456	314	560
Subparcelas tras Merge	292	208	361	230	156	286

Tabla 4.1: Promedio, mínimo y máximo del número de parcelas en la población de sujetos después de cada paso realizado, por hemisferio cerebral. (Elaboración Propia).

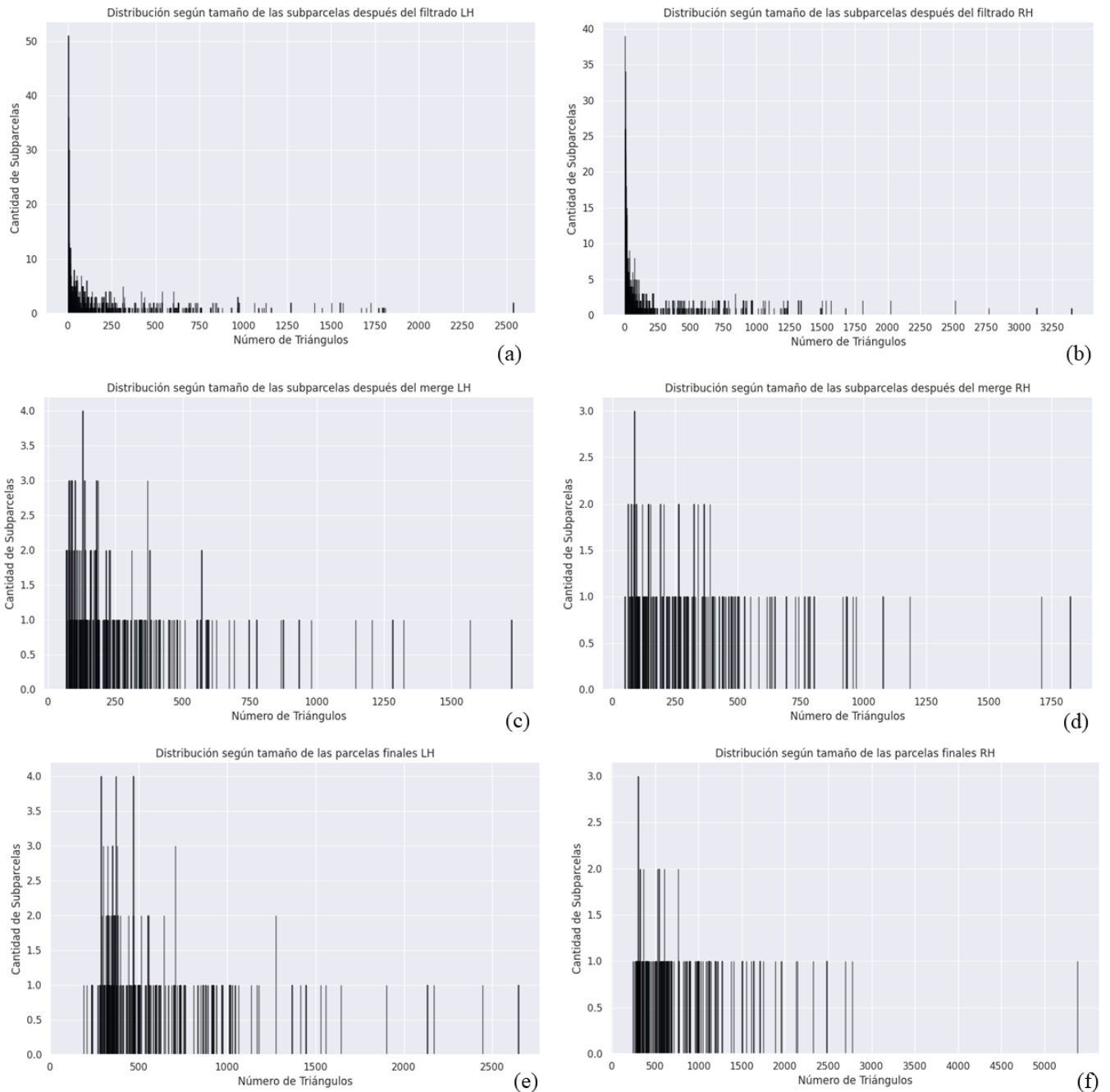


Fig. 4.22: En esta figura, podemos apreciar las distribuciones del tamaño de las subparcelas en cada etapa de procesamiento. Podemos ver que las subparcelas después del filtrado tienen un alto número de parcelas pequeñas, como se ve en (a) y (b). Luego, cuando realizamos el merge o combinación de subparcelas, baja bastante el número de parcelas. Los resultados por hemisferio están en (c) y (d). Por último, en (e) y (f) se aprecia la distribución de parcelas finales. (Elaboración propia).

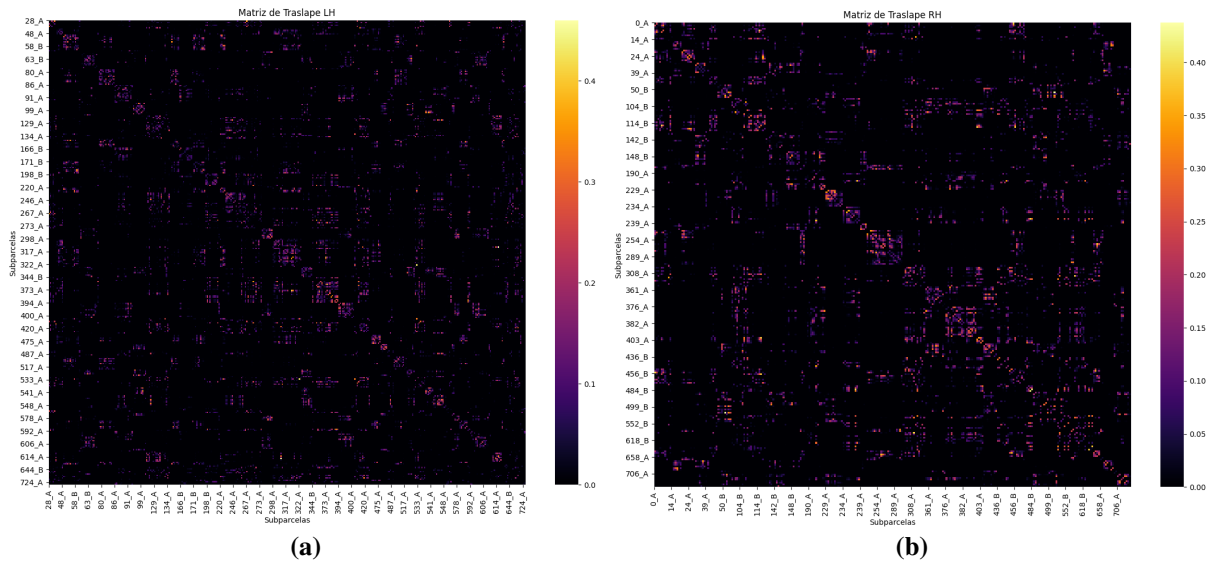


Fig. 4.23: En las imágenes (a) y (b) se puede apreciar las matrices de traslape entre centros de densidad de un sujeto del lóbulo izquierdo y derecho, respectivamente. Para interpretar estas, revisar sección 3.3.4 (Elaboración propia).

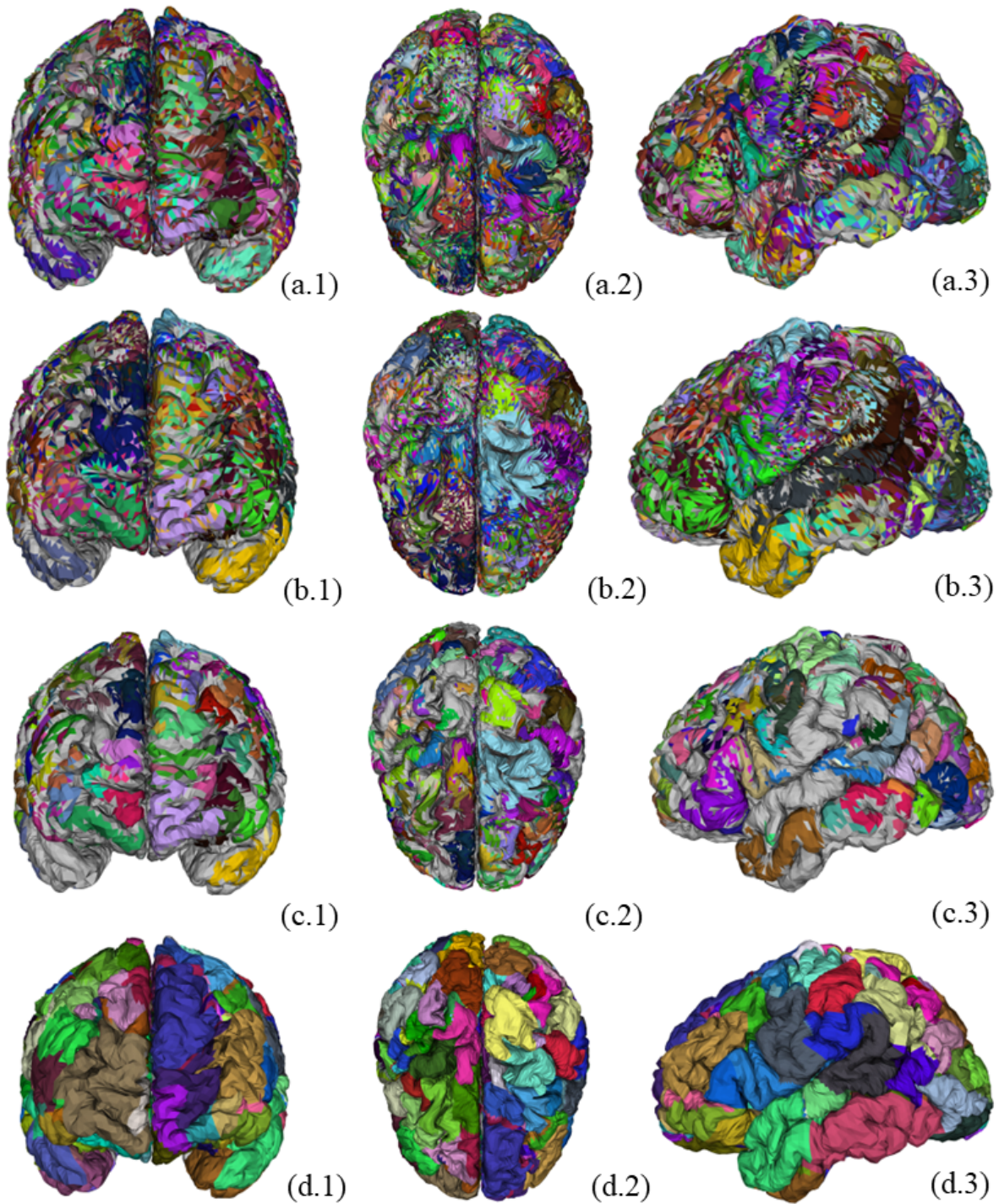


Fig. 4.24: En esta figura, podemos apreciar la estrategia completa que utilizamos para poder solucionar el traslape. Las subparcelas preliminares, subparcelas combinadas, componentes conexas de las subparcelas combinadas y parcelas dilatadas (con 30 iteraciones) se pueden apreciar las columnas a,b,c,d; respectivamente (Elaboración Propia).

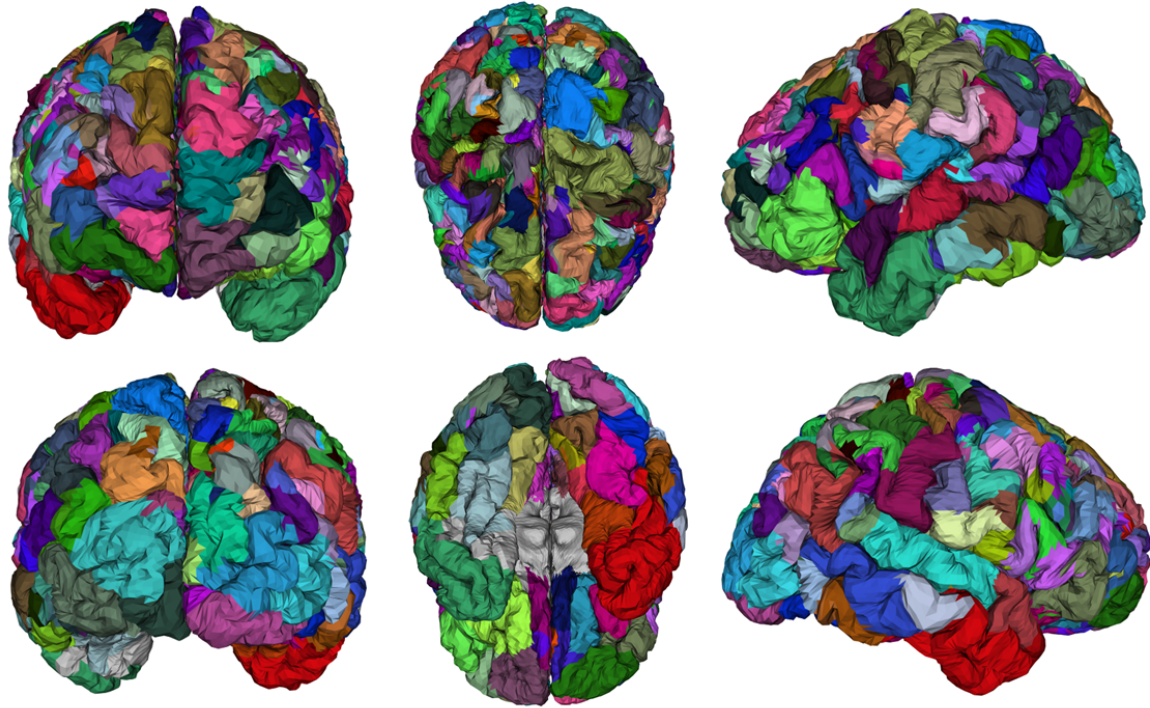


Fig. 4.25: Se muestran las parcelas finales obtenidas tras un total de 20 dilataciones.

En orden de comprobar que el algoritmo de dilatación funciona de manera correcta, se realizaron en total 20 iteraciones. En la figura 4.26 se puede apreciar cómo las parcelas van aumentando su tamaño y ocupando espacios sin asignar a medida que avanza el algoritmo.

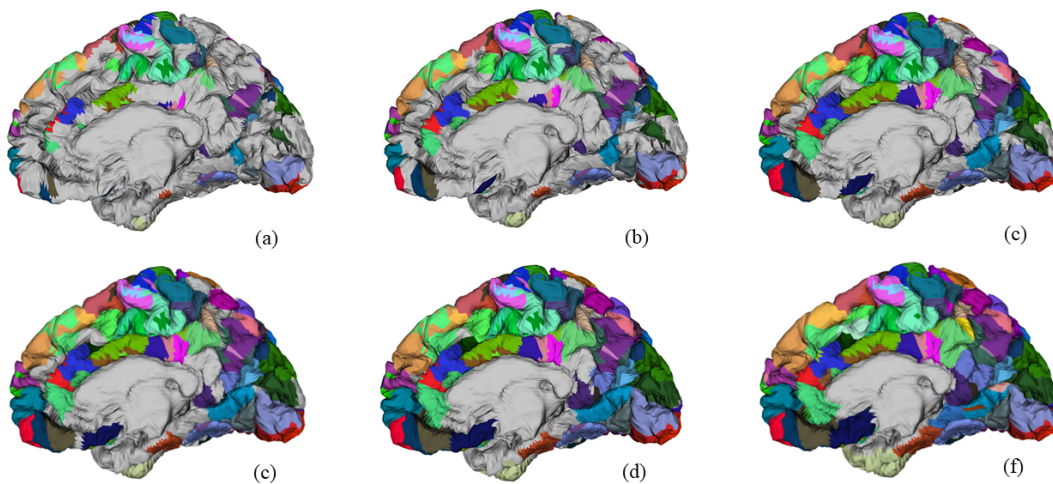


Fig. 4.26: En esta figura, podemos apreciar cómo evoluciona la dilatación a medida que aumentan las iteraciones de esta. Se muestra la iteración 0, 4, 9, 14 y 19 en las figuras a,b,c,d,e,f; respectivamente. (Elaboración propia).

4.2.4 Comparación de Parcelaciones con el Estado del Arte

Se comparó la similitud de las parcelas obtenidas con las parcelaciones del estado del arte. Estas corresponden a las parcelas de Brainnetome, Richards, Lefranc y Narciso ([50], [51], [11] y [36]). Podemos observar la cantidad de parcelas por lóbulo de estas parcelaciones en la tabla 4.2. Por otro lado, la tabla 4.3 nos muestra la cantidad de parcelas que contienen un coeficiente de dice mayor a 0.5. Es interesante apreciar que existe un mayor número de parcelas en el lado izquierdo para todas las parcelaciones del estado del arte, lo cual puede implicar que las funciones del lóbulo izquierdo se especializan más que las del lado derecho, debido a una mayor cantidad de zonas.

Parcelación	LH	RH
Parcelación creada	189	139
Brainnetome	126	125
Lefranc	125	112
Narciso	116	91
Richards	174	128

Tabla 4.2: Cantidad de parcelas iniciales de la parcelación creada y las parcelaciones del estado del arte.

Parcelación	LH	RH
Brainnetome	17	12
Lefranc	53	43
Narciso	39	46
Richards	157	107

Tabla 4.3: Número de parcelas con coeficiente DICE superior a 0.5 en la comparación entre la parcelación dil40 creada y las parcelaciones del estado del arte.

A continuación, se muestran las tablas comparativas basadas en el índice de similitud normalizado entre parcelaciones S_h y S_{global} mencionados en la sección 3.3.5 del capítulo Materiales y Métodos en la tabla 4.4. Se puede apreciar que los resultados de similitud son bastante altos para el algoritmo de Richards pese a que los datos utilizados en su investigación vienen de la base de datos ARCHI. Es importante notar que la metodología que aplicó Richards es bastante similar a la aplicada en esta investigación.

Parcelación	S_{lh}	S_{rh}	S_{global}
Brainnetome	0,054	0,045	0,050086
Lefranc	0,169	0,181	0,169912
Narciso	0,162	0,227	0,191441
Richards	0,525	0,489	0,489796

Tabla 4.4: Se muestra el índice de similitud normalizado que presenta la parcelación creada *dil40* con el resto de las parcelaciones del estado del arte para el lóbulo izquierdo, derecho y la parcelación con ambos lóbulos.

En las tablas 4.5, 4.8, 4.7 y 4.6 podemos ver los 10 pares de parcelas que presentan mayor coeficiente de DICE, entre las parcelas obtenidas y las parcelaciones del estado del arte mencionadas.

Parcelas Obtenidas (LH)	Parcelas Brainnetome (LH)	DICE (LH)	Parcelas Obtenidas (RH)	Parcelas Brainnetome (RH)	DICE (RH)
cluster_72_B	185	0.6203	cluster_618_A	140	0.5833
cluster_592_B	181	0.6132	cluster_184_A	144	0.5789
cluster_517_A	49	0.5795	cluster_378_A	140	0.5699
cluster_421_B	141	0.5663	cluster_289_B	50	0.5557
cluster_263_B	141	0.5657	cluster_142_B	88	0.5524
cluster_318_B	3	0.5484	cluster_379_B	140	0.5480
cluster_724_B	151	0.5472	cluster_191_B	144	0.5259
cluster_262_B	139	0.5285	cluster_563_A	186	0.5211
cluster_167_B	185	0.5246	cluster_233_A	196	0.5166
cluster_245_B	141	0.5237	cluster_79_A	50	0.5133

Tabla 4.5: Coeficientes DICE para Parcelas Obtenidas y Brainnetome en LH y RH

Parcelas Obtenidas (LH)	Parcelas Lefranc (LH)	DICE (LH)	Parcelas Obtenidas (RH)	Parcelas Lefranc (RH)	DICE (RH)
cluster_61_B	141	0.7797	cluster_364_B	663	0.7168
cluster_400_A	301	0.7738	cluster_345_A	652	0.7143
cluster_542_B	303	0.7189	cluster_141_B	451	0.7079
cluster_236_A	291	0.6724	cluster_114_B	653	0.7010
cluster_415_B	323	0.6720	cluster_36_B	451	0.6987
cluster_92_A	301	0.6681	cluster_42_B	653	0.6787
cluster_277_A	141	0.6611	cluster_116_B	653	0.6648
cluster_168_A	262	0.6558	cluster_406_B	663	0.6606
cluster_67_B	141	0.6511	cluster_281_A	641	0.6435
cluster_517_A	151	0.6412	cluster_0_A	641	0.6389

Tabla 4.6: Coeficientes DICE para Parcelas Obtenidas y Lefranc en LH y RH

Parcelas Obtenidas (LH)	Parcelas Narciso (LH)	DICE (LH)	Parcelas Obtenidas (RH)	Parcelas Narciso (RH)	DICE (RH)
cluster_400_A	16622112	0.8411	cluster_42_B	9256904	0.7456
cluster_318_B	6667636	0.8169	cluster_234_A	4647846	0.7338
cluster_262_A	12740310	0.7659	cluster_74_A	66541	0.7222
cluster_170_A	12740310	0.7612	cluster_47_B	66541	0.7188
cluster_261_A	12740310	0.7504	cluster_345_A	15670418	0.6879
cluster_645_B	12740310	0.7496	cluster_616_B	12207623	0.6768
cluster_90_B	16622112	0.7458	cluster_307_A	66541	0.6695
cluster_298_B	6667636	0.7208	cluster_116_B	9256904	0.6675
cluster_92_A	16622112	0.6839	cluster_455_B	66541	0.6621
cluster_603_A	16622112	0.6789	cluster_553_B	12207623	0.6567

Tabla 4.7: Coeficientes DICE para Parcelas Obtenidas y Narciso en LH y RH

Parcelas Obtenidas (LH)	Parcelas Richards (LH)	DICE (LH)	Parcelas Obtenidas (RH)	Parcelas Richards (RH)	DICE (RH)
cluster_400_A	parcel_SP1_IP1	0.7884	cluster_406_B	parcel_IP1_SP1	0.8772
cluster_80_B	parcel_2SF_SF	0.7786	cluster_364_B	parcel_IP1_SP1	0.8552
cluster_262_B	parcel_SP0_SM0	0.7762	cluster_152_B	parcel_IP0_MT0	0.8128
cluster_276_B	parcel_LO_Fu	0.7630	cluster_2_B	parcel_SF0_RMF0	0.7778
cluster_415_B	parcel_SP0_SM0	0.7552	cluster_229_A	parcel_ST_LO	0.7752
cluster_421_B	parcel_SP0_SM0	0.7508	cluster_378_A	parcel_PoC2_PrC2	0.7640
cluster_580_B	parcel_2SF_SF	0.7495	cluster_114_A	parcel_SF0_RMF0	0.7611
cluster_316_A	parcel_RMF1_CMF1	0.7336	cluster_27_B	parcel_SF0_RMF0	0.7595
cluster_59_B	parcel_RMF1_CMF1	0.7327	cluster_563_A	parcel_PrCu_2PoCi	0.7533
cluster_320_B	parcel_RMF1_CMF1	0.7313	cluster_56_A	parcel_IC_SF	0.7491

Tabla 4.8: Coeficientes DICE para Parcelas Obtenidas y Richards en LH y RH

Finalmente, las figuras 4.27, 4.28, 4.29, 4.30, 4.31, 4.32, 4.33, 4.34 presentan una ilustración de los 5 pares de parcelas con mayor coeficiente de DICE. En todas estas imágenes, se aprecia la parcelación creada en el lado izquierdo, y la parcelación a comparar en el lado derecho. Podemos apreciar que existen similitudes significativas en ciertas parcelas en todas las comparaciones realizadas.

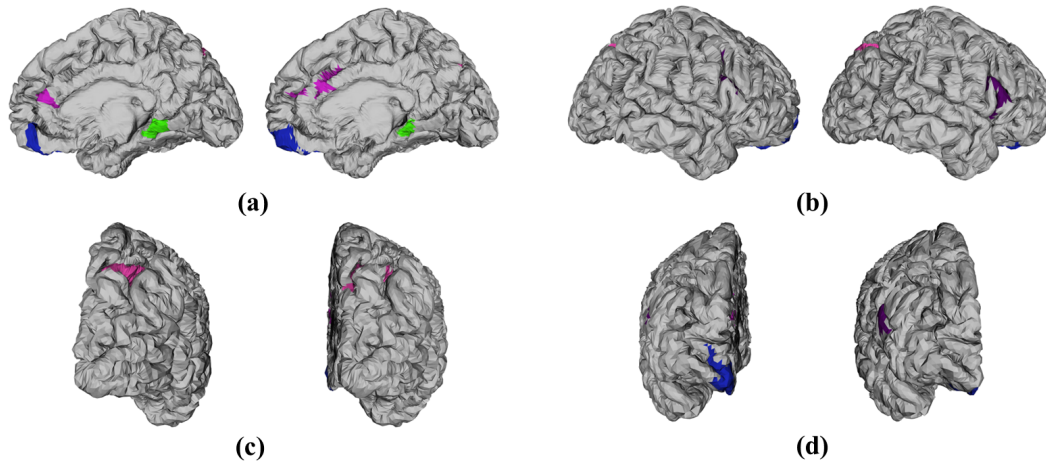


Fig. 4.27: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Brainnetome del lado izquierdo. (Elaboración propia).

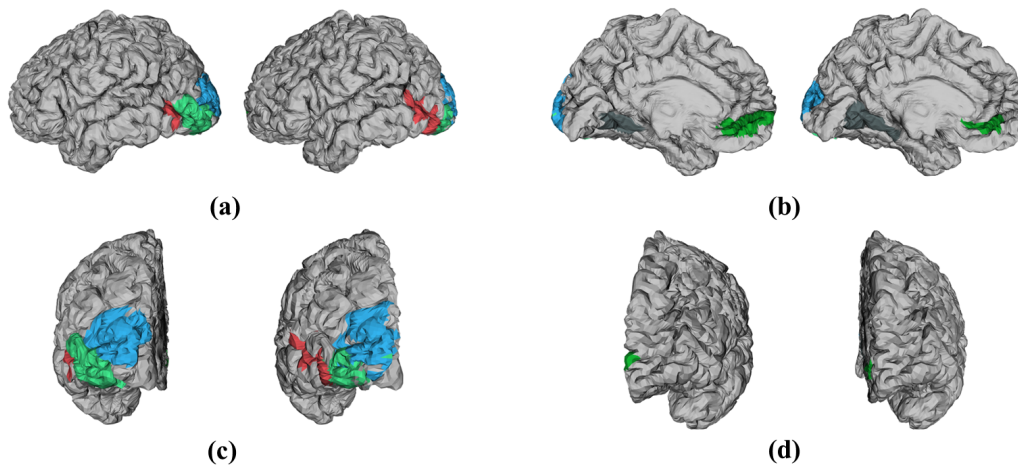


Fig. 4.28: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Brainnetome del lado izquierdo. (Elaboración propia).

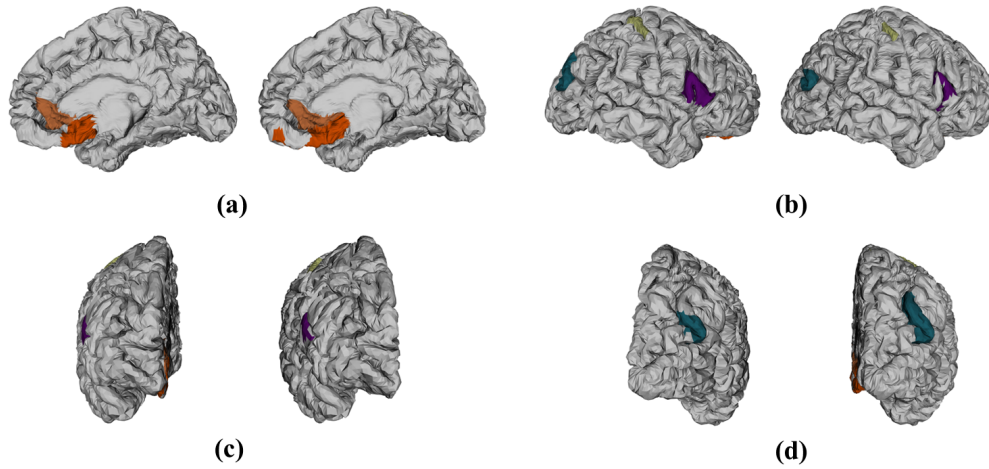


Fig. 4.29: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Lefranc del lado izquierdo. (Elaboración propia).

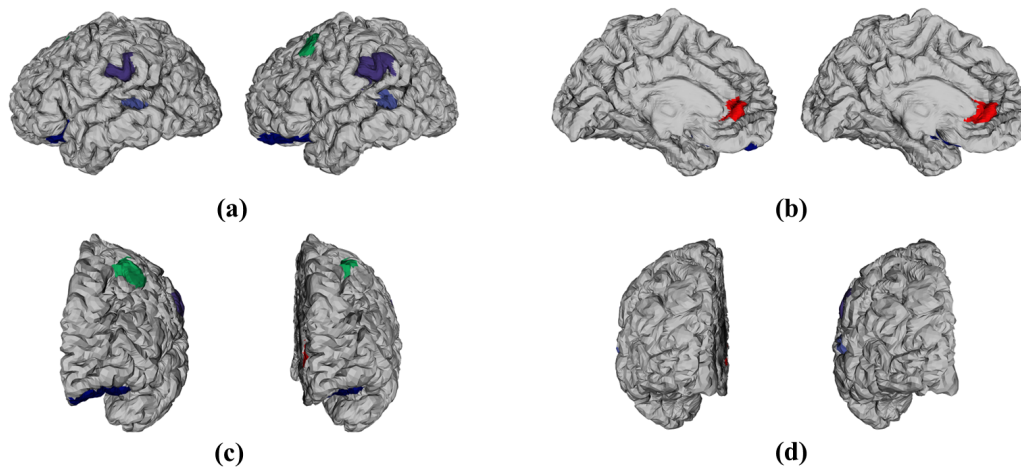


Fig. 4.30: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Lefranc del lado derecho. (Elaboración propia).

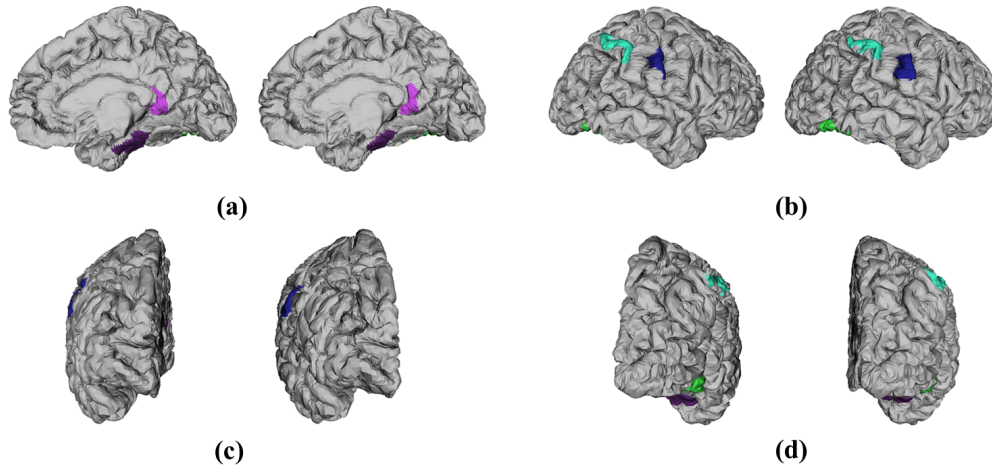


Fig. 4.31: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Narciso del lado izquierdo. (Elaboración propia).

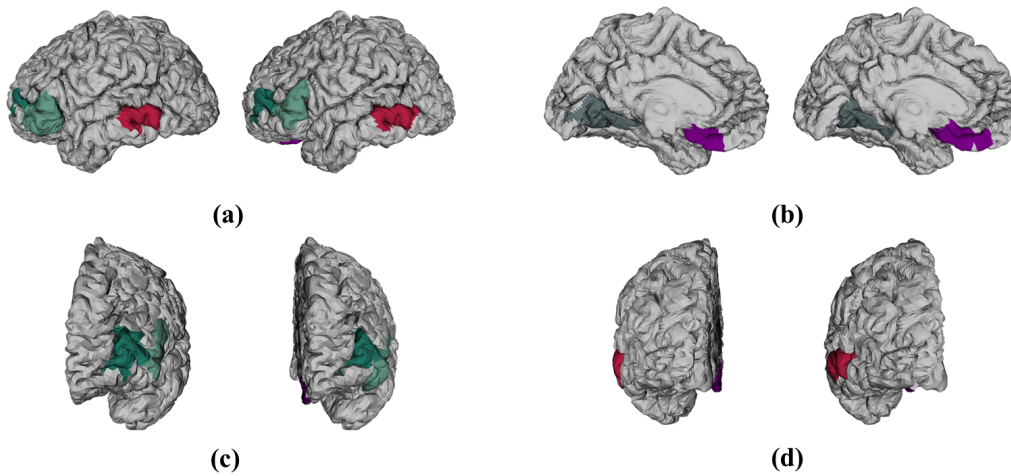


Fig. 4.32: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Narciso del lado derecho. (Elaboración propia).

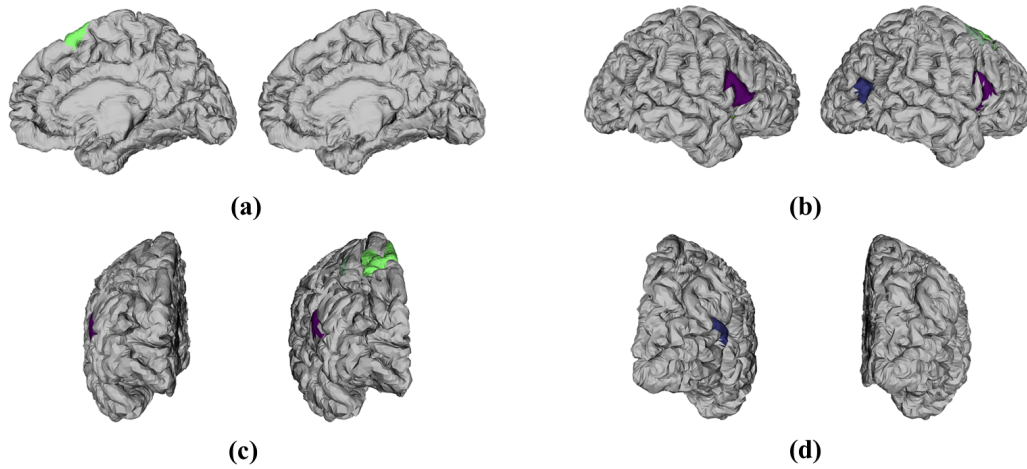


Fig. 4.33: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Richards del lado izquierdo. (Elaboración propia).

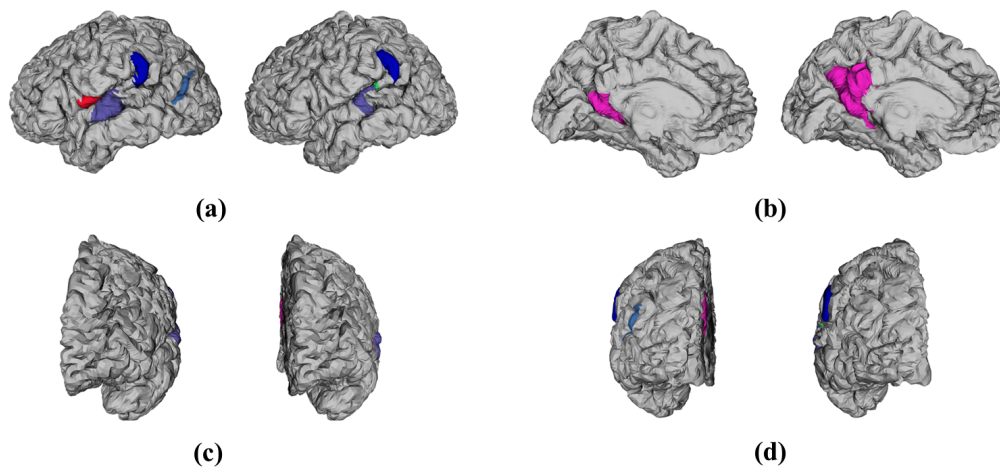


Fig. 4.34: En esta imagen, podemos apreciar la comparativa de las parcelas Generadas con la parcelación de Richards del lado derecho. (Elaboración propia).

Capítulo 5. Discusión, Conclusiones y Trabajo Futuro

5.1 Introducción

En este capítulo, se presenta el análisis de los resultados obtenidos en la investigación, donde de manera crítica se analizan todos los puntos a resolver en trabajos futuros, resaltando las limitaciones y áreas de mejora que deben ser consideradas para futuras investigaciones. Se plantearán estrategias para optimizar la obtención de clusters representativos y mejorar la robustez de los algoritmos empleados, y además, se describe el trabajo futuro propuesto, correspondiente a una parcelación jerárquica multiescala de estos datos.

5.2 Discusión

5.2.1 Preprocesamientos

Se escogió un umbral de $40mm$ para eliminar las fibras más pequeñas (entre $30mm$ y $40mm$) debido a la complejidad espacial que representa incluir los 3 millones de fibras. Estos valores se escogieron intentando conservar la mayor cantidad de fibras posibles utilizando al máximo la capacidad de cómputo del hardware. No considerar este umbral hace imposible la aplicación del algoritmo con la capacidad disponible.

5.2.2 Respecto a la Creación del Atlas Representativo de Fibras

El proceso de *clustering intrasujeto* utilizando el algoritmo FFClust permitió identificar una estructura consistente y compacta en los tractogramas individuales. La elección de los parámetros $d_{Rmax} = 6$ mm y $d_{Mmax} = 6$ mm, respaldada por investigaciones previas en la base de datos ARCHI/HARDI, parece adecuada para lograr clusters coherentes a nivel anatómico para las 834057 fibras que en promedio hay en la población de sujetos. Es posible explorar nuevos métodos para definir qué distancias son correctas para d_{Rmax} y d_{Mmax} . Por otro lado, la utilización del método de codo para seleccionar el número óptimo de clusters iniciales resultó efectiva, debido a la aparente compacticidad que se logró en los clusters. Podemos apreciar esto en la figura (4.18.c).

La etapa de *clustering intersujeto* permitió obtener clusters representativos a partir de los centroides de los clusters individuales. Este enfoque, basado en el algoritmo QuickBundles, preserva la individualidad de cada tractograma. La utilización de un umbral de representatividad ($rep_thresh = 88\%$) asegura que solo los clusters con alta presencia intersujeto sean considerados. Por otro lado, podemos ver que la utilización de un umbral $\theta = 10mm$ nos entrega un muy buen agrupamiento de centroides, como podemos apreciar en la figura (c.4.20). Además, si vemos con cuidado estos clusters, podemos ver que se identificó un cluster de centroides que pueden pertenecer al fascículo longitudinal superior. Este proceso garantiza una reducción del ruido al no considerar fibras individuales sino agrupaciones representativas en cada sujeto. Cabe mencionar que es posible optimizar los parámetros de esta etapa. En especial, el parámetro $theta_QB$.

La preferencia por QuickBundles en esta fase intersujeto radica en su capacidad para agrupar los centroides sin descartar ninguna representación intrasujeto, capturando una estructura de conectividad completa y representativa. Este paso final del proceso de clustering resalta la eficacia de combinar ambos algoritmos: FFClust para el análisis intrasujeto, eliminando el ruido, y QuickBundles para el análisis intersujeto, donde se preserva la representatividad total de las fibras mediante los centroides.

5.2.3 Respecto al Criterio de Manejo del Traslape

La identificación de los DC aportó una segmentación preliminar que permitió distinguir las zonas con una mayor densidad de conexiones dentro de cada subparcela. Este paso fue crucial para limitar el número de triángulos asignados a múltiples subparcelas y lograr una partición más estructurada.

La matriz de traslape y el grafo generado a partir de ella facilitaron la identificación de subparcelas con un alto grado de superposición, proponiendo la fusión de aquellas conectadas en el grafo y que superaban un umbral de traslape significativo. Este enfoque permitió simplificar la parcelación, eliminando redundancias en las subparcelas. Sin embargo, la dependencia en el umbral de traslape para definir la fusión podría limitar la precisión en áreas donde las conexiones son menos densas, pero igualmente importantes. Se deben buscar maneras para fijar este umbral de manera adaptativa, o utilizar enfoques más adaptativos que no requieran la definición de un umbral.

5.2.4 Respetto a la Dilatación Morfológica

La identificación y preservación de las componentes conexas principales para cada subparcela resultó efectiva para generar un mapa de subparcelas continuo y cohesionado, permitiendo expandir las regiones desde un solo lugar. Este procedimiento eliminó fragmentaciones y permitió definir claramente las zonas anatómicas, dando lugar a una parcelación cortical más representativa. Sin embargo, cabe señalar que la eliminación de componentes más pequeñas podría llevar a la pérdida de información anatómica relevante en las parcelaciones finales.

La dilatación morfológica aplicada iterativamente contribuyó a una distribución espacialmente coherente de las subparcelas, expandiendo sus límites. La estrategia de dilatar primero las subparcelas más pequeñas resultó particularmente útil para preservar la integridad de subparcelas con menor densidad de triángulos, lo que causó menos desvanecimiento de las parcelas más pequeñas. El criterio de limitación en el caso de que dos triángulos asignados se encontraran también ayudó a mejorar la confiabilidad del algoritmo, debido a que ahora, por cada iteración de la dilatación se revisa si cierto triángulo pertenece a cierto cluster antes de expandir dicha zona hacia cierto triángulo, etiquetando el cluster con la zona que más se repite en dicho triángulo en los datos crudos. De esta manera, se evitan errores de asignación. Esto representa la mejora al ADP de Joaquín Molina, cumpliendo con el objetivo inicial.

5.2.5 Respetto a la Comparación con las Parcelaciones del Estado del Arte

De la tabla 4.2 podemos observar que las parcelaciones del estado del arte contienen una mayor cantidad de parcelas en el hemisferio izquierdo (LH). Esto puede tener una base neurobiológica asociada a una **mayor complejidad estructural y funcional** de este hemisferio. Desde una perspectiva citoarquitectónica, esta asimetría podría reflejar una **mayor densidad y dispersión de la distribución de centros de densidad de WM** en ciertas regiones del LH, lo cual permite distinguir más sectores con patrones de conectividad únicos y reproducibles. Un mayor número de fibras o una organización más modular favorecería la identificación de límites entre regiones funcionales, generando una parcelación más fragmentada.

Esta mayor fragmentación podría deberse a la **especialización funcional del hemisferio izquierdo**, particularmente en funciones lingüísticas, analíticas y motoras finas, que requieren una arquitectura cortical más refinada y subdividida. Estudios estructurales mediante resonancia magnética han mostrado asimetrías volumétricas y de forma en regiones como el **planum temporale** [52] y el **área de Broca** [53], las cuales suelen estar más desarrolladas o más plegadas en comparación con sus homólogas derechas.

Además del coeficiente de DICE, se incorporaron métricas normalizadas de similitud (Ecuaciones 3.3.5 y 3.3.6) que cuantifican la proporción de subparcelas significativamente similares, ofreciendo una visión sintética de la calidad del emparejamiento entre parcelaciones. Los resultados obtenidos demuestran una alta correspondencia entre la parcelación generada mediante el algoritmo propuesto y las parcelaciones del estado del arte, en particular con el atlas de Richards, donde se observaron los mayores índices de similitud tanto hemisféricos como globales. Esta coincidencia sugiere que la estrategia basada en dilatación estructural permite capturar patrones de conectividad similares, pese a ser de bases de datos distintas.

5.3 Conclusión

En este trabajo, se desarrolló un algoritmo de parcelación cortical basado en conectividad estructural utilizando datos de tractografía probabilística y mallados de la superficie cortical de 100 sujetos de la base de datos HCP-YA. A través de un enfoque basado en clustering intrasujeto e intersujeto, se logró obtener clusters representativos de fibras de la población de sujetos, integrando técnicas robustas como FFClust y QuickBundles. Estos métodos permitieron reducir significativamente el ruido en los tractogramas individuales y agrupar las fibras de manera coherente, preservando la representatividad anatómica y funcional de las conexiones, cumpliendo así el primer objetivo.

El método propuesto para el manejo de traslapes (correspondiente al propuesto por Narciso López en [11], con algunos cambios) también resultó efectivo. La identificación de áreas de alta densidad de conexiones y la integración de una matriz de traslape entre estas permitió reducir la redundancia entre subparcelas y obtener una segmentación más coherente y anatómicamente representativa. Finalmente, se logró aportar a la solución ya existente creando un algoritmo de dilatación que considerara los datos crudos al momento de designar los triángulos que se van agregando a las parcelas a medida que avanza el algoritmo. La principal diferencia entre este enfoque y el enfoque de Narciso es que la agrupación de parcelas se basa en el perfil completo de traslape (una fila o columna), pudiendo considerar la agrupación del conjunto de parcelas más traslapadas en vez de agrupar por medio de pares de parcelas más traslapadas.

Finalmente, se logró cuantificar cada etapa mediante una gran variedad de histogramas, los cuales serán de gran ayuda al momento de comparar la salida obtenida a distintos parámetros iniciales. Además, la alta cantidad de parcelas finales obtenidas nos va a permitir generar parcelaciones jerárquicas más robustas, ya que una alta cantidad de parcelas implica que podemos generar una jerarquía desde un alto nivel de detalle.

5.4 Trabajo Futuro

Es necesario considerar los 3 millones de fibras completos en vez de realizar una etapa de filtrado, con el fin de evitar la eliminación de fibras y así obtener un análisis más completo. Esto se puede lograr utilizando un equipo que tenga una mayor cantidad de memoria RAM, y optimizando la forma en que se escriben los archivos. Una manera es realizar este proceso de escritura por medio de batches en vez de cargar toda la información de los centroides y clusters en la memoria RAM.

Mientras más compactos sean los clusters, más robusta será la parcelación final. Métricas de compacticidad de clusters como el índice de Davies–Bouldin [54] nos pueden ayudar a encontrar parámetros que logren obtener clusters tanto intrasujeto como intersujeto más compactos. Luego, es posible aplicar clustering a aquellos clusters que tengan mala compacticidad, generando así nuevos clusters más compactos.

Se debe encontrar alguna metodología para escoger los umbrales óptimos de definición de centros de densidad *thresh_DC* y el umbral de traslape significativo *overlap_thresh*, sin embargo, la falta de un ground truth dificulta esta tarea. También se debe explorar la posibilidad de aplicar enfoques basados en la distancia geodésica entre los centros de gravedad de los centros de densidad de las subparcelas, para evaluar qué parcelas deben juntarse según qué tan lejanos son los puntos donde se concentran más fibras. Es posible también mejorar el algoritmo de dilatación que se presenta en la metodología, creando bordes para la dilatación en base a la envolvente conexa de los triángulos que están contenidos en las subparcelas iniciales, evitando así que estas salgan hacia sectores donde dicha parcela no tiene conectividad, mejorando la robustez y confiabilidad del algoritmo. Es posible hacer esto condicionando los triángulos vecinos por cada nivel de iteración.

Como trabajo futuro, se plantea integrar este mismo con el algoritmo de distancia geodésica que presenta Yarelis Prieto [13] junto a un enfoque basado en la similitud entre los perfiles de conectividad de las parcelas. La idea es computar matrices de conectividad de las parcelas obtenidas, con el fin de evaluar qué parcelas tienen un patrón de conectividad similar. Estas matrices pueden ponderarse con una nueva matriz de distancias, que almacene la distancia geodésica entre los centros de gravedad de los centros de densidad. Es posible generar un grafo de afinidad a partir de la matriz ponderada, y generar parcelaciones multiescala a partir de este, utilizando clustering jerárquico aglomerativo. Es posible también encontrar la conectividad para los distintos niveles de granularidad obtenidos de este procedimiento.

Capítulo 5. Glosario

ADP	Algoritmo de Parcelación Cerebral.
CSD	Deconvolución Esférica Restringida , en inglés, Constrained Spherical Deconvolution
dMRI	Imagen de Resonancia Magnética por Difusión , en inglés, Diffusion Magnetic Resonance Imaging.
DSI	Espectro de Difusión , en inglés, Diffusion Spectrum Image
DTI	Modelo de Tensor de Difusión , en inglés, Diffusion Tensor Imaging
FA	Anisotropía Fraccional , en inglés, Fractional Anisotropy
fMRI	Imagen de Resonancia Magnética Funcional , en inglés, Functional Magnetic Resonance Imaging.
GM	Materia Gris , en inglés, Gray Matter.
HARDI	Imagen de Alta Resolución Angular , en inglés, High Angular Resolution Diffusion Imaging
HCP	Proyecto del Conectoma Humano , en inglés, Human Connectome Project.
HCP-YA	Proyecto del Conectoma Humano young adult , en inglés, Human Connectome Project Young Adult.
LH	Hemisferio Cerebral Izquierdo , en inglés, Left Hemisphere
LR	Fibras Iterhemisfericas de Izquierdo a Derecho , en inglés, Left to Right Intrahemispheric Fiber.
LH	Hemisferio Cerebral Derecho , en inglés, Right Hemisphere
RL	Fibras Iterhemisfericas de Derecho a Izquierdo , en inglés, Right to Left Intrahemispheric Fiber.
sMRI	Imagen de Resonancia Magnética Estructural , en inglés, Structural Magnetic Resonance Imaging.
WM	Materia Blanca , en inglés, White Matter.

Capítulo 6. Referencias

- [1] R. C. Craddock, S. Jbabdi, C. Yan, J. T. Vogelstein, F. X. Castellanos, A. Di Martino, C. Kelly, K. Heberlein, S. Colcombe, and M. P. Milham, “Imaging human connectomes at the macroscale,” *Nature Methods*, vol. 10, no. 6, pp. 524–539, 2013.
- [2] K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth, 1909, oCLC 14796098.
- [3] J. M. P.J. Basser and D. Lebihan, “Estimation of the effective self-diffusion tensor from the mri spin echo,” *Journal of Magnetic Resonance, Series B*, vol. 103, no. 3, pp. 247–254, 1994.
- [4] P. Moghimi, A. T. Dang, Q. Do, T. I. Netoff, K. O. Lim, and G. Atluri, “Evaluation of functional mri-based human brain parcellation: a review,” *Journal of Neurophysiology*, vol. 128, no. 1, pp. 197–217, 2022.
- [5] A. M. Dale, B. Fischl, and M. I. Sereno, “Cortical surface-based analysis: I. segmentation and surface reconstruction,” *NeuroImage*, vol. 9, no. 2, pp. 179–194, 1999. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1053811998903950>
- [6] C. Román, “Estudio de fibras de la materia blanca superficial en datos del proyecto conectoma humano,” Ph.D. dissertation, Universidad de Concepción, Concepción, Chile, 2021. [Online]. Available: <http://repositorio.udec.cl/jspui/handle/11594/5382>
- [7] K. Ugurbil, D. V. Essen *et al.*, “Human connectome project young adult study,” <https://www.humanconnectome.org/study/hcp-young-adult>, 2017, accessed: Nov 25, 2023.
- [8] J. I. Molina Muñoz, “Parcelación cortical inter-sujeto basada en conectividad cerebral y clustering jerárquico,” 2022, tesis presentada para optar al título de Ingeniero Civil Biomédico. [Online]. Available: <https://repositorio.udec.cl/handle/11594/11004>
- [9] A. Vázquez, N. López-López, A. Sánchez, J. Houenou, C. Poupon, J.-F. Mangin, C. Hernández, and P. Guevara, “Ffclust: Fast fiber clustering for large tractography datasets for a detailed study of brain connectivity,” *NeuroImage*, vol. 220, p. 117070, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1053811920305565>
- [10] E. Garyfallidis, M. Brett, M. M. Correia, G. B. Williams, and I. Nimmo-Smith, “Quickbundles, a method for tractography simplification,” *Frontiers in neuroscience*, vol. 6, p. 175, 2012.

- [11] N. López-López, A. Vázquez, J. Houenou, C. Poupon, J.-F. Mangin, S. Ladra, and P. Guevara, “From coarse to fine-grained parcellation of the cortical surface using a fiber-bundle atlas,” *Frontiers in Neuroinformatics*, vol. 14, p. 32, 2020.
- [12] J. Molina, C. Mendoza, C. Román, J. Houenou, C. Poupon, J. F. Mangin, W. El-Deredy, C. Hernández, and P. Guevara, “Group-wise cortical parcellation based on structural connectivity and hierarchical clustering,” in *18th International Symposium on Medical Information Processing and Analysis*, vol. 12567. SPIE, 2023, pp. 172–181.
- [13] Y. Prieto, J. Molina, M. Otero, J.-F. Mangin, C. Hernández, W. El-Deredy, and P. Guevara, “Multiscale cortical parcellation based on geodesic distance and hierarchical clustering,” pp. 1–5, 2023.
- [14] “Brainvisa / Anatomist,” <https://brainvisa.info>, 2023, página web de Brainvisa y Anátomist.
- [15] W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit (4th ed.)*. Kitware, 2006.
- [16] M. H. Ross and W. Pawlina, *Histología: Texto y Atlas Color con Biología Celular y Molecular*, 6th ed. Madrid, España: Editorial Médica Panamericana, 2014.
- [17] R. Snell, *Neuroanatomía Clínica*, 8th ed. Buenos Aires: Médica Panamericana, 2003.
- [18] W. Webb, “Organization of the nervous system,” in *Neurology for the Speech-Language Pathologist*. Elsevier, 2017, pp. 13–43.
- [19] M. L. Ries, C. M. Carlsson, H. A. Rowley, M. A. Sager, C. E. Gleason, S. Asthana, and S. C. Johnson, “Magnetic resonance imaging characterization of brain structure and function in mild cognitive impairment: A review,” *Journal of the American Geriatrics Society*, vol. 56, no. 5, pp. 920–934, 2008.
- [20] C. Thomsen, O. Henriksen, and P. Ring, “In vivo measurement of water self diffusion in the human brain by magnetic resonance imaging,” *Acta Radiol*, vol. 28, no. 3, pp. 353–361, May-Jun 1987.
- [21] H. Johansen-Berg and T. Behrens, in *Diffusion MRI: From Quantitative Measurement to In-Vivo Neuroanatomy*, 1st ed. Academic Press, 2009, ch. 2.
- [22] B. J. Jellison, A. S. Field, J. Medow, M. Lazar, M. S. Salamat, and A. L. Alexander, “Diffusion tensor imaging of cerebral white matter: A pictorial review of physics, fiber tract anatomy, and tumor imaging patterns,” *American Journal of Neuroradiology*, vol. 25, no. 3, pp. 356–369, 2004. [Online]. Available: <https://www.ajnr.org/content/25/3/356>

- [23] P. J. Basser, J. Mattiello, and D. LeBihan, “Estimation of the effective self-diffusion tensor from the nmr spin echo,” *Journal of Magnetic Resonance, Series B*, vol. 103, no. 3, pp. 247–254, 1994.
- [24] J. Zhang, M. Aggarwal, and S. Mori, “Structural insights into the rodent cns via diffusion tensor imaging,” *Trends in Neurosciences*, vol. 35, no. 7, pp. 412–421, Jul 2012.
- [25] F. Szczepankiewicz, S. Lasič, D. van Westen, P. C. Sundgren, E. Englund, C.-F. Westin, F. Ståhlberg, J. Lätt, D. Topgaard, and M. Nilsson, “Quantification of microscopic diffusion anisotropy disentangles effects of orientation dispersion from microstructure: Applications in healthy volunteers and in brain tumors,” *NeuroImage*, vol. 104, pp. 241–252, 2015.
- [26] V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff, “Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging,” *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1377–1386, 2005.
- [27] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, “Regularized, fast, and robust analytical q-ball imaging,” *Magnetic Resonance in Medicine*, vol. 58, no. 3, pp. 497–510, 2007.
- [28] M. Descoteaux, *High Angular Resolution Diffusion Imaging (HARDI)*. John Wiley Sons, Ltd, 2015, pp. 1–25. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/047134608X.W8258>
- [29] C. Mendoza Sánchez, “Algoritmos para el filtrado de fibras cerebrales y su efecto en la reproducibilidad de la segmentación automática,” Master’s thesis, Universidad de Concepción, Concepción, Chile, 2024. [Online]. Available: <http://repositorio.udec.cl/jspui/handle/11594/11774>
- [30] D. Sculley, “Web-scale k-means clustering.” New York, NY, USA: Association for Computing Machinery, 2010. [Online]. Available: <https://doi.org/10.1145/1772690.1772862>
- [31] T. M. Kodinariya, P. R. Makwana *et al.*, “Review on determining number of cluster in k-means clustering,” *International Journal*, vol. 1, no. 6, pp. 90–95, 2013.
- [32] E. Garyfallidis, M.-A. Côté, F. Rheault, and M. Descoteaux, “Quickbundlesx: sequential clustering of millions of streamlines in multiple levels of detail at record execution time,” *24th International Society of Magnetic Resonance in Medicine (ISMRM)*, 2016.
- [33] L. Tait, A. Özkan, M. J. Szul, and J. Zhang, “A systematic evaluation of source reconstruction of resting meg of the human brain with a new high-resolution atlas: Performance, precision, and parcellation,” *Human Brain Mapping*, vol. 42, no. 14, pp. 4685–4707, 2021.

- [34] A. N. Santana *et al.*, “Using deep learning and resting-state fmri to classify chronic pain conditions,” *Frontiers in Neuroscience*, vol. 13, 2019. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnins.2019.01313>
- [35] I. Diez, P. Bonifazi, I. Escudero, B. Mateos, M. A. Muñoz, S. Stramaglia, and J. M. Cortes, “A novel brain partition highlights the modular skeleton shared by structure and function,” *Scientific reports*, vol. 5, no. 1, p. 10532, 2015.
- [36] S. Lefranc, “Groupwise connectivity-based parcellation of the whole human cortical surface using watershed-driven dimension reduction,” *Medical Image Analysis*, vol. 30, pp. 11–29, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841516000049>
- [37] D. Moreno-Dominguez, A. Anwander, and T. R. Knösche, “A hierarchical method for whole-brain connectivity-based parcellation,” *Human Brain Mapping*, vol. 35, no. 10, pp. 5000–5025, 2014. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.22528>
- [38] S. N. Sotiropoulos, S. Moeller, S. Jbabdi, J. Xu, J. L. Andersson, E. J. Auerbach, E. Yacoub, D. Feinberg, K. Setsompop, L. L. Wald, T. E. J. Behrens, K. Ugurbil, and C. Lenglet, “Effects of image reconstruction on fiber orientation mapping from multichannel diffusion mri: Reducing the noise floor using sense,” *Magnetic Resonance in Medicine*, vol. 70, no. 6, pp. 1682–1689, 2013. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.24623>
- [39] J. L. Andersson, S. Skare, and J. Ashburner, “How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging,” *NeuroImage*, vol. 20, no. 2, pp. 870–888, 2003.
- [40] J. L. Andersson and S. N. Sotiropoulos, “Non-parametric representation and prediction of single- and multi-shell diffusion-weighted mri data using gaussian processes,” *NeuroImage*, vol. 122, pp. 166–176, 2015.
- [41] J. Andersson and S. Sotiropoulos, “An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging,” *NeuroImage*, vol. 125, pp. 1063–1078, 2016.
- [42] J.-D. Tournier, F. Calamante, and A. Connelly, “Mrtrix3: A fast, flexible and open software framework for medical imaging and diffusion mri,” *NeuroImage*, vol. 202, p. 116137, 2019.
- [43] D. Christiaens, M. Reisert, T. Dhollander, S. Sunaert, P. Suetens, and F. Maes, “Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model,” *NeuroImage*, vol. 123, pp. 89–101, 2015.

- [44] J.-D. Tournier, F. Calamante, and A. Connelly, “Robust determination of the fibre orientation distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution,” *NeuroImage*, vol. 35, no. 4, pp. 1459–1472, 2007.
- [45] H. U. Martinos Center for Biomedical Imaging, MGH, “Freesurfer tutorial: Output data - freeview,” https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/OutputData_freeview, 2024, accessed: 2024-09-22.
- [46] BrainVisa, “Freesurfer to brainvisa conversion pipeline,” <https://brainvisa.info/axon-4.7/en/processes/freesurferToBrainvisaConversionPipeline.html>, 2024, accessed: 2024-09-22.
- [47] “Wavefront obj file format,” Library of Congress, accessed: 22/09/24. [Online]. Available: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000507.shtml>
- [48] F. Silva, “Método de parcelación cortical basado en una representación mediante grafos de las conexiones cerebrales,” Memoria de Título, Universidad de Concepción, Facultad de Ingeniería, Departamento de Ingeniería Eléctrica, Abril 2019.
- [49] F. Silva, M. Guevara, C. Poupon, J.-F. Mangin, C. Hernández, and P. Guevara, “Cortical surface parcellation based on graph representation of short fiber bundle connections,” in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, pp. 1479–1482.
- [50] L. Fan, H. Li, J. Zhuo, Y. Zhang, J. Wang, L. Chen, Z. Yang, C. Chu, S. Xie, A. R. Laird, P. T. Fox, S. B. Eickhoff, C. Yu, and T. Jiang, “The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture,” *Cerebral Cortex*, vol. 26, no. 8, pp. 3508–3526, 07 2016. [Online]. Available: <https://doi.org/10.1093/cercor/bhw157>
- [51] R. Richards, “Adaptación de métodos de parcelación cortical para el soporte de múltiples bases de datos de neuroimágenes,” 2022, memoria de título para optar al título de Ingeniero Civil Biomédico, Universidad de Concepción, 2022.
- [52] J. P. Larsen, H. Odegaard, T. H. Grude, and T. Høien, “Magnetic resonance imaging - a method of studying the size and asymmetry of the planum temporale,” *Acta Neurologica Scandinavica*, vol. 80, no. 5, pp. 438–443, 1989.
- [53] A. L. Foundas, C. M. Leonard, R. L. Gilmore, E. B. Fennell, and K. M. Heilman, “Pars triangularis asymmetry and language dominance,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 2, pp. 719–722, 1996.

- [54] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.

UNIVERSIDAD DE CONCEPCION – FACULTAD DE INGENIERIA
RESUMEN DE MEMORIA DE TITULO

Departamento : Departamento de Ingeniería Eléctrica
Carrera : Ingeniería Civil Biomédica
Nombre del memorista : **Martín Antonio López Aguilera**
Título de la memoria : Algoritmo de Parcelación Cerebral Intersujeto basado en Datos del Proyecto del Conectoma Humano
Fecha de la presentación oral : 28 de Abril
Profesor(es) Guía : **Pamela Guevara Alvez, Cecilia Hernández Rivas**
Profesor(es) Revisor(es) : Miguel Figueroa Toro
Concepto :
Calificación :

Resumen

En ese estudio se desarrolla una metodología para la parcelación cortical de la corteza cerebral humana basada en Imágenes de Resonancia Magnética (MRI), con el objetivo de analizar el conectoma humano. El enfoque consiste en la segmentación de la corteza cerebral en regiones con patrones de conectividad estructural, derivados de tractogramas y mallados corticales. Los tractogramas son representaciones tridimensionales de las trayectorias de las fibras cerebrales, y los mallados representan la superficie de la corteza cerebral. Los tractogramas y mallados se calcularon a partir de imágenes de Resonancia Magnética Estructural (sMRI) y Resonancia Magnética por Difusión (dMRI) pertenecientes a la base de datos Human Connectome Project Young Adult (HCPYA). Estos tienen 3 millones de fibras y 81920 triángulos por hemisferio en sus 100 sujetos, respectivamente. Como preprocesamiento, se filtran las fibras pequeñas en base a un umbral, y se aplica un filtro media móvil para suavizar. Luego, se identifican los clusters representativos de la población de sujetos mediante dos etapas de clustering de fibras. Primero se aplica el algoritmo FFClust a los tractogramas de cada sujeto, obteniendo clusters de fibras intrasujeto y su centroide respectivo. Luego, se migran estos centroides a un espacio común, donde se obtienen clusters de centroides al aplicar el algoritmo QuickBundles. Para obtener clusters representativos de la población de 100 sujetos, se seleccionaron aquellos clusters intersujeto que contienen centroides de al menos 88 sujetos. Finalmente, se almacena para cada sujeto los clusters de fibras (obtenidos de FFClust) que corresponden a los clusters de centroides (obtenidos de QuickBundles), obteniendo así sólo los clusters representativos de cada sujeto. Para cada sujeto, se obtuvieron 1025 clusters en promedio, en un rango de 664 a 1053 clusters por sujeto. Para asignar estas fibras a la corteza, el algoritmo proyecta cada fibra sobre el mallado cortical, identificando el triángulo inicial y final de cada trayectoria. Esto permite etiquetar triángulos con sus respectivos clusters, formando subparcelas. Estas subparcelas están fuertemente superpuestas, lo que significa que tienen triángulos en común. Para solucionar esto, se determina la probabilidad de que un triángulo sea de cierta subparcela, contando cuántas veces una fibra de cierta subparcela cruza el triángulo y su vecindad. Luego, se definen los centros de densidad como aquellos triángulos que tienen mayor probabilidad en cada subparcela. Las subparcelas que contienen cierto traslape entre sus centros de densidad son combinadas como una sola. Finalmente, se calculan las componentes conexas principales de cada subparcela, y se realiza dilatación morfológica iterativa para obtener parcelas finales. La parcelación final incluye 189 parcelas en el hemisferio izquierdo y 139 en el derecho.