



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



# **ANÁLISIS DE IMÁGENES DE RESONANCIA MAGNÉTICA POR DIFUSIÓN DE ADOLESCENTES CON AUTISMO**

POR

**Natalia Valentina Vidal Barraza**

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

Profesora guía  
Pamela Guevara Alvez

Comisión  
1. Cecilia Hernández  
2. Jaime Pinto

Agosto 16, 2023  
Concepción  
(Chile)

© 2023 Natalia Valentina Vidal Barraza

© 2023 Natalia Valentina Vidal Barraza

Ninguna parte de esta memoria puede reproducirse o transmitirse bajo ninguna forma o por ningún medio o procedimiento, sin permiso por escrito del autor.

## Agradecimientos

*De manera sincera, expreso mi profundo agradecimiento a mis padres Leontina y Ramón, quienes con su amor incondicional, valiosas enseñanzas, constante apoyo y firme confianza en mí, han forjado la persona que soy hoy. Asimismo, no puedo dejar de mencionar a mis hermanos, Javier y Felipe, cuya presencia, hermandad y amistad han sido pilares en cada paso de mi vida.*

*Agradezco de corazón a Sebastián, cuya compañía y alegría han sido esenciales para enfrentar los desafíos en el camino. Su perspectiva siempre valiosa y el afecto sincero que compartimos me han ayudado a llegar más lejos y no decaer ante cada bajo en esta curva de aprendizaje.*

*A mis amigas y amigos, pero de manera especial a Cristina, María Paz, Mayra y María Fernanda, por ser parte fundamental de mis vivencias. A lo largo de este tiempo, han demostrado que la amistad está siempre presente para celebrar los triunfos y brindar apoyo en los momentos difíciles. Espero que nuestros caminos continúen cruzándose en las distintas etapas que nos depara la vida.*

*Gracias profesora Pamela y profesor Esteban, por ser referentes e inspiración en mi carrera. Su guía y respaldo han sido fundamentales en mi crecimiento académico en diferentes etapas de mi formación.*

*Por último, quiero agradecer a los miembros del Laboratorio de Análisis de Imágenes Médicas de la Universidad de Concepción, a FONDECYT 1221665, ANILLO ACT210053, Basal FB0008 (AC3E) y Basal FB210017 (CENIA), cuya asistencia hicieron posible llevar a cabo esta investigación.*

## Resumen

El uso de neuroimágenes ha experimentado un crecimiento significativo en las últimas décadas, brindando una amplia gama de técnicas e investigaciones que respaldan el estudio, diagnóstico y seguimiento de enfermedades y trastornos cerebrales.

Esta investigación desarrolla y adapta códigos para trabajar con una base de datos de Imágenes de Resonancia Magnética por Difusión (dMRI) de adolescentes con y sin autismo. A través de un detallado proceso de análisis, se examinará la microestructura de la materia blanca para identificar características distintivas entre los jóvenes con autismo y aquellos neurotípicos. La información obtenida puede mejorar la comprensión y el tratamiento de esta patología, lo que permitirá un diagnóstico más preciso y un tratamiento temprano para mejorar la calidad de vida de las personas con autismo y sus familias.

Mediante la aplicación de conocimientos en álgebra, estadística, imágenes médicas, reconocimiento de patrones y neurociencia, se han desarrollado algoritmos para procesar los datos de tractografía obtenidos previamente de las dMRI de cada sujeto. Estos algoritmos aplican transformaciones espaciales para llevar los datos a un espacio común, remuestrean la cantidad de puntos por fibra, segmentan las fibras en 36 fascículos, revierten las transformaciones mencionadas anteriormente, y obtienen los valores medios de cada fascículo en los mapas de intensidad de ADC, FA y GFA. Luego, a dichos valores medios se les aplicaron tests estadísticos y algoritmos de clasificación mediante Machine Learning con el objetivo de identificar posibles diferencias significativas entre los sujetos.

A pesar de los desafíos enfrentados, esta investigación ha logrado avanzar en el estudio de la materia blanca en la patología mencionada. Si bien los resultados de las pruebas estadísticas no concuerdan completamente con las conclusiones basadas en los modelos de clasificación, se obtuvo información relevante sobre la normalidad de los valores de sujetos con autismo, así como diferencias significativas en fascículos Cingulados, Corticoespinales y Talámicos. También se destacó la importancia de clasificar los sujetos utilizando los valores de GFA y la combinación de valores ADC y GFA.

No obstante, se sugiere realizar estudios adicionales en el futuro, expandiendo la base de datos y considerando aspectos específicos del espectro autista, como el grado de autismo reflejado en dificultades en la comunicación, comportamientos repetitivos, dificultades sociales, sensibilidades sensoriales, entre otros. Esto permitiría obtener resultados más concluyentes y precisos.

## Abstract

The use of neuroimaging has experienced significant growth in recent decades, providing a wide range of techniques and research that support the study, diagnosis, and monitoring of brain diseases and disorders.

This research develops and adapts codes to work with a database of Diffusion Magnetic Resonance Imaging (dMRI) images of adolescents with and without autism. Through a detailed process of analysis, the microstructure of the white matter will be examined to identify distinctive features between young people with autism and neurotypical individuals. The information obtained can improve the understanding and treatment of this condition, enabling more accurate diagnosis and early intervention to enhance the quality of life for individuals with autism and their families.

By applying knowledge in algebra, statistics, medical imaging, pattern recognition, and neuroscience, algorithms have been developed to process the tractography data previously obtained from each subject's dMRI. These algorithms apply spatial transformations to bring the data to a common space, resample the fiber points, segment the fibers into 36 fascicles, reverse the aforementioned transformations, and obtain the mean values of each fascicle in the ADC, FA, and GFA intensity maps. Subsequently, these mean values underwent statistical tests and classification algorithms using Machine Learning to identify potential significant differences between subjects.

Despite the challenges faced, this research has made significant progress in studying the white matter in the mentioned condition. Although the results of the statistical tests do not fully agree with the conclusions based on the classification models, relevant information about the normality of values in subjects with autism was obtained, as well as significant differences in the Cingulate, Corticospinal, and Thalamic fascicles. The importance of classifying subjects using GFA values and the combination of ADC and GFA values was also highlighted.

However, further studies are recommended in the future, expanding the database and considering specific aspects of the autism spectrum, such as the degree of autism reflected in communication difficulties, repetitive behaviors, social challenges, sensory sensitivities, among others. This would allow for more conclusive and precise results.

# Tabla de contenidos

<b>CAPITULO 1. INTRODUCCIÓN.....</b>	<b>13</b>
1.1 INTRODUCCIÓN GENERAL .....	13
1.2 OBJETIVOS .....	13
1.2.1 <i>Objetivo General</i> .....	13
1.2.2 <i>Objetivos Específicos</i> .....	14
1.3 ALCANCES Y LIMITACIONES .....	14
1.4 METODOLOGÍA.....	15
<b>CAPITULO 2. MARCO TEÓRICO .....</b>	<b>16</b>
2.1 INTRODUCCIÓN .....	16
2.2 SISTEMA NERVIOSO CENTRAL.....	16
2.2.1 <i>Estructura interna del cerebro</i> .....	16
2.2.1.1 <i>Materia blanca</i> . . . . .	16
2.3 NEUROLOGÍA .....	17
2.3.1 <i>Trastorno del Espectro Autista</i> .....	17
2.4 RESONANCIA MAGNÉTICA.....	18
2.5 RESONANCIA MAGNÉTICA POR DIFUSIÓN.....	19
2.5.1 <i>Modelos de difusión local y tractografía</i> .....	20
2.5.1.1 <i>Imágenes por Tensor de Difusión</i> . . . . .	21
2.5.1.2 <i>Imágenes de Difusión de Alta Resolución Angular</i> . . . . .	21
2.5.1.3 <i>Imágenes por Q-Ball</i> . . . . .	22
2.5.1.4 <i>Tractografía Cerebral</i> . . . . .	23
2.5.2 <i>Medidas de difusividad y anisotropía</i> .....	23
2.5.2.1 <i>Coficiente de Difusividad Aparente</i> . . . . .	23
2.5.2.2 <i>Anisotropía Fraccional</i> . . . . .	23
2.5.2.3 <i>Anisotropía Fraccional Generalizada</i> . . . . .	24
2.5.3 <i>Sistemas de coordenadas</i> .....	24
2.5.3.1 <i>Talairach</i> . . . . .	25
2.6 PREPROCESAMIENTOS DE LAS DMRI.....	25

2.6.1	<i>Operaciones</i> .....	26
2.7	ESTUDIO DE WM EN PACIENTES CON AUTISMO .....	26
2.8	MULTIPLICACIÓN MATRICIAL .....	28
2.8.1	<i>Transformaciones 3D</i> .....	29
2.8.1.1	Matriz de traslación 3D . . . . .	29
2.8.1.2	Matriz de escalado . . . . .	29
2.8.1.3	Matriz de rotación . . . . .	30
2.9	ESTADÍSTICA .....	30
2.9.1	<i>Estadística descriptiva</i> .....	30
2.9.1.1	Distribución . . . . .	30
2.9.1.2	Varianza . . . . .	31
2.9.2	<i>Estadística inferencial</i> .....	31
2.9.2.1	T-test . . . . .	32
2.9.2.2	Prueba de Mann Whitney U . . . . .	32
2.10	MACHINE LEARNING .....	32
2.10.1	<i>Algoritmos de Aprendizaje Supervisado</i> .....	33
2.10.1.1	Árboles de Decisión . . . . .	33
2.10.1.2	Máquina de Vectores de Soporte . . . . .	33
2.10.1.3	Validación cruzada . . . . .	34
2.10.1.4	Métricas de evaluación . . . . .	34
2.11	DISCUSIÓN .....	35
<b>CAPITULO 3. DEFINICIÓN DEL PIPELINE DE PROCESAMIENTO .....</b>		<b>36</b>
3.1	INTRODUCCIÓN .....	36
3.2	BASE DE DATOS .....	36
3.3	ATLAS DE SEGMENTACIÓN .....	36
3.4	BRAINVISA .....	37
3.5	CÓDIGOS EXISTENTES .....	38
3.5.1	<i>Manipulación de archivos bundles</i> .....	38
3.5.2	<i>Aplicación de transformación espacial</i> .....	38
3.5.3	<i>Segmentación de fascículos</i> .....	39
3.5.4	<i>Obtención de valores promedios de medidas de difusión y anisotropía</i> .....	40
3.6	BIBLIOTECAS DE PYTHON .....	42
3.6.1	<i>DIPY</i> .....	42
3.6.2	<i>SCIPY</i> .....	42

3.6.3	<i>Scikit-learn</i> .....	43
3.7	DISCUSIÓN .....	44
<b>CAPITULO 4. APLICACIÓN DEL PIPELINE .....</b>		<b>45</b>
4.1	INTRODUCCIÓN .....	45
4.2	CONVERSIÓN DE FORMATO DE IMAGEN .....	45
4.3	TRANSFORMACIONES ESPACIALES .....	46
4.3.1	<i>Corrección de rotación errónea</i> .....	47
4.4	REMUESTREO DE FIBRAS .....	48
4.5	SEGMENTACIÓN DE FASCÍCULOS .....	50
4.5.1	<i>Método 'S1'</i> .....	50
4.5.2	<i>Método 'S2'</i> .....	50
4.6	TRANSFORMACIÓN INVERSA .....	52
4.7	APLICACIÓN DE MÁSCARAS .....	52
4.8	TEST ESTADÍSTICOS .....	53
4.8.1	<i>Pruebas de estadística descriptiva</i> .....	53
4.8.2	<i>Pruebas de estadística inferencial</i> .....	54
4.9	MACHINE LEARNING .....	55
4.10	DISCUSIÓN .....	56
<b>CAPITULO 5. RESULTADOS .....</b>		<b>58</b>
5.1	INTRODUCCIÓN .....	58
5.2	MÉTODOS DE SEGMENTACIÓN .....	58
5.3	APLICACIÓN DE MÁSCARAS .....	61
5.4	PROMEDIOS DE DIFUSIVIDAD Y ANISOTROPÍA .....	62
5.5	PRUEBAS ESTADÍSTICAS .....	63
5.6	CLASIFICACIÓN POR APRENDIZAJE AUTOMÁTICO .....	65
5.7	DISCUSIÓN .....	66
<b>CAPITULO 6. CONCLUSIONES .....</b>		<b>69</b>
6.1	DISCUSIÓN .....	69
6.2	CONCLUSIONES .....	70
6.3	TRABAJOS FUTUROS .....	71
<b>CAPITULO 5. GLOSARIO .....</b>		<b>72</b>



**CAPITULO 6. REFERENCIAS ..... 73**

**ANEXO. .... 78**

# Índice de Tablas

4.1	Ángulos de rotación obtenidos . . . . .	48
5.2	Comparación estadística de las medidas de difusión entre sujetos con y sin autismo . . . . .	64
5.3	Resultados de modelos de clasificación por Machine Learning . . . . .	65
A.1	Comparación de porcentajes de fibras segmentadas - Parte 1 . . . . .	78
A.2	Comparación de porcentajes de fibras segmentadas - Parte 2 . . . . .	78
A.3	Comparación de porcentajes de fibras segmentadas - Parte 3 . . . . .	78
A.4	Comparación de porcentajes de fibras segmentadas - Parte 4 . . . . .	78
A.5	Pruebas de normalidad en valores ADC . . . . .	79
A.6	Pruebas de normalidad en valores FA . . . . .	80
A.7	Pruebas de normalidad en valores GFA . . . . .	81
A.8	Pruebas de homogeneidad en valores ADC, FA y GFA . . . . .	82
A.9	Resumen de resultados de pruebas estadísticas descriptivas . . . . .	83
A.10	Resultados de modelos de clasificación aplicados a valores ADC . . . . .	84
A.11	Resultados de modelos de clasificación aplicados a valores FA . . . . .	85
A.12	Resultados de modelos de clasificación aplicados a valores GFA . . . . .	86
A.13	Resultados de modelos de clasificación aplicados a valores ADC y FA . . . . .	87
A.14	Resultados de modelos de clasificación aplicados a valores ADC y GFA . . . . .	88
A.15	Resultados de modelos de clasificación aplicados a valores FA y GFA . . . . .	89
A.16	Resultados de modelos de clasificación aplicados a valores ADC, FA y GFA . . . . .	90
A.17	Hiperparámetros destacados del modelo Decision Trees . . . . .	91
A.18	Hiperparámetros destacados del modelo Support Vector Machine . . . . .	92

# Índice de Figuras

1.1	Pipeline de la metodología . . . . .	15
2.2	Estructuras anatómicas . . . . .	17
2.3	Sección transversal del cerebro . . . . .	18
2.4	Imágenes MRI potenciadas en T1 y T2 . . . . .	19
2.5	Tipos de difusividad . . . . .	20
2.6	Comparación entre Tensor de Difusión y Función de Distribución de Orientación . . . . .	22
2.7	Medidas de difusividad y anisotropía . . . . .	24
2.8	Vista de una MRI en el Espacio MNI y Talairach . . . . .	25
2.9	Pipeline de preprocesamiento . . . . .	26
2.10	Fascículos con anomalías en WM según la literatura . . . . .	27
2.11	Fascículos correlacionados según la literatura al comportamiento autista . . . . .	27
2.12	Vista lóbulos Temporal e Ínsula . . . . .	28
3.13	Atlas de 36 fascículos . . . . .	37
3.14	Pipeline para aplicar una transformación espacial . . . . .	39
3.15	Obtención de valores ADC, FA y GFA de cada fascículo. . . . .	41
4.16	Conversión iterativa de formato . . . . .	45
4.17	Aplicación de la transformación espacial . . . . .	46
4.18	Representación visual de la detección de rotación errónea . . . . .	47
4.19	Ejemplo de corrección rotacional . . . . .	48
4.20	Cantidad de fibras por tractografía . . . . .	49
4.21	Cantidad de fibras según rango de puntos . . . . .	49
4.22	Ejemplo de remuestreo de 21 puntos equidistantes . . . . .	50
4.23	Diagrama de segmentación 'S1' . . . . .	51
4.24	Diagrama de metodología empleada en la segmentación 'S2' . . . . .	51
4.25	Implementación de fascículos como máscaras . . . . .	52
4.26	Implementación de pruebas estadísticas descriptivas . . . . .	53
4.27	Aplicación de pruebas estadísticas inferenciales . . . . .	54

4.28	Aplicación de modelos de clasificación de Machine Learning . . . . .	56
5.29	Comparación visual de fascículos segmentados - Parte 1 . . . . .	59
5.30	Comparación de fascículos segmentados - Parte 2 . . . . .	59
5.31	Comparación de cantidad de fibras por fascículo - Grupo 1 . . . . .	60
5.32	Comparación de cantidad de fibras por fascículo - Grupo 2 . . . . .	60
5.33	Comparación de cantidad de fibras por fascículo - Grupo 3 . . . . .	61
5.34	Ejemplo de aplicación de máscara para el fascículo Talámico Motor Izquierdo . . . . .	61
5.35	Valores promedios de ADC y FA por fascículo . . . . .	62
5.36	Valores promedios de GFA por fascículo . . . . .	63

## Capítulo 1. Introducción

---

### 1.1 Introducción General

El Trastorno del Espectro Autista (TEA) es un trastorno del desarrollo neurológico que afecta a la comunicación social y al comportamiento. Puede verse reflejado por medio de alteraciones en la materia blanca del cerebro, siendo esta una sustancia compuesta por millones de fibras ordenadas en fascículos [1] [2]. El diagnóstico temprano de este trastorno es esencial para iniciar tratamientos y terapias que mejoren la calidad de vida del paciente [3]. Sin embargo, actualmente el diagnóstico se basa en exámenes clínicos que evalúan alteraciones en la interacción social, la comunicación y el comportamiento.

El uso de técnicas de neuroimagen propias de las Imágenes de Resonancia Magnética por Difusión (dMRI) ha cobrado importancia en el estudio del TEA [4]. Pues la tractografía, que es una técnica de representación de fibras a partir del procesamiento de las dMRI, permite obtener información detallada sobre la conectividad neuronal y las diferencias en los fascículos cerebrales de personas autistas y no autistas [5]. Esto podría conducir a un enfoque más cuantitativo para el diagnóstico de este trastorno.

Es esencial tener en cuenta que los métodos de procesamiento de dMRI y de archivos de tractografías deben adaptarse a la base de datos con la que se trabaja. La información proporcionada, los tipos de archivos y las características de los sujetos influyen en la manera en que se deben manipular los datos para obtener resultados precisos y confiables. Adaptar los métodos de procesamiento a las características de la base de datos asegura que el análisis de los datos dMRI sea riguroso.

### 1.2 Objetivos

#### 1.2.1 Objetivo General

Crear y aplicar un pipeline de análisis de fibras cerebrales que una, mejore y/o adapte códigos existentes para llevar a cabo procesamientos a Imágenes de Resonancia Magnética por Difusión para estudios clínicos de una base de datos HARDI de 18 controles y 19 pacientes con autismo.

Este Pipeline debe permitir el análisis de las alteraciones de la materia blanca en pacientes con autismo, además de estar ordenado y documentado para su uso en futuras investigaciones.

## 1.2.2 Objetivos Específicos

- Estudiar pipelines y analizar scripts existentes de procesamiento de datos de dMRI y tractografías, para mejorarlos y/o adaptarlos.
- Implementar y aplicar algoritmos para alinear los datos de tractografía de la base de datos al espacio Talairach, realizar el remuestreo de sus fibras, segmentar fascículos en base a un atlas y obtener medidas promedio de FA, GFA y ADC para los fascículos de cada sujeto.
- Aplicar tests estadísticos y algoritmos de clasificación para determinar alteraciones en los fascículos de la materia blanca en pacientes con autismo, y lograr predecir su diagnóstico según los algoritmos de Machine Learning.
- Evaluar los resultados y concluir al respecto de las características que diferencian ciertos fascículos entre personas con, y sin, autismo.

## 1.3 Alcances y Limitaciones

Para la creación del pipeline, se debe realizar un proyecto depurado y entendible, para eventualmente ser utilizado por futuros lectores. Es requisito estudiar los trabajos previos y proyectos relacionados, con el propósito de fusionarlos y extenderlos.

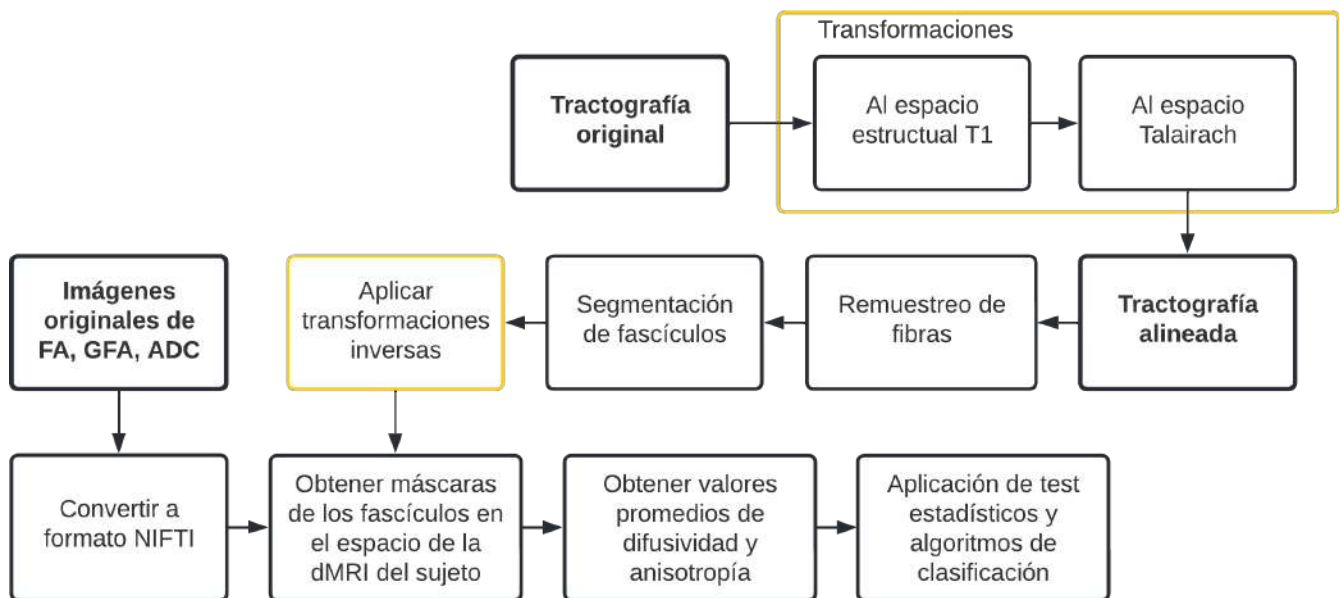
- i. Se utilizará una base de datos HARDI de 18 controles y 19 pacientes con autismo, provista por colaboradores del Centro Neurospin, Francia [6].
- ii. La base de datos contiene las tractografías calculadas, las transformaciones de los espacios a T1 y de T1 a Talairach, así como las imágenes de FA, GFA y ADC de cada sujeto.
- iii. Mayormente, los scripts serán desarrollados y ejecutados en Python, a excepción de la segmentación que está escrita en lenguaje C y requiere del sistema operativo Ubuntu para su ejecución.
- iv. Se usará un atlas de 36 fascículos de fibras largas existente para la segmentación [7].

## 1.4 Metodología

Se analiza el contenido de la base de datos HARDI de adolescentes con autismo [6]. Una vez identificados los datos a utilizar, se aplican dos transformaciones a las tractografías utilizando las matrices disponibles, con el fin de trasladarlas, escalarlas y rotarlas al espacio Talairach.

Posteriormente, se remuestra cada fibra normalizando los datos y se procede a segmentar los 36 fascículos presentes en el atlas de fibras largas [7]. Ya obtenidos los fascículos como archivos bundles, se invierten las transformaciones para que vuelvan al espacio original. Para obtener los valores promedio de difusividad y anisotropía, se convierten los archivos de imagen FA, GFA y ADC del formato GIS al formato NIFTI, y se obtienen máscaras de los fascículos para conocer las intensidades que forman parte de los segmentos.

Con los resultados recopilados, se aplican test estadísticos y modelos de clasificación para analizar posibles alteraciones en la materia blanca. Todo lo anterior se ve representado en la Figura 1.1.



**Fig. 1.1:** Pipeline que representa la metodología empleada para el procesamiento de archivos e información de cada sujeto de la base de datos.

## Capítulo 2. Marco Teórico

---

### 2.1 Introducción

En este capítulo se presenta información relevante sobre el estudio de la materia blanca del cerebro en relación con el trastorno del espectro autista utilizando resonancias magnéticas por difusión (dMRI). El objetivo es obtener una comprensión de las técnicas de neurociencia computacional, especialmente en cuanto a los procesamientos de dMRI y tractografías. La dMRI es una técnica de neuroimagen no invasiva que proporciona información sobre la microestructura de la materia blanca, como la anisotropía fraccional, la difusividad y el coeficiente de difusión aparente, entre otros parámetros. Por otro lado, la tractografía es una técnica para reconstruir las fibras y conexiones del cerebro.

### 2.2 Sistema nervioso central

El sistema nervioso central (SNC), ver Figura 4.22a, es el principal centro de control y coordinación del cuerpo humano. Es responsable de la cognición, memoria, emoción y la conciencia [1].

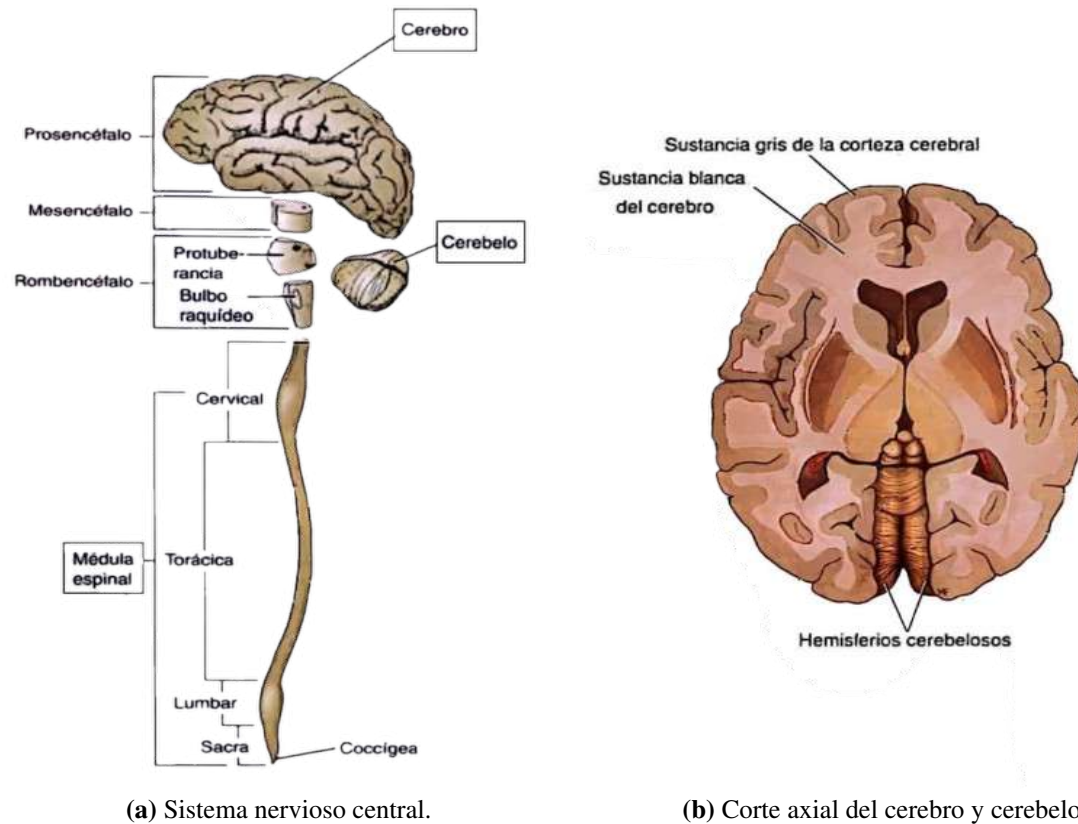
#### 2.2.1 Estructura interna del cerebro

El cerebro es el órgano responsable de dar órdenes al resto del cuerpo, que pueden ser de respuesta, movimiento, sensaciones, emociones, entre otras. Está protegido por estructuras como el cráneo, meninges, líquido cefalorraquídeo y una barrera hematoencefálica. Se divide en los hemisferios derecho e izquierdo, y lo conforman principalmente dos sustancias denominadas Materia Blanca (WM) y Materia Gris (GM), ver Figura 4.22b, [1] [2].

##### 2.2.1.1 Materia blanca

Está formada por conjuntos de fibras nerviosas mielínicas de distintos tamaños. Debido a la presencia de material lipídico en las vainas de mielina que cubren estas estructuras, la sustancia adquiere un color blanco. Las fibras, también conocidas como axones, pueden agruparse para formar tractos o haces, que también se denominan fascículos nerviosos, ver Figura 2.3, [2].





Adaptado de [2]

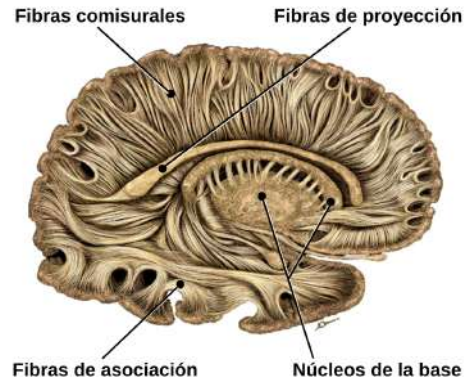
**Fig. 2.2:** Estructuras anatómicas. En a) se ve el Sistema Nervioso central (SNC), conformado principalmente por el cerebro, cerebelo y la médula espinal. En b) se aprecia la Materia Gris (GM) en la corteza cerebral, y la Materia Blanca (WM) en las estructuras subcorticales, también conocidas como Sustancias Gris y Blanca.

## 2.3 Neurología

Es una especialidad médica que estudia, diagnostica y trata los síndromes y trastornos neurológicos [9]. El Síndrome de Tourette, Desorden de Hiperactividad, Déficit Atencional, Trastorno del Espectro Autista (TEA), entre otros, son trastornos que se pueden detectar tanto en la infancia como en la edad adulta y pueden tener un impacto significativo en la vida de las personas [3].

### 2.3.1 Trastorno del Espectro Autista

Según los centros de control de enfermedades, aproximadamente el 1 % de la población mundial, es decir, 75 millones de personas, presenta un trastorno del espectro autista [10]. Los investigadores del área esperan poder estudiarlo mejor basándose en las dMRI.



Adaptado de [8]

**Fig. 2.3:** Ilustración creada por Daniel Casanova Martínez. Presenta una sección transversal del cerebro en la que se aprecia la sustancia blanca conformada por las fibras cerebrales que provienen y/o se dirigen a la corteza.

Se menciona en [4] la notable repercusión en los tractos cerebrales que se encargan de la cognición social en personas con TEA, pero que parecieran contradictorios, ya que existen diversos estudios de TBSS (Estadísticas Espaciales Basadas en Tractos) que muestran un aumento en la anisotropía fraccional en las regiones frontales, parietales y temporales en el caso de los niños, que luego tienden a reducirse en los adultos. Es aún más interesante notar que los cambios que solían observarse en las personas con TEA también están presentes en sus hermanos, que a pesar de compartir dichas características, no poseen un diagnóstico formal. Esto sugiere que la genética juega un papel importante en el trastorno del espectro autista y puede ser una pista en la búsqueda de tratamientos efectivos [3].

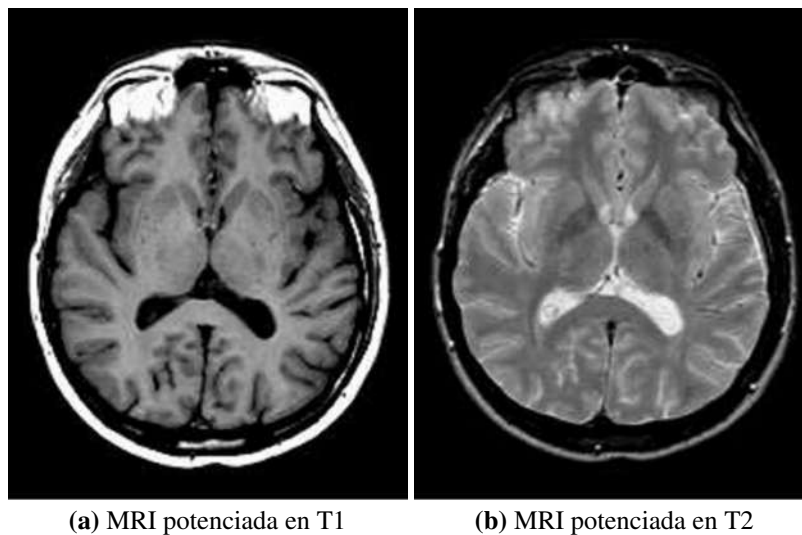
## 2.4 Resonancia Magnética

La Resonancia Magnética es una técnica médica no invasiva que produce imágenes detalladas del cuerpo humano (MRI). Es una prueba sensible y superior para la detección de enfermedades, ya sean tumores, infecciones, entre otras. Se basa en las propiedades de los tejidos al someter la zona ante un campo magnético fuerte. A diferencia de los rayos X y la tomografía computarizada, la MRI utiliza la radiación en el rango de radiofrecuencia presente en nuestro entorno y es inofensiva para los tejidos [11].

Se basa en la resonancia de los protones de las moléculas de agua, cuya frecuencia de precisión es proporcional al campo magnético luego de aplicar un pulso de radiofrecuencia, una vez que las estructuras involucradas vuelven a su estado de reposo, emiten una señal que se transforma en una escala de grises para construir imágenes. El tiempo que tarda un tejido en regresar al reposo se llama tiempo de relajación, y puede variar, afectando el resultado de la imagen [11].

La resolución de la imagen depende del muestreo y filtro de señales y del ruido. Los tejidos y estructuras tienen distintos niveles de contraste producto de varios factores como la sensibilidad de la MRI a las densidades de protones, la temperatura, el movimiento de los protones, el desplazamiento químico en las frecuencias de Larmor, la heterogeneidad del tejido y los tiempos de relajación T1 y T2 [12].

- T1: Tiempo de relajación longitudinal. Es el tiempo que le toma a los protones de un tejido recuperar su estado normal, y ocurre a lo largo del campo magnético. Cuando se ven las imágenes de RM potenciadas en T1, las sustancias con un T1 largo, como el agua, se aprecian con un color oscuro, mientras que las sustancias con un T1 corto aparecen brillantes, ver Figura 2.4a [13] [14].
- T2: Denominado tiempo de relajación transversal. Es el tiempo que le toma a los protones de un tejido dejar de girar en el mismo plano, recuperando su orientación transversal. En las imágenes potenciadas en T2, los tejidos con T2 largo, como los líquidos en general, lucen más brillantes, mientras que los huesos, que tienen T2 corto, se ven más oscuros, ver Figura 2.4b [13] [14].



Adaptado de [13].

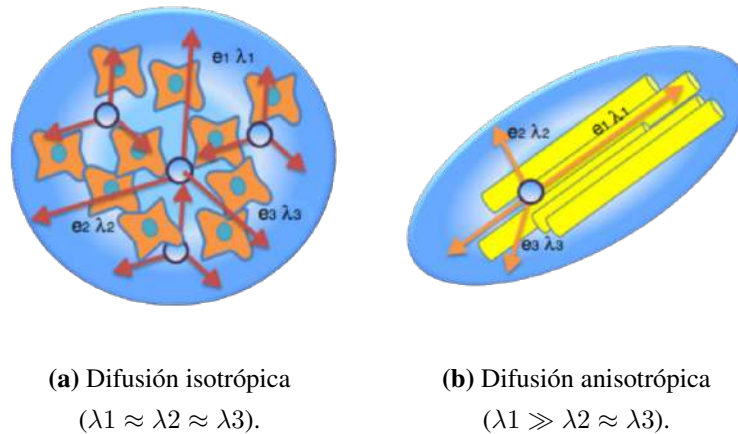
**Fig. 2.4:** Imágenes MRI potenciadas en distintos tiempos de relajación. En (a) el hueso se ve brillante (T1 corto), y los líquidos se ven oscuros (T1 largo). En (b) el hueso se ve oscuro (T2 corto), y los líquidos brillantes (T2 largo).

## 2.5 Resonancia Magnética por Difusión

Las dMRI o DWI son secuencias de imágenes que permiten estudiar la difusión de las moléculas de agua en los tejidos. Esta propiedad se refiere al movimiento caótico y aleatorio de las moléculas de agua, conocido como movimiento browniano. A diferencia de las MRI, incluyen una nueva dimensión en las imágenes, la cual contiene la información del movimiento del agua en diferentes direcciones [15] [16].

Las moléculas intracelulares poseen una difusión restringida, y mientras mayor sea la proporción intracelular, mayor será la restricción y será menos intenso el tejido [15].

La difusión puede ser isotrópica o anisotrópica dependiendo de las direcciones de movimiento de las moléculas en 3 dimensiones. La difusión isotrópica ocurre cuando el movimiento tiene igual probabilidad de dirigirse a cada una de las direcciones del espacio, ver Figura 2.5a, mientras que la anisotropía se presenta cuando existe una dirección dominante, ver Figura 2.5b. Esta última ocurre en el caso de las fibras cerebrales, en las que la dirección dominante es a lo largo del eje mayor de los axones [17].



Adaptado de [17]

**Fig. 2.5:** Tipos de difusividad. Siendo  $\lambda_i$  la magnitud de los vectores propios principales que representan la dirección de la difusión en las moléculas de agua. En (a) las moléculas se mueven en todas las direcciones con igual probabilidad, es decir, con magnitudes similares. En (b), las moléculas se mueven predominantemente en una dirección, en este caso,  $\lambda_1$ .

### 2.5.1 Modelos de difusión local y tractografía

Las dMRI cuentan con diversos algoritmos propios que permiten estudiar la difusión del agua y brindan información sobre las microestructuras. Cada método posee ventajas y limitaciones. Algunos de los más utilizados son Imágenes por Tensor de Difusión (DTI) y modelos a partir de la adquisición de Imágenes de Difusión de Alta Resolución Angular (HARDI), como las Imágenes por Q-Ball (QBI).

La intensidad de una señal de dMRI está determinada por la dirección de las fibras y del gradiente aplicado. Estas direcciones no suelen coincidir, por lo que las propiedades de la difusión se describen matemáticamente mediante modelos de difusión local [5].

### **2.5.1.1 Imágenes por Tensor de Difusión**

El modelo más simple es el Tensor de Difusión (DT). Este es una matriz cuyos valores corresponden a una orientación del gradiente y a una orientación de la célula [14]. Se utiliza en las DTI para calcular los valores propios y las direcciones principales de difusión en la sustancia blanca [16] [18].

En comparación con una imagen de difusión convencional, esta técnica nos brinda información más detallada sobre la estructura y orientación de las fibras nerviosas. Cuantifica la anisotropía de difusión y proporciona una estimación de la dirección principal de las fibras [5] [17].

Aun así, DTI cuenta con una limitación importante. Debido a que el modelo de tensor solo puede resolver una dirección de las fibras dentro de cada vóxel, ver Figura 2.6, no puede resolver el cruce, flexión o torsión de las fibras dentro de esta unidad. Por ende, lo vuelve inadecuado para representar la arquitectura en regiones con patrones complejos. Sin embargo, hoy en día se sabe que esto se puede resolver con las técnicas de Imágenes de Difusión de Alta Resolución Angular, como el modelo de Imágenes Q-Ball, que resuelve mejor los cruces de fibras [17] [19].

### **2.5.1.2 Imágenes de Difusión de Alta Resolución Angular**

Inicialmente, es importante tener en cuenta que HARDI no es en sí una técnica para estimar o representar las orientaciones de las fibras, sino que se trata de una estrategia de adquisición. Si bien es posible utilizar datos obtenidos con HARDI para realizar una reconstrucción DTI, esta técnica es más utilizada para métodos que resuelven fibras que se cruzan o curvan [18].

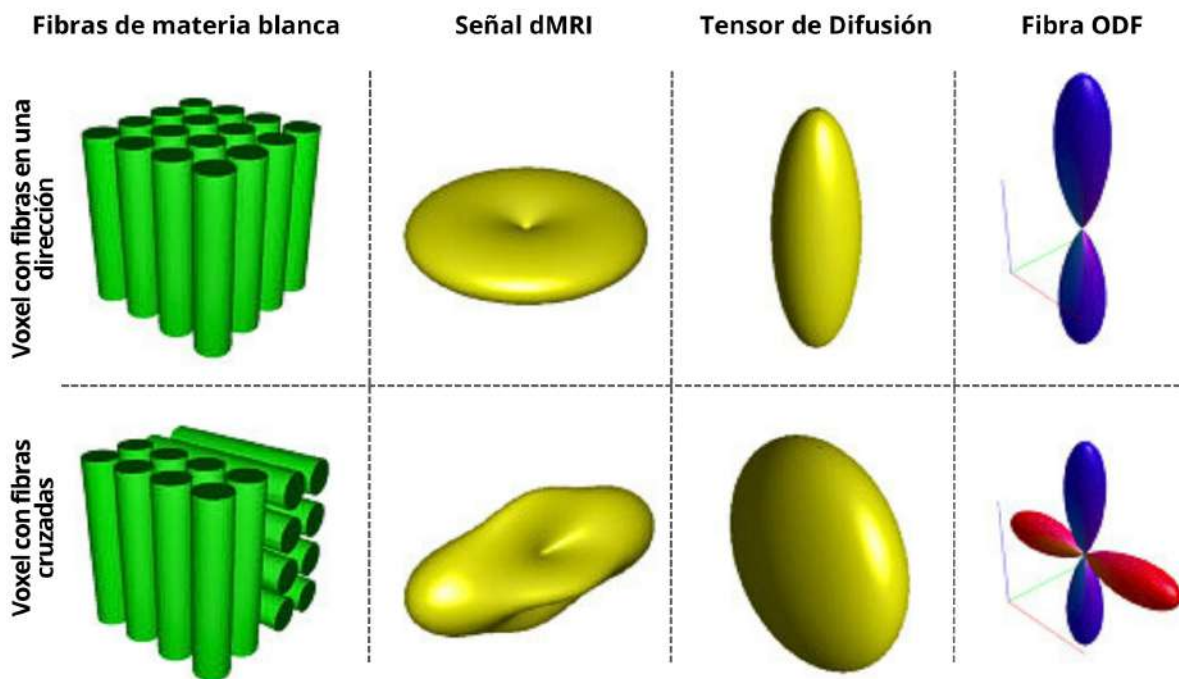
Los parámetros que caracterizan esta técnica son dos, el número de direcciones únicas y el valor  $b$ , el cual se utiliza para cuantificar la difusión del agua en el tejido. HARDI suele usar un mayor número de direcciones de gradiente dependiendo del método utilizado. Cabe mencionar que al aumentar el número de direcciones, mejora la precisión angular alcanzable [5] [18].

Existen diferentes variantes de HARDI que permiten modelar con mayor precisión la arquitectura de las fibras, aunque algunas son más precisas que otras. Sin embargo, es importante tener en cuenta que este método también presenta algunas limitaciones, por ejemplo, su sensibilidad al ruido [5] [20].

### 2.5.1.3 Imágenes por Q-Ball

También conocida como QBI, es un modelo basado en la adquisición HARDI, y a diferencia de la DTI, no describe la difusión por medio de un tensor, sino que lo hace mediante la Función de Distribución de Orientación (ODF). Esta función se calcula a partir de la señal medida a través de diferentes gradientes, y permite identificar múltiples direcciones de fibras en un vóxel, ver Figura 2.6, proporcionando información sobre su orientación y distribución [5] [19].

Q-Ball es una buena herramienta para detectar regiones cerebrales con patrones complejos de cruce de fibras, por lo que es bastante utilizada en estudios de conectividad y tractografías [5]. A diferencia de DTI, requiere un mayor tiempo de adquisición y procesamiento de datos, y en cuanto a otros modelos de HARDI, en regiones con curvaturas de fibras en el interior de un vóxel, QBI podría ser incapaz de distinguir adecuadamente entre ellas [20] [19].



Adaptado de [21]

**Fig. 2.6:** Comparación entre Tensor de Difusión (DT) y Función de Distribución de Orientación (ODF). Mientras que el DT asume una única dirección de difusión en cada voxel, la ODF permite la detección de múltiples direcciones de difusión y una visualización más detallada de la estructura de la materia blanca del cerebro.

### **2.5.1.4 Tractografía Cerebral**

Herramienta fundamental para comprender la estructura y conectividad del cerebro. Da al investigador oportunidades de comparar y contrastar los resultados entre diferentes individuos, regiones cerebrales y con otras técnicas [5]. Es la representación tridimensional de las fibras del cerebro, a través de la determinación de la dirección dominante de movimiento del agua en cada vóxel generada por modelos como DTI, Q-Ball, entre otras [17] [22]. Es valiosa en la evaluación y el tratamiento de una amplia variedad de condiciones neurológicas y psiquiátricas. Pues muchas enfermedades y patologías afectan a la sustancia blanca, como la Esclerosis Múltiple, demencias, encefalitis, TEA, entre otros [22].

## **2.5.2 Medidas de difusividad y anisotropía**

A partir de las técnicas de dMRI, se pueden calcular varios parámetros, como las medidas de difusión, que corresponden a Difusividad Media (MD), Difusividad Axial (AD), Difusividad Radial (RD) y Coeficiente de Difusión Aparente (ADC), y medidas de anisotropía, como Anisotropía Fraccional (FA) y una modificación de esta derivada del modelo Q-Ball, llamada Anisotropía Fraccional Generalizada (GFA). En esta ocasión se enfatiza en lo siguiente.

### **2.5.2.1 Coeficiente de Difusividad Aparente**

ADC se utiliza para estimar el grado de restricción de la difusión del agua en los tejidos. Es una medición cuantitativa absoluta del movimiento traslacional del agua. Las estructuras con un alto índice de difusión presentan un alto valor de ADC, y en consecuencia, brillarán más. Por otro lado, las estructuras con difusión restringida como WM y GM aparecen hipointensas, es decir, oscuras, ver Figura 2.7a. Esta medida nos permite corroborar que una visible hiperintensidad es debida netamente a la restricción y no a otro factor [16] [17].

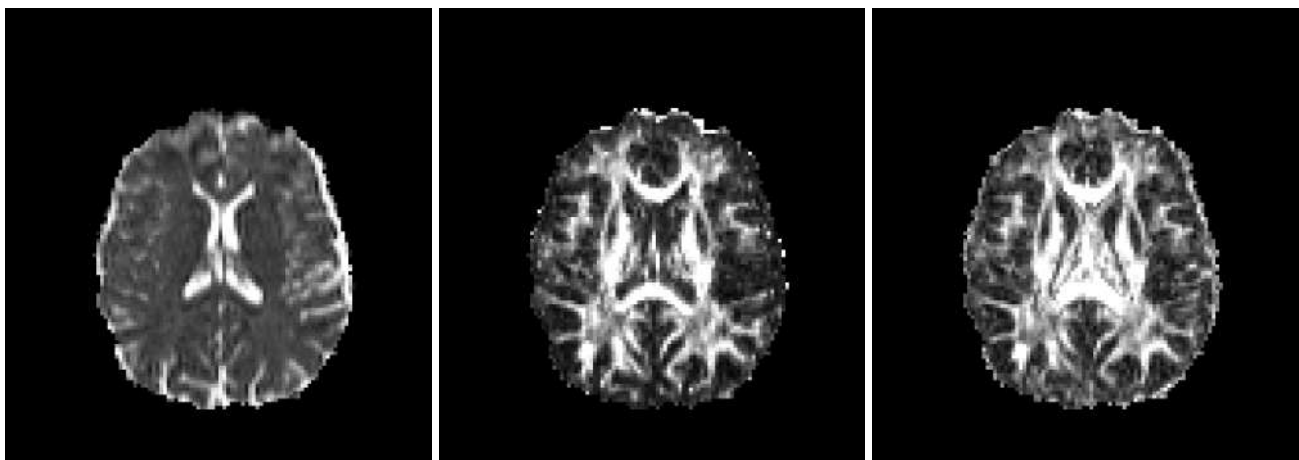
### **2.5.2.2 Anisotropía Fraccional**

Es la relación entre el componente anisotrópico del tensor de difusión (DT) y el tensor de difusión completo. Entrega información sobre la orientación y coherencia de la difusión de agua en los tejidos, y su grado de anisotropía. Dicho valor varía entre 0 (difusión isotrópica máxima) y 1 (difusión anisotrópica máxima). Las estructuras con alto valor FA son brillantes, y con bajo valor, oscuras, ver Figura 2.7b [16].

Este parámetro es una herramienta útil para la identificación temprana de patologías cerebrales y para la monitorización de enfermedades [18]. Cabe mencionar, que el análisis de las medidas de ADC y FA permiten especialmente estudiar la maduración cerebral normal y anormal en los niños. Ya que, respectivamente, los valores de ADC disminuyen con la maduración y los valores de FA aumentan [16].

### 2.5.2.3 Anisotropía Fraccional Generalizada

Está asociado con la actividad neuronal regional y en la conexión cerebral en reposo, especialmente en el sistema límbico, es decir, en el grupo de estructuras que dirigen las emociones, aprendizaje y memoria [23] [24]. En otras palabras, el nivel de GFA puede reflejar la forma en que dichas áreas se comunican con otras partes del cerebro en estado de reposo. Ver Figura 2.7c.



(a) Coeficiente de Difusividad Aparente

(b) Anisotropía Fraccional

(c) Anisotropía Fraccional Generalizada

**Fig. 2.7:** Medidas de difusividad y anisotropía del sujeto as100351 de la base de datos HARDI de adolescentes con autismo, capturadas en el corte 26 del plano axial.

### 2.5.3 Sistemas de coordenadas

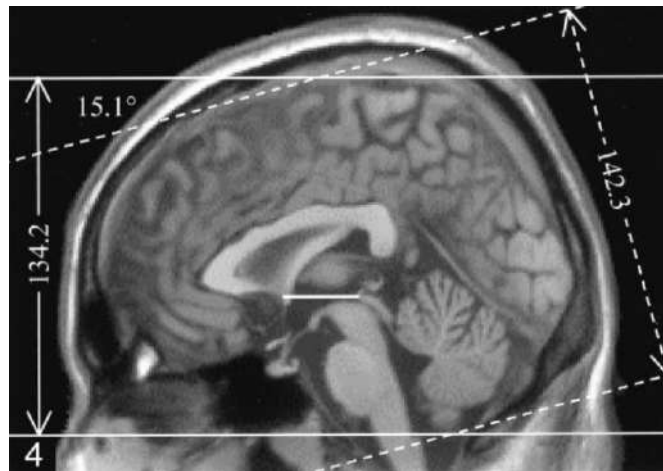
Para tener una visión general y no individual del cerebro, a menudo se combinan imágenes de muchos sujetos para formar mapas de consenso del cerebro [25]. Por ejemplo, para llevar a cabo procedimientos y estudios utilizando un conjunto de tractografías, es esencial que todas estén en un sistema de coordenadas común. Si bien no garantiza que una subestructura se encuentre en el mismo punto para cada sujeto, este enfoque permite automatizar el trabajo, ahorrando una cantidad significativa de tiempo. Los dos espacios más utilizados en la comunidad de neurociencia son el espacio Talairach y el espacio del Instituto Neurológico de Montreal (MNI) [26].



### 2.5.3.1 Talairach

La primera versión del atlas Talairach se publicó en 1967 y constaba en un atlas de varios cerebros de distintos sujetos, por ende, no era ideal para la normalización espacial. En 1988 se creó una segunda versión basada en el cerebro de una sola persona, que sí sirvió para la normalización espacial, y fue posteriormente modificada y mejorada en las versiones de 1993 y 1998 [25].

Su orientación está definida considerando la Comisura Anterior (AC), conectando por una línea en el eje y, desde el punto más superior de AC y el más inferior de la Comisura Posterior (PC), formando el modelo estándar AC-PC [26] [25]. Cabe mencionar que, las plantillas de Talairach y MNI no coinciden perfectamente, debido a las diferencias en forma y tamaño del cerebro, ver Figura 2.8, [26] [27].



Adaptado de [28]

**Fig. 2.8:** Se aprecia la vista de una MRI en el Espacio MNI y Talairach. Las líneas continuas representan la percepción en Talairach, y las líneas entrecortadas, la vista en MNI.

## 2.6 Preprocesamientos de las dMRI

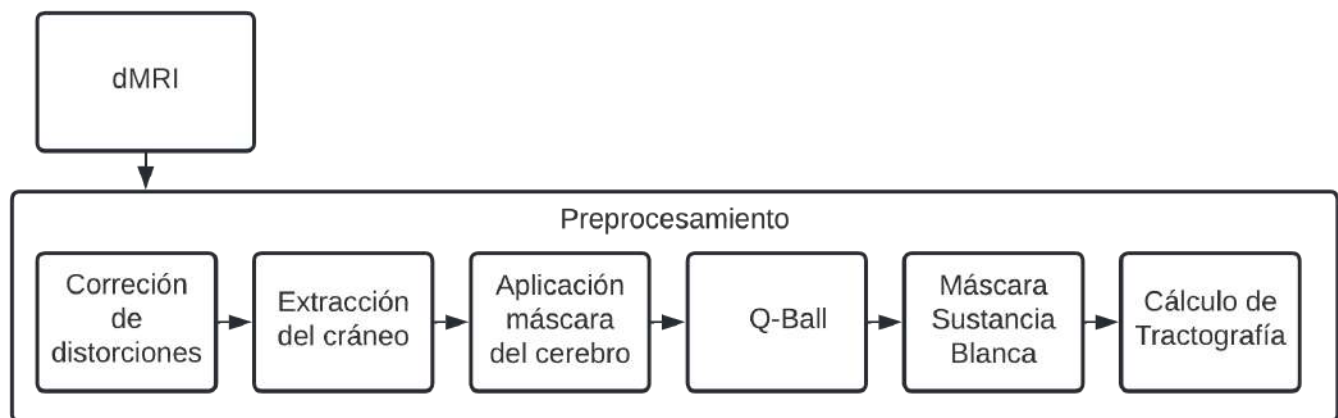
Constan de técnicas y operaciones que se les aplican a los datos antes de ser procesados para su análisis. Son fundamentales, ya que permiten asegurar la calidad e integridad de estos, corrigiendo posibles errores o inconsistencias. Su ejecución permite obtener resultados precisos y confiables.

En muchos casos, estas operaciones se llevan a cabo en un esquema secuencial conocido como Pipeline. En este, los datos se someten a una serie de etapas de procesamiento, donde cada una transforma los datos de una manera específica y los pasa a la siguiente [29].

## 2.6.1 Operaciones

Usualmente, el preprocesamiento de datos dMRI consta de un pipeline con scripts de extracción de elementos irrelevantes, limpieza de ruido, aplicación de técnicas propias, entre otras. Ver Figura 2.9.

La corrección de movimientos y distorsiones se enfoca en abordar eventos que hayan ocurrido durante la adquisición de datos que puedan haber afectado la calidad de la información. Por otro lado, la extracción del cerebro a través de una máscara y la eliminación previa de estructuras no relacionadas con él, permiten aplicar de manera adecuada los algoritmos para caracterizar la arquitectura de la materia blanca y, de esta manera, obtener tractografías precisas.



**Fig. 2.9:** Diagrama que representa el pipeline de preprocesamiento para las imágenes dMRI, pasando por la corrección de imágenes, hasta la obtención de la tractografía.

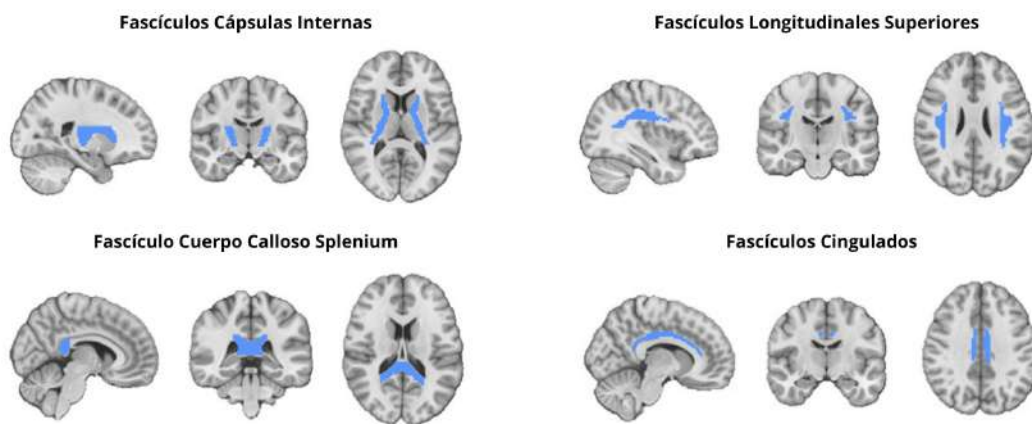
## 2.7 Estudio de WM en pacientes con autismo

En [30], se evalúa la materia blanca mediante los valores de Anisotropía Fraccional (FA) en una población de 194 niños, 125 pacientes con TEA y 69 controles, cuyas edades varían entre los 2.5 y 7 años. De lo cual se obtiene que los niños autistas mostraban un valor más bajo de FA en los fascículos Cingulados, Logintudinales Superiores, Cápsula Interna y Cuerpo Calloso Splenium, ver Figura 2.10.

Por otro lado, en [31] realizan estudios de WM en una población homogénea de 58 hombres, 27 pacientes con TEA y 31 controles, del cual se obtiene una correlación entre las ADOS, que en español se traduce como Programa de Observación para el Diagnóstico del Autismo, que consta de una serie de actividades que permiten observar y evaluar el comportamiento social, comunicativo e interacción social de una persona, y la conectividad temporal, frontal y parietal, ver Figura 2.11, lo anterior por medio del

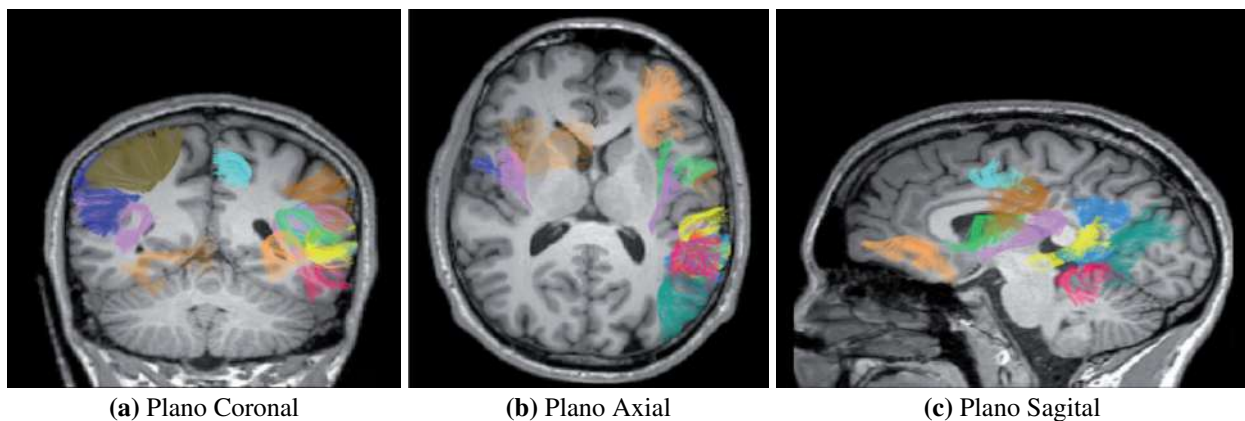
análisis de valores promedio de la Anisotropía Fraccional Generalizada (GFA). Obtuvieron atipicidades específicas en el lóbulo temporal, ver Figura 2.12a y la ínsula, ver Figura 2.12b, cuyo resultado se asoció significativamente con manifestaciones clínicas del trastorno del espectro autista, como conciencia social, estructura del lenguaje, habilidades pragmáticas y empatía.

Mientras que en [32] se evaluó en un grupo de 110 sujetos, 53 pacientes con TEA y 57 controles, de edades que variaban entre los 6 y 25 años, mediante los valores de FA y MD en dos grupos separados, los niños menores de 12.5 años y los adolescentes/adultos mayores o iguales a 12.5 años. Se encontró anomalías de la Materia Blanca Superficial en las cortezas parietales mediales bilaterales que abarcan el precuneus y la circunvolución posterior del cíngulo y también en la unión temporoparietal derecha.



Adaptado de [30]

**Fig. 2.10:** Fascículos con anomalías en WM, según estudios de la literatura sobre los valores FA [30].



Adaptado de [31]

**Fig. 2.11:** Fascículos pertenecientes a la conectividad temporal, frontal y parietal, correlacionados, según [31], al comportamiento de sujetos autistas.



(a) Lóbulo Temporal.  
Adaptado de [33]

(b) Lóbulo de la Ínsula.  
Adaptado de [34]

**Fig. 2.12:** Vista lóbulos Temporal e Ínsula desde el plano sagital.

## 2.8 Multiplicación matricial

Para multiplicar matrices es necesario que las dimensiones de estas sean compatibles. Es decir, una matriz  $A$  de tamaño  $n \times m$  puede ser multiplicada por cualquier matriz  $B$  de tamaño  $m \times p$ , siempre y cuando la cantidad de columnas de la primera matriz, sea equivalente a la cantidad de filas de la segunda matriz. Como resultado, se obtiene una nueva matriz  $C$  de tamaño  $n \times p$ , ecuación (2.8.1) [35].

$$A_{n \times m} \cdot B_{m \times p} = C_{n \times p} \quad (2.8.1)$$

Siendo  $A$ ,  $B$  y  $C$  matrices compatibles para multiplicar e  $In$  la matriz identidad, se destacan las siguientes propiedades y consideraciones [35]:

$$\text{Ley asociativa: } A(BC) = (AB)C$$

$$\text{Ley distributiva: } A(B+C) = AB + AC$$

$$\text{Propiedad de invertida: } A \cdot A^{-1} = In$$

$$A(In) = In(A)$$

$$AB \neq BA$$

## 2.8.1 Transformaciones 3D

Existen diferentes matrices de transformación para aplicar a elementos en un espacio tridimensional. Estas transformaciones pueden ser aplicadas de manera individual o combinada, y su representación comúnmente se realiza mediante matrices de tamaño  $4 \times 4$ . Esto se debe a que ciertas transformaciones, como la traslación, requieren de una columna o fila adicional para su correcta aplicación. En consecuencia, las coordenadas tridimensionales se representan como matrices vectoriales  $4 \times 1$  con un componente extra, cuyo valor es equivalente a 1 [36].

### 2.8.1.1 Matriz de traslación 3D

Esta matriz modifica la posición en el espacio. La última columna almacena la cantidad de traslación en cada componente por medio de  $\mathbf{T}_x$ ,  $\mathbf{T}_y$  y  $\mathbf{T}_z$ , como se observa en la ecuación (2.8.2) [36].

$$\begin{bmatrix} 1 & 0 & 0 & \mathbf{T}_x \\ 0 & 1 & 0 & \mathbf{T}_y \\ 0 & 0 & 1 & \mathbf{T}_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \equiv \begin{bmatrix} x + \mathbf{T}_x \\ y + \mathbf{T}_y \\ z + \mathbf{T}_z \\ 1 \end{bmatrix} \quad (2.8.2)$$

### 2.8.1.2 Matriz de escalado

Esta matriz modifica las dimensiones en cada eje del espacio por medio de los valores  $\mathbf{E}_x$ ,  $\mathbf{E}_y$  y  $\mathbf{E}_z$ , que multiplican a su respectiva coordenada asociada, como se observa en la ecuación (2.8.3) [36].

$$\begin{bmatrix} \mathbf{E}_x & 0 & 0 & 0 \\ 0 & \mathbf{E}_y & 0 & 0 \\ 0 & 0 & \mathbf{E}_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \mathbf{T}_x(x) \\ \mathbf{T}_y(y) \\ \mathbf{T}_z(z) \\ 1 \end{bmatrix} \quad (2.8.3)$$

### 2.8.1.3 Matriz de rotación

Modifica la rotación 3D bloqueando un solo eje. Por ende, existen 3 variantes para realizar esta transformación. Considerando un ángulo  $\alpha$ , la rotación en el eje 'x' está dada por la ecuación (2.8.4), la rotación en el eje 'y', por la ecuación (2.8.5), y al rededor del eje 'z', por la ecuación (2.8.6) [35] [36].

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & -\text{sen}(\alpha) & 0 \\ 0 & \text{sen}(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8.4)$$

$$\begin{bmatrix} \cos(\alpha) & 0 & \text{sen}(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ -\text{sen}(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8.5)$$

$$\begin{bmatrix} \cos(\alpha) & -\text{sen}(\alpha) & 0 & 0 \\ \text{sen}(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8.6)$$

## 2.9 Estadística

### 2.9.1 Estadística descriptiva

Se enfoca en describir y resumir los datos mediante diversas medidas, tales como la media, mediana, desviación estándar, entre otras. Proporciona una visión general de los datos y su distribución [37].

#### 2.9.1.1 Distribución

Es la forma en que se distribuyen los valores de una variable en un conjunto de datos. Describe la frecuencia o probabilidad de ocurrencia de cada valor o rango de valores de la variable [37]. A continuación se presentan los comúnmente utilizados para analizar el comportamiento de la distribución.

- Prueba de Shapiro-Wilk: Esta prueba se utiliza para contrastar la normalidad cuando el tamaño de la muestra es menor a 50. Usualmente, se considera un valor de significancia de referencia de 0.05. Los "p-values" mayores a este valor indican una distribución normal, mientras que los "p-values" menores indican lo contrario [38].
- Prueba de simetría: También conocida como Prueba de Skewness, evalúa la igualdad o desigualdad de los extremos de una distribución con relación a su media. Al realizar esta prueba, se obtiene una medida numérica que indica si los datos son simétricos, con un valor cercano a cero, o si presenta una asimetría hacia la izquierda o derecha [38].
- Prueba de curtosis: Evalúa la concentración de los datos alrededor de la media y la forma de los extremos de la distribución. Proporciona una medida numérica que indica si la distribución de los datos se asemeja a una distribución normal, es decir, curtosis cercana a cero [38].

### 2.9.1.2 Varianza

Medida de dispersión que indica la variabilidad de un conjunto de datos con respecto a su media. Si la varianza es alta, significa que los datos están más dispersos y alejados, en el caso contrario, indica que los datos están más cercanos y agrupados [37] [39].

- Prueba de Levene: Prueba estadística utilizada para evaluar si las varianzas de dos o más grupos son iguales o diferentes. Su objetivo es determinar si hay evidencia suficiente para afirmar o rechazar la hipótesis calculando una estadística de prueba basada en las diferencias absolutas entre los valores individuales y la media de cada grupo [39].
- Prueba de Bartlett: También se utiliza para evaluar si las varianzas de dos o más grupos son iguales o diferentes. Su diferencia radica en el cálculo asociado, Bartlett se basa en la comparación de las sumas de cuadrados de las diferencias entre los valores individuales y la media de cada grupo [39].

### 2.9.2 Estadística inferencial

Se refiere a la técnica utilizada para realizar inferencias y tomar decisiones basadas en los resultados obtenidos a partir de técnicas descriptivas. Su objetivo es obtener conclusiones en base a la información recopilada [37]. Mediante esta técnica se puede realizar una comparación de dos muestras para determinar si existen diferencias significativas entre ellas. Existen pruebas paramétricas como no paramétricas. La elección entre una u otra depende de las características de los datos y los supuestos asociados [40].

### 2.9.2.1 T-test

Prueba paramétrica que requiere que las muestras sean independientes, estén compuestas por valores continuos, sigan una distribución normal y tengan varianzas homogéneas entre ellas. Al realizar el t-test, se calcula una medida llamada "valor t" que compara la diferencia entre las medias de los dos grupos con la variabilidad dentro de cada grupo. Luego, se establece un nivel de significancia y se compara con el "p-value" obtenido para determinar si se rechaza o se acepta la hipótesis planteada [41].

### 2.9.2.2 Prueba de Mann Whitney U

Esta es una prueba no paramétrica. A diferencia del t-test, no requiere que las muestras independientes sigan una distribución normal. Si bien proporciona una solución cuando los supuestos no se cumplen, no es tan precisa como el t-test. Esta prueba se basa en la clasificación de los datos en rangos y compara las sumas de los rangos entre los dos grupos. El resultado de la prueba es un valor llamado U, que se utiliza para determinar si hay una diferencia significativa entre los grupos [42].

## 2.10 Machine Learning

Conocido en español como Aprendizaje Automático, es la ciencia de crear algoritmos que le permitan al computador aprender a partir de datos sin ser explícitamente programados para ello, mejorando el rendimiento de una tarea a través de la experiencia [43]. Dichos algoritmos se pueden clasificar según el tipo y cantidad de supervisión que reciben durante el entrenamiento:

- Aprendizaje supervisado: El algoritmo utiliza un conjunto de entrenamiento que incluye las soluciones deseadas, denominadas etiquetas.
- Aprendizaje no supervisado: Los datos de entrenamiento no están etiquetados, es decir, el algoritmo intenta aprender solo. Busca encontrar patrones o estructuras en los datos de entrada.
- Aprendizaje semisupervisado: Algunos algoritmos pueden trabajar con datos parcialmente etiquetados, ya que etiquetar los datos puede llevar demasiado tiempo y/o ser costoso.
- Aprendizaje por refuerzo: El sistema de aprendizaje puede observar el entorno, seleccionar y realizar acciones para obtener recompensas o penalizaciones a cambio. De esta forma, debe aprender por sí mismo cuál es la mejor estrategia para obtener la mayor recompensa.



## 2.10.1 Algoritmos de Aprendizaje Supervisado

Existen diversos algoritmos de aprendizaje supervisado, cada uno con sus propias características y aplicaciones. A continuación se exploran dos modelos ampliamente utilizados para clasificación.

### 2.10.1.1 Árboles de Decisión

Pueden realizar tareas de clasificación, regresión e incluso tareas de múltiples salidas. Son algoritmos potentes, capaces de ajustarse a conjuntos de datos complejos. Este modelo funciona dividiendo repetidamente el conjunto de datos en función de las características relevantes, creando así, una estructura de árbol que permite realizar predicciones [43]. Para construir un modelo robusto y preciso se pueden variar los hiperparámetros utilizados. Cuatro de ellos son:

- **Profundidad máxima:** Controla el número de niveles o divisiones que puede tener el árbol. Un valor alto de profundidad implica un árbol más complejo con más divisiones. Por otro lado, un valor muy bajo limita la complejidad del árbol y puede llevar a un ajuste deficiente.
- **Mínimo de muestras por hoja:** Este parámetro establece el número mínimo de muestras requeridas en una hoja del árbol. Un valor alto resultará en hojas más generales, mientras que un valor bajo, generará hojas más específicas.
- **Mínimo de muestras para dividir:** Número mínimo de muestras requeridas para dividir un nodo interno del árbol. Un valor alto resultará en un árbol con menos divisiones, es decir, más generalizado. Un valor bajo generará un árbol con más divisiones, lo que lo hará más específico.
- **Máximas características:** Número máximo de características que se considerarán al buscar la mejor división en cada nodo del árbol. Un valor alto puede permitir más opciones de división, mientras que un valor bajo limita las opciones de división y puede llevar a un árbol más simple.

### 2.10.1.2 Máquina de Vectores de Soporte

También denominado SVM por sus siglas en inglés, es un modelo versátil utilizado para clasificación, regresión y detección de valores atípicos. Es especialmente adecuado para conjuntos de datos complejos de tamaño pequeño o mediano. Busca separar las clases en el espacio de características [43]. Para ello considera varios hiperparámetros. Tres de los más utilizados e importantes son:

- **C:** Controla el equilibrio entre conseguir un bajo error de entrenamiento y de prueba. Un alto valor de C implica un margen más estrecho, lo que puede resultar en una clasificación más precisa, pero un valor extremadamente alto de C puede llevar al sobreajuste del modelo. Por otro lado, un C bajo implica un margen amplio, lo cual podría implicar una clasificación menos precisa.
- **Kernel:** Función utilizada para transformar los datos de entrada en un espacio de mayor dimensión, lo que facilita su separación. Algunas opciones de kernel son *polinómico*, *base radial*, y *lineal*.
- **Gamma:** Controla el alcance de influencia de un solo ejemplo de entrenamiento. Un valor bajo de gamma indica un alcance amplio. Por otro lado, un valor de gamma alto indica un alcance estrecho, es decir, que la influencia de cada ejemplo de entrenamiento se limita a sus vecinos cercanos.

### 2.10.1.3 Validación cruzada

Técnica utilizada para evaluar el rendimiento de un modelo. Los datos se dividen en conjuntos de entrenamiento y validación, iterando a través de ellos, permitiendo evaluar el modelo en diferentes combinaciones de datos, evitando el sesgo y obteniendo una estimación más confiable. Una de sus aplicaciones es la selección de los mejores hiperparámetros de un modelo a través de la búsqueda en rejilla [43].

### 2.10.1.4 Métricas de evaluación

Elementos clave para evaluar la efectividad y rendimiento de los modelos de ML. Brindan una comprensión más completa de cómo clasifica y generaliza los datos. Las métricas más utilizadas son:

- **Accuracy:** Muestras clasificadas correctamente en relación con el número total de muestras. Un valor cercano a 1 indica una mayor precisión en las clasificaciones [43]. Ecuación (2.10.1).
- **Precision:** Muestras positivas correctamente predichas en relación con el número total de muestras clasificadas como positivas. Un valor cercano a 1 indica que la mayoría de las muestras clasificadas como positivas son realmente positivas [43]. Ecuación (2.10.2).
- **Recall:** Muestras positivas correctamente predichas en relación con el número total de muestras positivas reales. Valor  $\approx 1$ , indica mayor acierto en las predicciones positivas [43]. Ecuación (2.10.3).
- **F1 Score:** Combina precisión y recall en un valor, e indica el equilibrio entre ambos. Varía entre 0 y 1, cuán mayor sea, indica un mejor rendimiento [43]. Ecuación (2.10.4).

Verdaderos Positivos (VP), son las muestras positivas que han sido correctamente clasificadas. Verdaderos Negativos (VN), muestras negativas que han sido correctamente clasificadas. Falsos Positivos (FP) son las muestras negativas que han sido incorrectamente clasificadas como positivas. Falsos Negativos (FN), son la cantidad de muestras positivas que han sido incorrectamente clasificadas como negativas.

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.10.1)$$

$$Precision = \frac{VP}{VP + FP} \quad (2.10.2)$$

$$Recall = \frac{VP}{VP + FN} \quad (2.10.3)$$

$$F1Score = 2 \times \left( \frac{Precision \times Recall}{Precision + Recall} \right) \quad (2.10.4)$$

## 2.11 Discusión

Las imágenes de resonancia magnética son de suma importancia en el ámbito clínico, mientras que las imágenes de resonancia magnética por difusión (dMRI) juegan un papel fundamental en la investigación. Se pueden utilizar para conocer más de la conectividad de la materia blanca, estudiar patologías y/o detectar anomalías microestructurales en el cerebro, sobre todo con sus respectivas medidas ADC, FA, GFA, entre otras. En este sentido, técnicas como la Imagen de Tensor de Difusión (DTI) y la Imagen de Q-Ball (QBI) son ampliamente utilizadas. Aunque la DTI es más común debido a su simplicidad y rapidez, presenta limitaciones en la representación de cruces y curvaturas de las fibras en comparación con modelos basados en la adquisición de imágenes de difusión de alta resolución (HARDI), como la técnica Q-Ball. Esta última tiene la capacidad de representar de manera más precisa la arquitectura de la materia blanca.

Además, ML se presenta como una herramienta poderosa para clasificar y descubrir relaciones complejas en conjuntos de datos. Su capacidad para analizar grandes volúmenes de información y encontrar patrones que podrían pasar desapercibidos en análisis estadísticos tradicionales es especialmente valiosa. La elección del modelo de aprendizaje automático más apropiado dependerá de la naturaleza del problema, la cantidad y calidad de los datos.

## Capítulo 3. Definición del pipeline de procesamiento

---

### 3.1 Introducción

Se describe el diseño llevado a cabo para cumplir con los objetivos de la investigación. Se realizó un análisis exhaustivo de la base de datos HARDI de adolescentes con autismo, y además se evaluaron diferentes software, scripts, herramientas y bibliotecas necesarias para la implementación del pipeline.

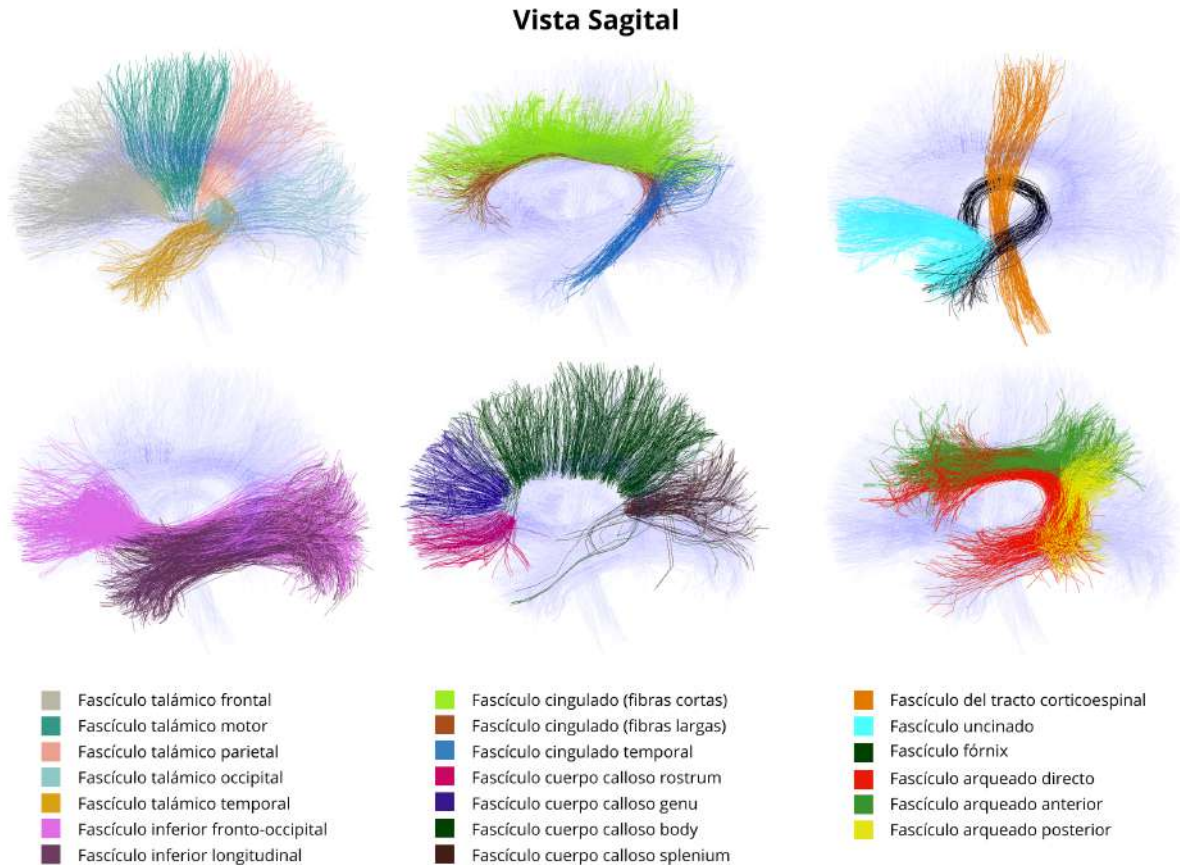
### 3.2 Base de datos

Base de datos HARDI de adolescentes con autismo, provista por colaboradores del Centro Neurospin, Francia, y facilitada por la profesora guía Pamela Guevara, cuenta con datos de 37 sujetos, 18 controles y 19 pacientes [6]. Para esta investigación se han utilizado los siguientes archivos:

- Listado de sujetos: Almacenado en un archivo de texto, posee la lista anonimizada de los 37 sujetos, diferenciando entre sí los 19 pacientes con autismo y los 18 controles de sujetos sin autismo.
- Tractografías: Fueron obtenidas basándose en el modelo Q-Ball. Su información se encuentra en archivos bundles y bundlesdata. Las imágenes dMRI presentan un corte en su sección inferior proveniente de una modificación que recibieron el año 2011 para otros procesos.
- Matrices de transformadas: Son dos matrices de 4x3 resultantes de una combinación de transformaciones, tales como escalado 3D, traslado y rotación. Una transforma la tractografía al espacio estructural T1, y la otra lo hace posteriormente al espacio Talairach.
- Imágenes FA, GFA y ADC: Medidas de Anisotropía Fraccional, Anisotropía Fraccional Generalizada y Coeficiente de Difusividad Aparente. La información de cada una está contenida en 3 archivos del formato GIS propio del Software BrainVisa, cuyas extensiones son .ima, que contiene la imagen, .dim, su dimensión y geometría, y .minf, la información de metadatos adicionales [44].

### 3.3 Atlas de segmentación

Atlas en espacio Talairach de 36 fascículos bundles basados en la técnica de adquisición HARDI. Compuesto por tractos interhemisféricos, arqueado, cuerpo caloso, uncinado, cíngulo, radiaciones talámicas, longitudinal inferior, fronto-occipital inferior, fórnix y tracto corticoespinal, ver Figura 3.13 [7]



**Fig. 3.13:** Atlas de 36 fascículos agrupados y vistos desde el plano Sagital, cada color representa un fascículo diferente. Los fascículos, a excepción de los del cuerpo calloso, se encuentran en ambos hemisferios. Para apreciar su ubicación, en la figura solo se muestran los fascículos del lado izquierdo.

### 3.4 BrainVisa

Software de neuroimagen de código abierto que se utiliza para la visualización, procesamiento y análisis de datos de imágenes médicas del cerebro. Permite la manipulación de imágenes estructurales y funcionales, proporciona una amplia gama de herramientas para la medición, segmentación y registro de imágenes cerebrales [45]. Se destacan:

- Aims Converter: Permite convertir archivos de imágenes médicas de un formato a otro de manera automática y en cuestión de segundos.
- Anatomist: Permite visualizar imágenes médicas y examinar su información. Además, puede integrarse con otros módulos de BrainVisa para realizar análisis avanzados.

## 3.5 Códigos existentes

Se presentan los códigos suministrados por el grupo de análisis de imágenes médicas de la Universidad de Concepción que serán empleados. Estos scripts están escritos en lenguaje Python o C, dos lenguajes de programación ampliamente utilizados en una variedad de aplicaciones, incluyendo el desarrollo de sistemas operativos, programación de microcontroladores, compiladores, aplicaciones de alto rendimiento, análisis de datos, automatización de tareas, entre otros.

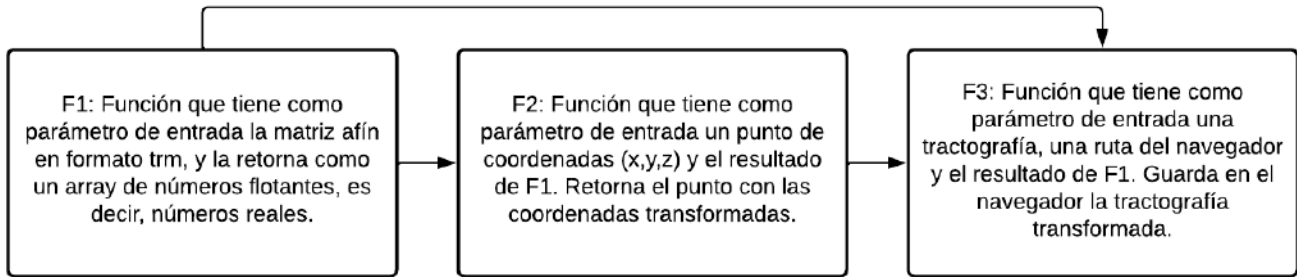
### 3.5.1 Manipulación de archivos bundles

Código escrito en lenguaje Python y consta de 8 funciones. Fue creado para trabajar con archivos de tractografía. Entre las funciones más básicas y ampliamente utilizadas se encuentran la escritura y lectura de archivos bundles, los cuales contienen coordenadas e información de reconstrucción de la tractografía.

### 3.5.2 Aplicación de transformación espacial

Pipeline escrito en lenguaje Python. Consiste en tres funciones interconectadas, donde la salida de una función se convierte en la entrada de la siguiente. Este pipeline requiere un archivo bundle y una matriz afín como entradas. A grandes rasgos, la matriz afín está compuesta por componentes de traslación, escalado y rotación. Al multiplicar esta matriz por cada punto de un elemento tridimensional, se generan cambios en su posición y orientación [46].

Para comprender en profundidad la metodología que emplea, ver Figura 3.14. En esta, la función F1 se encarga de inicializar la matriz afín  $4 \times 3$  denominada 'MA' en el código, dicha matriz se transforma y se extiende en una matriz cuadrada  $4 \times 4$ , como se muestra en la ecuación (3.5.1). F2 recibe como parámetros de entrada el resultado de F1 y, en particular, las coordenadas  $(x,y,z)$  de un punto. Estas coordenadas también se someten a una transformación al agregar un elemento final equivalente a 1, convirtiéndolas posteriormente en una matriz  $4 \times 1$  denominada 'P', como se muestra en la ecuación (3.5.2). Por último, la función F3 se encarga de inicializar la tractografía que se desea transformar espacialmente, y utiliza las funciones F1 y F2 para multiplicar la matriz afín por cada punto 'P' del archivo, según se aprecia en la ecuación (3.5.3). Como resultado, la matriz resultante se vuelve a convertir al formato de coordenadas, como se observa en la ecuación (3.5.4).



**Fig. 3.14:** Pipeline para aplicar una transformación espacial. F1, F2 y F3 son las 3 funciones pertenecientes al script de Aplicación de transformada.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix} \rightarrow \begin{bmatrix} d & e & f & a \\ g & h & i & b \\ j & k & l & c \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv MA \quad (3.5.1)$$

$$(x \ y \ z) \rightarrow [x \ y \ z \ \mathbf{1}] \rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \equiv P \quad (3.5.2)$$

$$MA \times P \Leftrightarrow \begin{bmatrix} d & e & f & a \\ g & h & i & b \\ j & k & l & c \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \equiv \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \quad (3.5.3)$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \rightarrow (x' \ y' \ z') \quad (3.5.4)$$

### 3.5.3 Segmentación de fascículos

Proceso esencial en el análisis de imágenes médicas, especialmente en el campo de la neuroimagen. Su objetivo principal es identificar y separar estructuras de interés en el cerebro humano, representado por datos de tractografías, en segmentos conocidos como fascículos. Esta técnica permite un estudio individualizado de cada fascículo y el análisis de sus métricas representativas. A continuación, se mencionan dos métodos, para efectos prácticos, se denominarán S1 y S2.

- Método 'S1': Script escrito en lenguaje C. Las fibras del sujeto se clasifican mediante el cálculo de la distancia Euclidiana máxima entre estas y cada centroide del atlas utilizado para segmentar. Este código es óptimo en su ejecución debido al descarte progresivo de fibras según sus distancias del punto central, y debido a su ejecución paralela, ya que las fibras y los centroides del atlas se evalúan en paralelo mediante múltiples hilos, también conocidos como threads [47].
- Método 'S2': Script escrito en lenguaje Python propuesto en [48]. Comprime la información de las fibras en un espacio de menor dimensionalidad llamado espacio latente utilizando el algoritmo Autoencoder de Deep Learning. En dicho espacio, mediante el algoritmo de ML Radius Neighbors Classifier, se segmentan los fascículos determinados por un atlas de referencia. Cabe mencionar, la precisión de la segmentación depende del ajuste del parámetro radial de búsqueda.

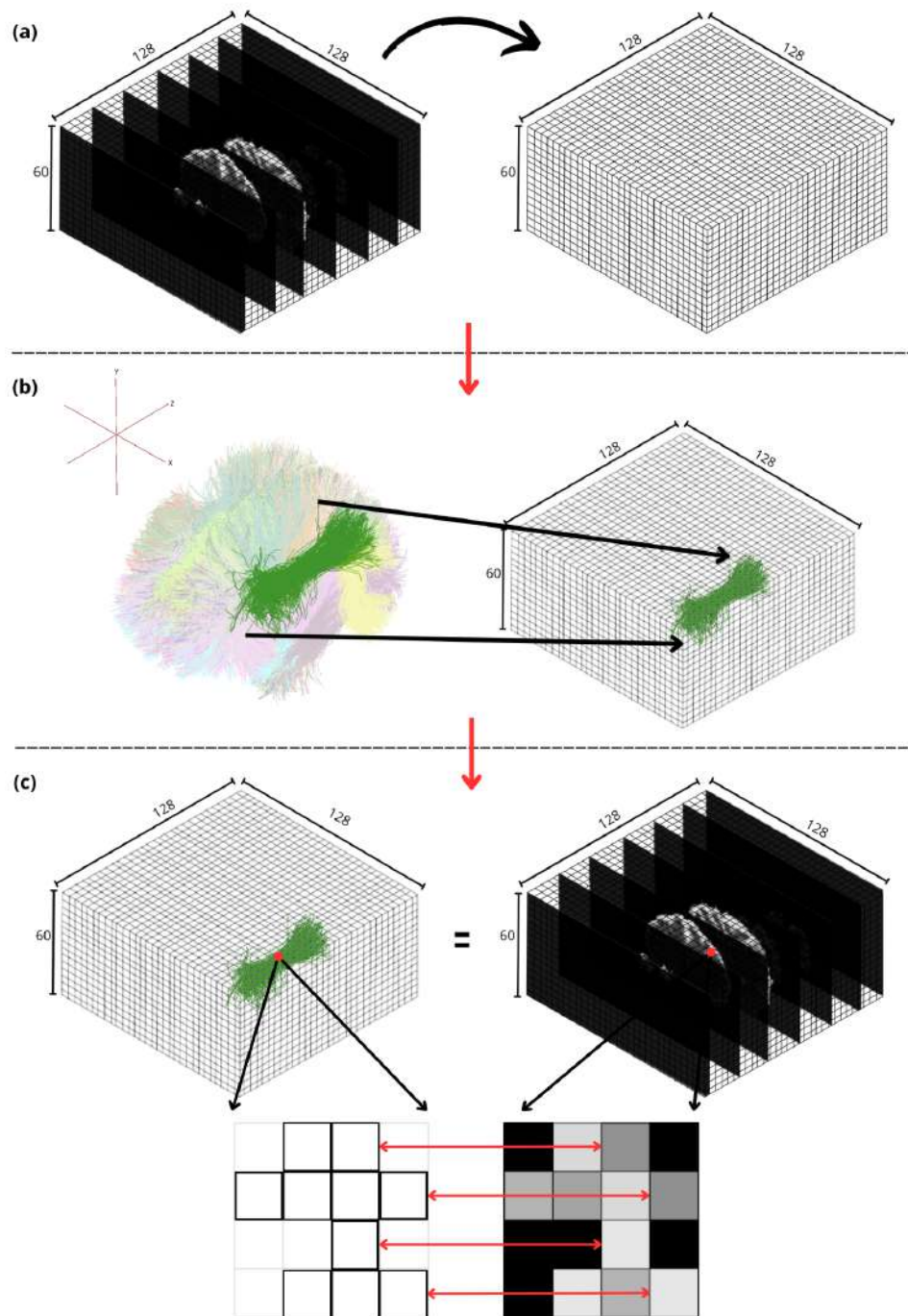
### 3.5.4 Obtención de valores promedios de medidas de difusión y anisotropía

Código escrito en lenguaje Python que permite localizar un fascículo en las imágenes de difusividad y anisotropía. La información de intensidad correspondiente al espacio en el que se posiciona el fascículo se guarda en una matriz del tamaño de la imagen [46]. Para utilizar este código, es necesario conocer el tamaño de los vóxeles en las dMRI, tener los fascículos segmentados en el espacio original de cada sujeto, y las imágenes, también conocidas como mapas de intensidad, de las medidas que se deseen estudiar.

Se define el orden de los ejes en el espacio y se asigna a una matriz que almacena la imagen. Luego, inicia un ciclo iterativo anidado que recorre cada fascículo, creando una matriz de ceros con las mismas dimensiones, ver Figura 3.15a, y recorre también cada fibra y punto de esta. A cada punto, se asigna una ubicación precisa dentro de la matriz tridimensional inicial, como se muestra en la Figura 3.15b. Posteriormente, se asigna el valor del vóxel a la matriz de ceros, como se ilustra en la Figura 3.15c.

Es posible que varios puntos compartan el mismo vóxel, pero el código solo sobrescribirá el valor en la nueva matriz sin afectar la representación. Además, la asignación de espacio al fascículo puede variar dependiendo de la ubicación espacial de la dMRI. La situación óptima ocurre cuando la dMRI está posicionada cerca del centro del sistema de coordenadas en el cuadrante positivo. En caso de que esta, y por consiguiente, la tractografía, no cumpla con esta disposición, se puede agregar un valor de desfase fijo. Una vez que se han asignado todos los valores de los vóxeles que atraviesa cada punto del fascículo a la nueva matriz, se calcula el promedio de intensidad al leer todos los elementos, excluyendo los ceros.





**Fig. 3.15:** Obtención de valores de intensidad de cada fascículo. En (a), se inicializa la imagen representativa de la medida de difusión o anisotropía como una matriz, seguido a esto, se crea una matriz vacía con las mismas dimensiones. En (b), se ubica el fascículo en la matriz de ceros, es decir, se reconoce qué coordenadas de la matriz ocupa el fascículo. En (c), buscamos dichas coordenadas en la matriz de la imagen inicializada en (a) y se asigna su contenido, es decir, el valor del vóxel, a la matriz vacía.

## 3.6 Bibliotecas de Python

Existen diversas bibliotecas de Python que permiten realizar procesos específicos en el campo del análisis y ciencia de datos, ofreciendo una gran cantidad de funciones y herramientas que facilitan la implementación de algoritmos, manipulación de información y visualización de resultados. Entre las más conocidas y utilizadas se encuentran las bibliotecas `os`, `math`, `numpy`, `matplotlib`, `nibabel` y `pandas`.

### 3.6.1 DIPY

Esta biblioteca se utiliza para procesar y analizar imágenes médicas 3D y 4D, datos de tractografía con el módulo `'tracking.streamline'`, entre otros. Contiene una amplia variedad de herramientas y algoritmos que permiten realizar tareas como la normalización espacial, el procesamiento de señales, el aprendizaje automático y el análisis estadístico de estas imágenes [49]. Se destaca:

- `'set_number_of_points'`: Comando que se utiliza para remuestrear los puntos de una fibra. Lee una lista de puntos y ajusta su cantidad a un valor proporcionado como parámetro de entrada, distribuyéndolos de manera equidistante.

### 3.6.2 SCIPY

Biblioteca ampliamente utilizada para realizar cálculos científicos y matemáticos. Incluye varios módulos con funcionalidades diversas, los cuales abarcan procesos de optimización, álgebra lineal, procesamiento de señales, estadísticas y más. En particular, el módulo `'stats'` se centra en el análisis estadístico [50]. Algunos de los comandos destacados de este módulo son:

- `'skew'`: Calcula el valor de la asimetría de una distribución estadística. Si dicho valor es cercano a 0, indica que la distribución es aproximadamente simétrica.
- `'kurtosis'`: Calcula el valor de curtosis de una distribución estadística. Si el valor de curtosis es igual o cercano a cero, la distribución sigue una forma de campana similar a la distribución normal.
- `'shapiro'`: Realiza la prueba de Shapiro-Wilk para evaluar la normalidad de una distribución estadística. Entrega el valor de prueba y el p-value. Si el valor de prueba es cercano a 1 y el p-value es mayor que un umbral predefinido, generalmente 0.05, la distribución sigue una forma de campana similar a la distribución normal.

- *'levene'*: Realiza la prueba de Levene para evaluar la igualdad de varianzas entre dos o más grupos. Se establece un valor de significancia, generalmente 0.05, y si el valor p obtenido es menor que este, indica que las varianzas no son homogéneas.
- *'bartlett'*: Realiza la prueba de Bartlett, y al igual que el comando anterior, evalúa la igualdad de varianzas entre dos o más grupos, siguiendo la misma lógica e hipótesis.
- *'ttest\_ind'*: Realiza la prueba T de Student para muestras independientes. Se establece un valor de significancia, generalmente 0.05, y si el valor p obtenido es menor que este, indica que hay diferencias estadísticamente significativas entre los grupos.
- *'mannwhitneyu'*: Realizar la prueba de Mann-Whitney U. Al igual que el comando anterior, se establece un valor de significancia, generalmente 0.05, y sigue la misma lógica e hipótesis.

### 3.6.3 Scikit-learn

Conocida como *'sklearn'*, es una biblioteca de aprendizaje automático (ML), la cual proporciona una amplia variedad de algoritmos y herramientas para realizar diversas tareas, tales como clasificación, regresión, reducción de dimensionalidad, entre otras [51]. Algunos de sus comandos disponibles son:

- *'SVC'*: Del módulo *'svm'*, diseñado para resolver problemas de clasificación por medio de Máquinas de Vectores de Soporte (SVM).
- *'DecisionTreeClassifier'*: Comando del módulo *'tree'*, se utiliza para crear un clasificador a partir del modelo de Árboles de Decisiones.
- *'train\_test\_split'*: Función del módulo *'model\_selection'*. Se utiliza para dividir en porcentajes especificados un conjunto de datos en conjuntos de entrenamiento y prueba. Además, se puede especificar el valor del parámetro *random\_state* para reconocer la aleatoriedad de la división.
- *'GridSearchCV'*: Del módulo *'model\_selection'*. Sirve para encontrar la mejor combinación de parámetros evaluando el rendimiento de un modelo.
- *'KFold'*: Del módulo *'model\_selection'*. Se utiliza dentro de *'GridSearchCV'* para dividir un conjunto de datos en subconjuntos llamados pliegues y realizar validación cruzada.
- *'accuracy\_score'*: Función del módulo *'metrics'*. Calcula la precisión de un modelo de clasificación, siguiendo la lógica de la ecuación (2.10.1).

- *'precision\_score'*: Del módulo *'metrics'*. También calcula la precisión de un modelo de clasificación, mediante otro enfoque, siguiendo la lógica de la ecuación (2.10.2).
- *'recall\_score'*: Función del módulo *'metrics'*. Obtiene la sensibilidad de un modelo de clasificación, siguiendo la lógica de la ecuación (2.10.3).
- *'f1\_score'*: Del módulo *'metrics'*. Se utiliza para calcular el valor F1 de un modelo de clasificación. Proporciona una medida equilibrada del rendimiento del modelo, siguiendo la ecuación (2.10.4).

### 3.7 Discusión

Durante el diseño del pipeline de procesamiento de datos, se encontraron ciertas consideraciones y desafíos que debían abordarse. En primer lugar, la base de datos utilizada había sido procesada previamente, lo que implicaba lidiar con una gran cantidad de archivos y un contenido complejo de comprender para una estudiante nueva en el campo de la neurociencia.

Eventualmente se reconocieron los archivos necesarios a utilizar, y se observó que los datos de tractografía presentaban un corte en la zona inferior, específicamente en el cerebelo, el cual procedía de una alteración en las dMRI originales. Afortunadamente, se determinó que esta área no era relevante para los objetivos del presente estudio.

Es importante tener en cuenta que trabajar con una base de datos modificada no es lo óptimo, pero se destaca la existencia de una buena representatividad entre los sujetos con, y sin autismo. Sin embargo, es importante mencionar que debido al tamaño relativamente pequeño de la muestra, compuesta por 37 sujetos, podría haber limitaciones en la generalización de los resultados y un mayor riesgo de sesgo.

Además, al tratarse de sujetos adolescentes, había que considerar la posibilidad de incompatibilidad con ciertos atlas de fascículos o metodologías utilizadas. Para abordar esta diferencia etaria manifestada en el cerebro y sus estructuras, se optó por utilizar un atlas de fibras largas y realizar una transformación espacial adecuada, lo que permitió equiparar las diferencias anatómicas entre los sujetos.

Si bien existen otras medidas para evaluar la difusividad y anisotropía, las medidas ADC, FA y GFA son suficientes para analizar estas características. Cabe mencionar que las imágenes originales estaban en formato GIS, sin embargo, se optó por convertirlas al formato NIFTI debido a que resultaba más conveniente y sencillo realizar la conversión del archivo, en lugar de adaptar los scripts.

## Capítulo 4. Aplicación del pipeline

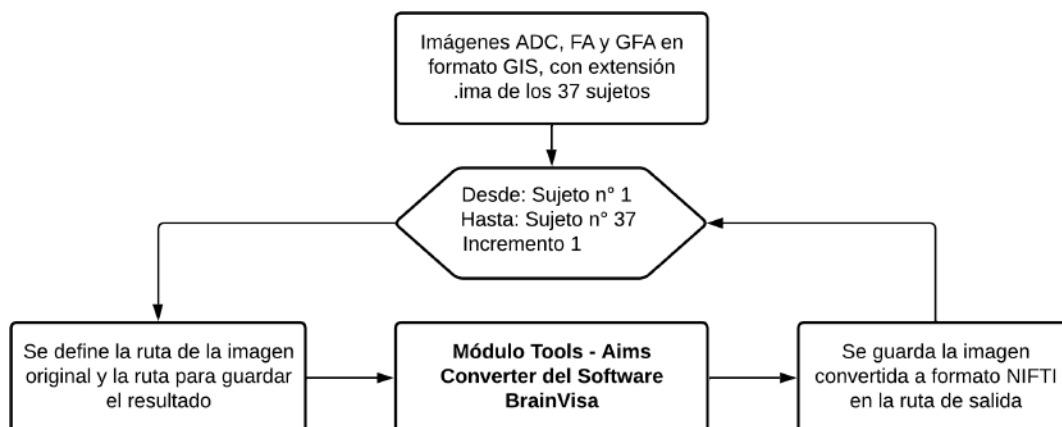
### 4.1 Introducción

Se describe la aplicación del pipeline propuesto en la Figura 1.1. Para ello se utilizaron los datos e información mencionados en el capítulo anterior, adaptando los scripts y archivos que fuesen necesarios para mejorar la obtención de medidas promedio de difusividad y anisotropía de los fascículos de cada sujeto, permitiendo evaluar sus características de distinción más relevantes.

### 4.2 Conversión de formato de imagen

El formato GIS de las imágenes FA, GFA y ADC no es compatible con el código disponible para obtener los valores de intensidad. Por ende, se transforma el archivo de extensión IMA, archivo que contiene la imagen, al formato NIFTI, uno de los más utilizados en imágenes médicas y que además es apropiado para este proceso.

Esta fase del Pipeline se lleva a cabo mediante la herramienta Aims Converter del software BrainVisa, y se itera para realizar el proceso automáticamente en cada uno de los sujetos, indicando la dirección, nombre, tipo de archivo de entrada y salida, ver Figura 4.16.



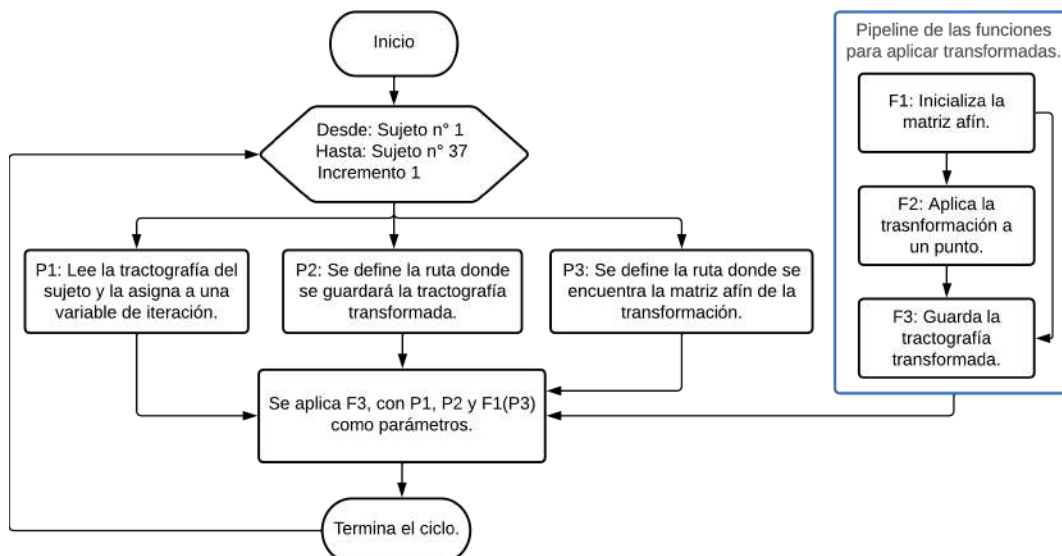
**Fig. 4.16:** Conversión iterativa de formato GIS (ima) a NIFTI (nii) para las imágenes de los 37 sujetos de la base de datos HARDI.

### 4.3 Transformaciones espaciales

Es necesario que los archivos de tractografía se encuentren en un solo espacio normalizado. Para lograr esto, se realizan dos transformaciones sucesivas mediante sus respectivas matrices afines. La primera para transformar el espacio de difusión del sujeto al espacio T1 del mismo, y luego, del espacio T1 al espacio normalizado Talairach. Para automatizar este proceso, se utiliza un ciclo 'for' que adapta el código de la aplicación de transformada. La metodología se muestra en la Figura 4.17.

Se emplea álgebra matricial para fusionar ambas transformaciones en una. Se denota la matriz de transformación a T1 como  $MA1$  y la matriz afín de Talairach como  $MA2$ . Ambas matrices siguen la forma indicada en la ecuación (3.5.1), por lo que es posible multiplicarlas entre sí para obtener una única matriz afín  $MA'$ . La transformación espacial se aplica multiplicando una matriz afín por cada coordenada expresada como un vector  $P$  de dimensiones  $4 \times 1$ , como se ve en la ecuación (3.5.3). Por lo tanto, aplicar primero  $MA1$  al vector  $P$ , y luego  $MA2$  al resultado anterior, se puede representar mediante la ecuación (4.3.1). Es importante recordar que la multiplicación matricial no es conmutativa, pero sí es asociativa. Dicho esto, la fórmula propuesta es equivalente a la ecuación (4.3.2), y a su vez, a la ecuación (4.3.3)

Finalmente, el diagrama de la Figura 4.17 es adaptado, permitiendo ingresar ambas matrices en vez de una sola, y se ajusta la función F1 agregando el comando '*np.dot*', el cual realiza la multiplicación de las matrices  $MA1$  y  $MA2$  para obtener  $MA'$ . El resto del script permanece sin cambios.



**Fig. 4.17:** Aplicación de la transformación espacial. Inicialmente, ocurren dos procesos. P1 inicializa la tractografía, P2 indica la ruta de salida, y P3 inicializa la ruta de la matriz afín. Una vez que hayan ejecutado P1, P2 y P3, se aplican las funciones F1, F2 y F3, que se explican con mayor claridad en la Figura 3.14.

$$MA2_{4 \times 4} \cdot (MA1_{4 \times 4} \cdot P_{4 \times 1}) = P'_{4 \times 1} \quad (4.3.1)$$

$$(MA2_{4 \times 4} \cdot MA1_{4 \times p}) \cdot P_{4 \times 1} = P'_{4 \times 1} \quad (4.3.2)$$

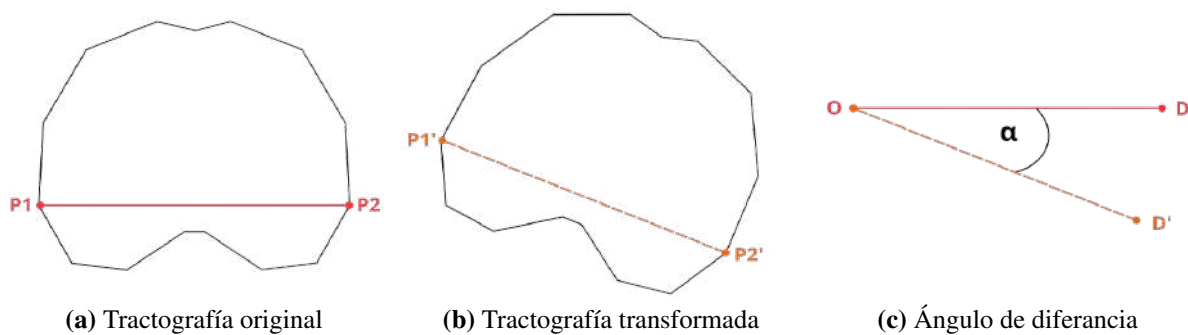
$$MA'_{4 \times 4} \cdot P_{4 \times 1} = P'_{4 \times 1} \quad (4.3.3)$$

### 4.3.1 Corrección de rotación errónea

Considerando que las tractografías presentan un corte en la sección del cerebelo, era posible que su comportamiento ante ciertos procesamientos no fuera el óptimo. Al aplicar las transformaciones, se pudo observar que 12 de las 37 tractografías presentaban una rotación errónea en el eje 'y'.

Para determinar el ángulo a corregir, se leyó la tractografía original centrada en 'y', identificando la ubicación del punto mínimo ( $P1$ ) y máximo ( $P2$ ) del eje 'x', ver Figura 4.18a. Luego, se leyó la tractografía transformada y se ubicaron los puntos ( $P1'$ ) y ( $P2'$ ), ver Figura 4.18b. Una vez obtenidos ambos segmentos, se les realizó una traslación simple al origen del plano, ver Figura 4.18c.

Dentro de la función arcocoseno, ecuación (4.3.4), se obtiene el producto escalar entre  $D$  y  $D'$  utilizando el comando '*np.dot*', mientras que en el denominador, se multiplican las magnitudes de los vectores, obteniendo el valor del ángulo. Finalmente, el saber si se deberá aplicar de manera antihoraria u horaria, se determina con una condicional. Si la coordenada en 'y' de  $D$  es mayor que la coordenada en 'y' de  $D'$ , el ángulo es positivo, de lo contrario, es negativo. Los valores resultantes se registran en la Tabla 4.1.



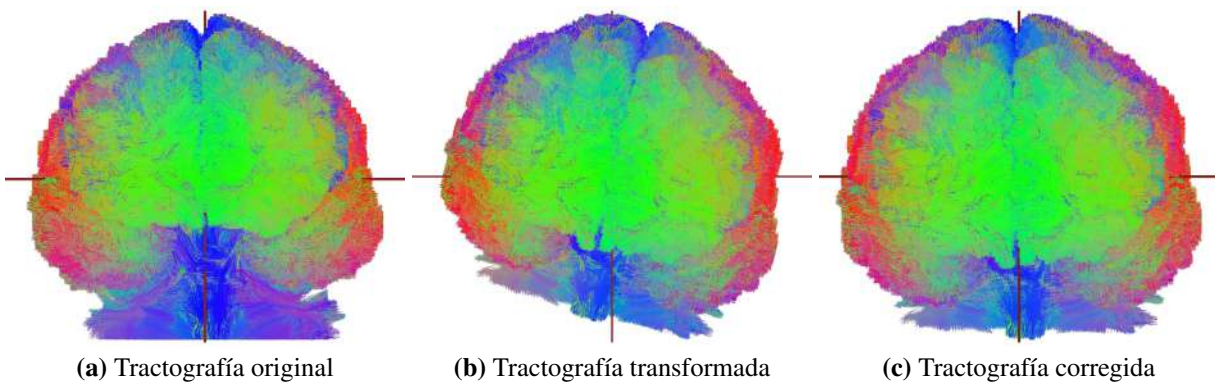
**Fig. 4.18:** Representación visual de la detección del ángulo de rotación. En (a) se observan los puntos mínimo y máximo en el eje 'x' de la tractografía original. En (b) se aprecian los mismos puntos en la tractografía resultante de la transformación. En (c) se aprecian ambos segmentos compartiendo el mismo origen y ángulo.

$$\alpha = \arccos\left(\frac{D \cdot D'}{|D| \times |D'|}\right) \quad (4.3.4)$$

Una vez obtenido el ángulo de rotación necesario, se sigue la misma lógica planteada en la Figura 4.17. La matriz afín aplicada corresponde a la matriz de rotación en el eje 'y' mostrada en la ecuación (2.8.5). Un ejemplo de este proceso se ilustra en la Figura 4.19.

Sujeto	bw	ch	ch	db	ef	ig	jp	lk	nl	ru	sb3	sb0
Ángulo	4°	4°	8°	14°	-15°	6°	6°	-4°	9°	4°	6°	-5°

**Tabla 4.1:** Ángulos de rotación obtenidos en grados para aplicar la corrección en el eje 'y' según el sujeto.



**Fig. 4.19:** Ejemplo de corrección rotacional en el sujeto db100111, alias 'db'. En (a) se aprecia el estado original. En (b) se observa el resultado de la transformación espacial a Talairach con una rotación incorrecta en el eje 'y'. En (c) se observa la tractografía transformada al espacio Talairach con la rotación alineada correctamente.

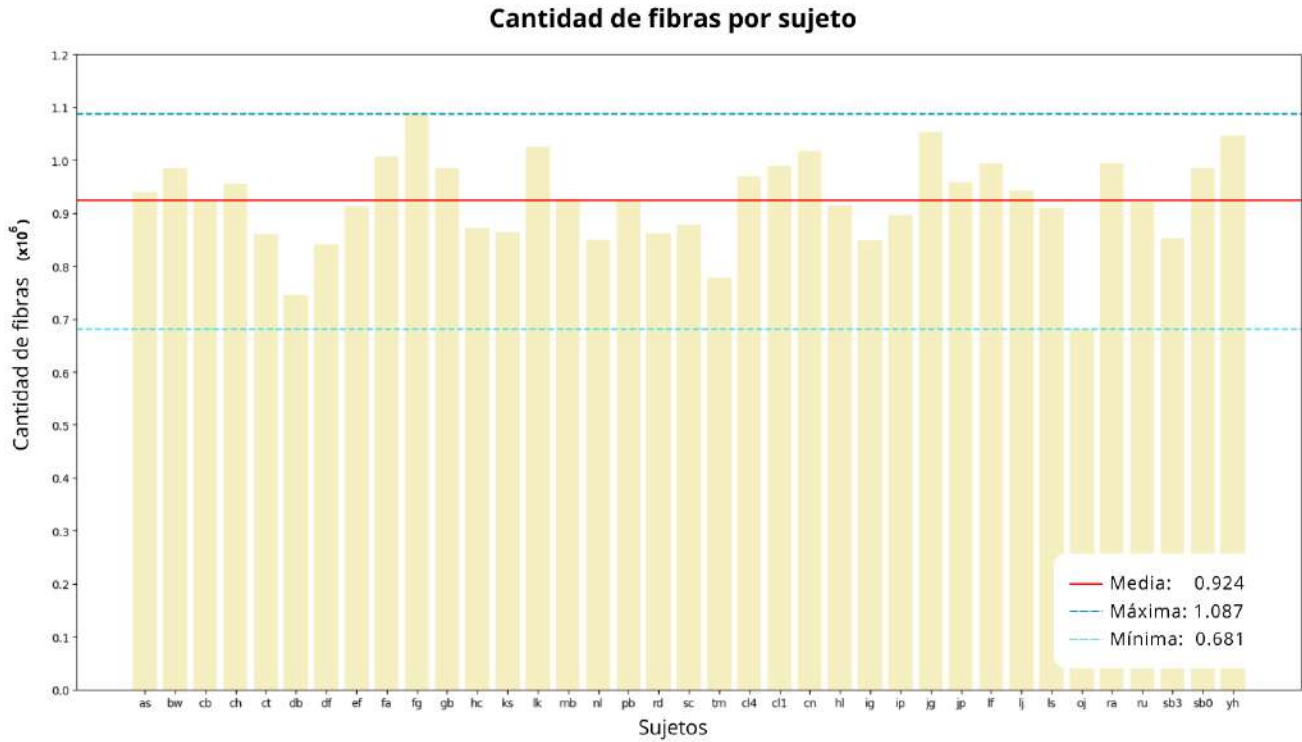
## 4.4 Remuestreo de fibras

Para determinar la secuencia adecuada entre las transformaciones espaciales y el remuestreo de fibras, se realizaron análisis matemáticos. En la base de datos, se encontró que la variación en el número de fibras en los datos de tractografía oscila entre 681 mil y 1,087 millones, con un promedio de 924 mil, ver Figura 4.20. Además, se registró el promedio de la cantidad de puntos que conformaban las fibras, las cuales variaban de 2 hasta 62 puntos, ver Figura 4.21.

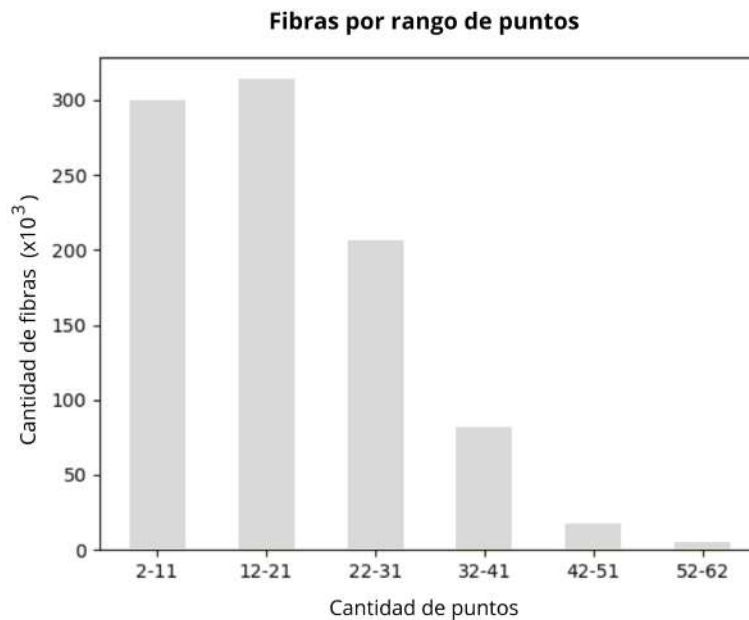
Al analizar el porcentaje de fibras en función de la cantidad de puntos que las conformaban, se obtuvo que el 63.8 % de las fibras están compuestas por menos de 21 puntos. Teniendo en cuenta que las transformaciones son procesos que se realizan punto a punto, se decidió remuestrear los archivos de tractografía después de dicha operación para optimizar el tiempo y el trabajo involucrado.

El valor de muestreo puede variar a discreción del usuario. En este caso, se aplicó un remuestreo de 21 puntos equidistantes en cada fibra, ver Figura 4.22, y se guardó el resultado para la segmentación 'S1'.

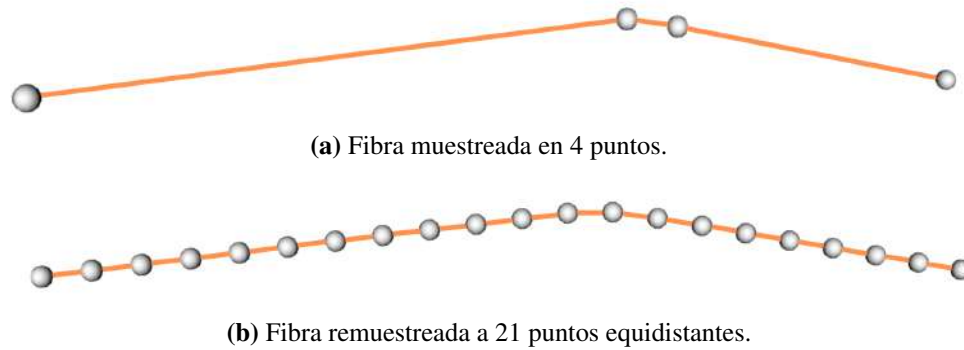




**Fig. 4.20:** Cantidad de fibras por tractografía de los 37 sujetos. En el eje 'y' se encuentran valores de 1 es a 1 millón. Además, se marca con una línea horizontal el promedio, el valor máximo y mínimo.



**Fig. 4.21:** Cantidad de fibras según rango de puntos. Para una mejor representación, la cantidad de puntos se abordó en rangos de 10, con excepción del último que incluye 11 debido a la cantidad impar.



**Fig. 4.22:** Remuestreo de 21 puntos equidistantes. En (a) se aprecia la fibra transformada previamente al espacio Talairach, muestreada a 4 puntos no equidistantes. En (b) se observa la misma fibra luego de ser remuestreada.

## 4.5 Segmentación de fascículos

Se implementaron dos métodos distintos para la segmentación de fascículos. En ambas ocasiones se utilizó el atlas de 36 fascículos de fibras largas en el espacio Talairach, ver Figura 3.13.

### 4.5.1 Método 'S1'

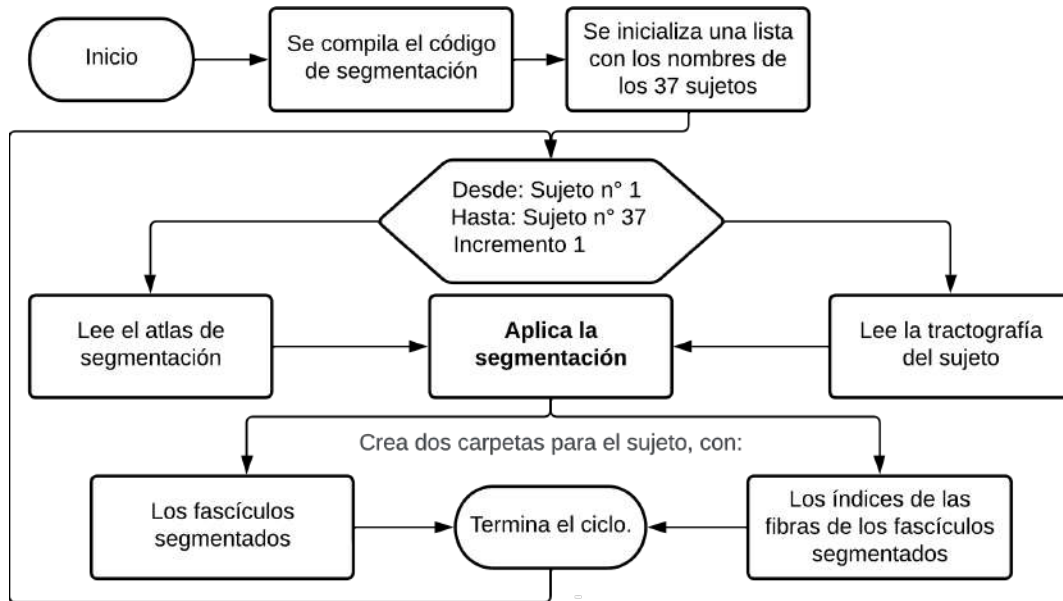
Mediante el sistema operativo Ubuntu, y los datos de tractografía en el espacio Talairach con las fibras muestreadas en 21 puntos, se ejecuta la línea `'g++ -std=c++14 -O3 main.cpp -o main -fopenmp -ffast-math'` que compila el código de segmentación `'main.cpp'` utilizando el estándar de C++14, habilita la optimización de nivel 3, la compatibilidad con OpenMP y aplica optimizaciones matemáticas.

A continuación, se inicia un ciclo for que especifica la ubicación del archivo de tractografía, la dirección del atlas a utilizar y el archivo de texto con su información. También se ordena la creación de carpetas para almacenar los fascículos y sus índices, siguiendo el diagrama mostrado en la Figura 4.23.

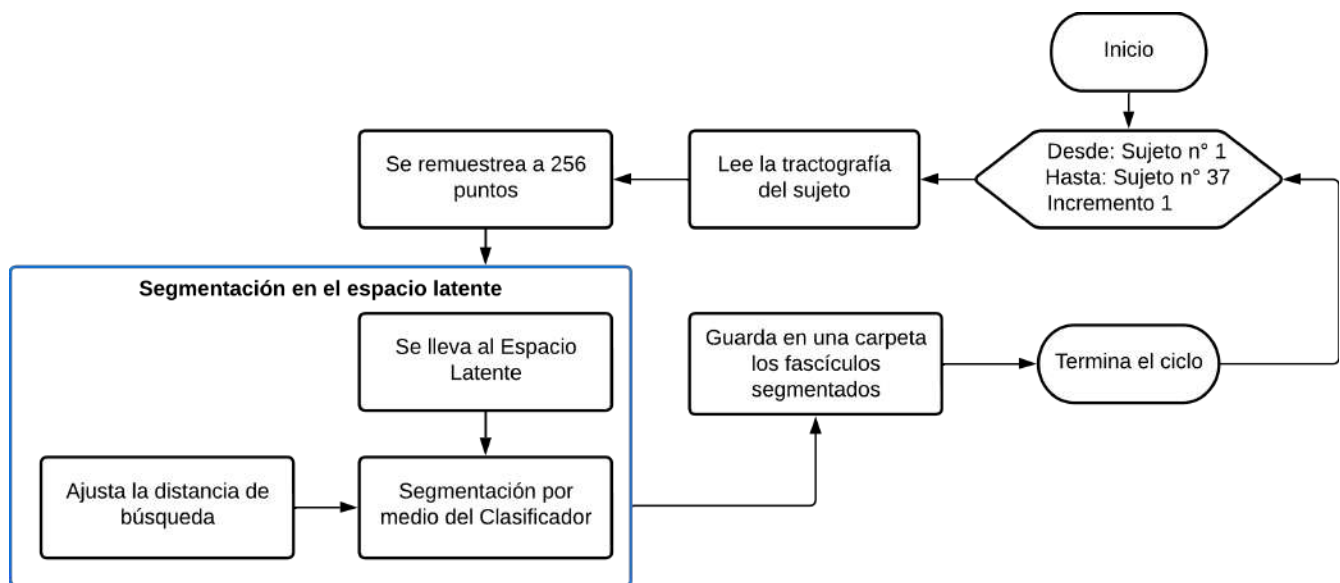
### 4.5.2 Método 'S2'

Para ejecutar el método de segmentación creado y entrenado en [48], se requirió que los datos de tractografía se encontraran en el espacio Talairach con las fibras muestreadas en 256 puntos. Sin embargo, antes de ser utilizados en este script, los datos de tractografía se remuestrearon y guardaron en 21 puntos. Por ello, el código se adapta, y posterior a que lee los datos, se aplica el comando de remuestreo.

Una vez normalizados apropiadamente los datos de tractografía, se especifica el parámetro del clasificador referente a la distancia de búsqueda, y el código proporcionado procede a clasificar las fibras en los fascículos correspondientes, guardándolas finalmente en una carpeta de resultados por sujeto. Para repetir lo anterior, se realizó un ciclo 'for' como se aprecia en la Figura 4.24.



**Fig. 4.23:** Diagrama de metodología empleada en la segmentación 'S1'. Inicialmente se ejecuta el código de segmentación y el listado de sujetos, para luego iterar un ciclo for.



**Fig. 4.24:** Diagrama de metodología empleada para el método de segmentación 'S2'.

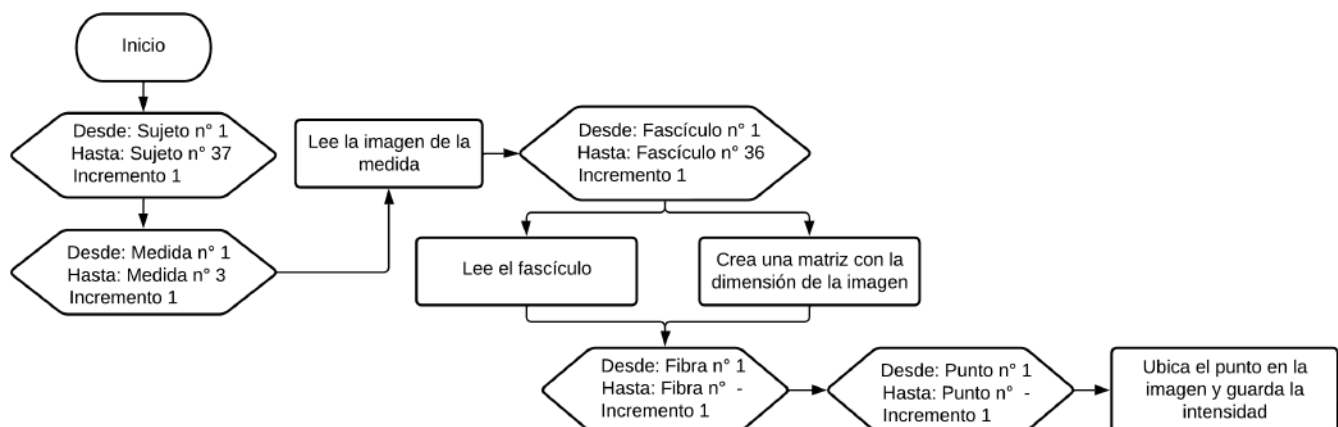
## 4.6 Transformación inversa

Las imágenes de difusividad y anisotropía se encuentran en el espacio propio de la dMRI de cada sujeto, al igual que las tractografías originales. Dado que necesitamos ubicar los fascículos en dichas imágenes, debemos aplicar las transformaciones inversas para llevar los fascículos a su espacio propio. Para ello, se realiza una modificación en el código previamente descrito para aplicar las transformaciones, ver Figura 4.17. La metodología es la misma, pero en la ecuación (3.5.1) se reemplaza la matriz  $MA$  por su inversa  $MA^{-1}$ . Esta matriz se obtiene utilizando el comando `'linalg.inv'` de la biblioteca Numpy en Python, siguiendo la propiedad de invertida, donde  $In$  representa la matriz identidad.

Con esto en mente, en primer lugar, se aplica la inversa de la rotación a los 12 sujetos que recibieron una rotación adicional en el eje 'y'. Luego, a los 37 sujetos se les aplica la transformación inversa de la matriz combinada que realiza la transformación a Talairach en cada caso.

## 4.7 Aplicación de máscaras

Los valores de ADC, FA y GFA corresponden a las intensidades de las imágenes conocidas como mapas de intensidades, por ende, se ubican los fascículos en dichas imágenes para obtener su promedio. Este proceso se lleva a cabo utilizando los fascículos segmentados, para cada una de las medidas. Una vez extraídas las intensidades equivalentes en cada punto, de cada fibra, de cada fascículo en la imagen, se calculan los promedios de valores diferentes de cero, y estos se almacenan en un archivo Excel mediante un DataFrame. La metodología empleada en esta aplicación se muestra en el diagrama de la Figura 4.25.



**Fig. 4.25:** Implementación de fascículos como máscaras en las imágenes ADC, FA y GFA. Para cada sujeto, se leyeron y ubicaron los puntos correspondientes a cada fascículo en las imágenes. Cabe mencionar, el valor máximo de iteración en los últimos ciclos no se define, ya que depende de la información de recopilada antes.

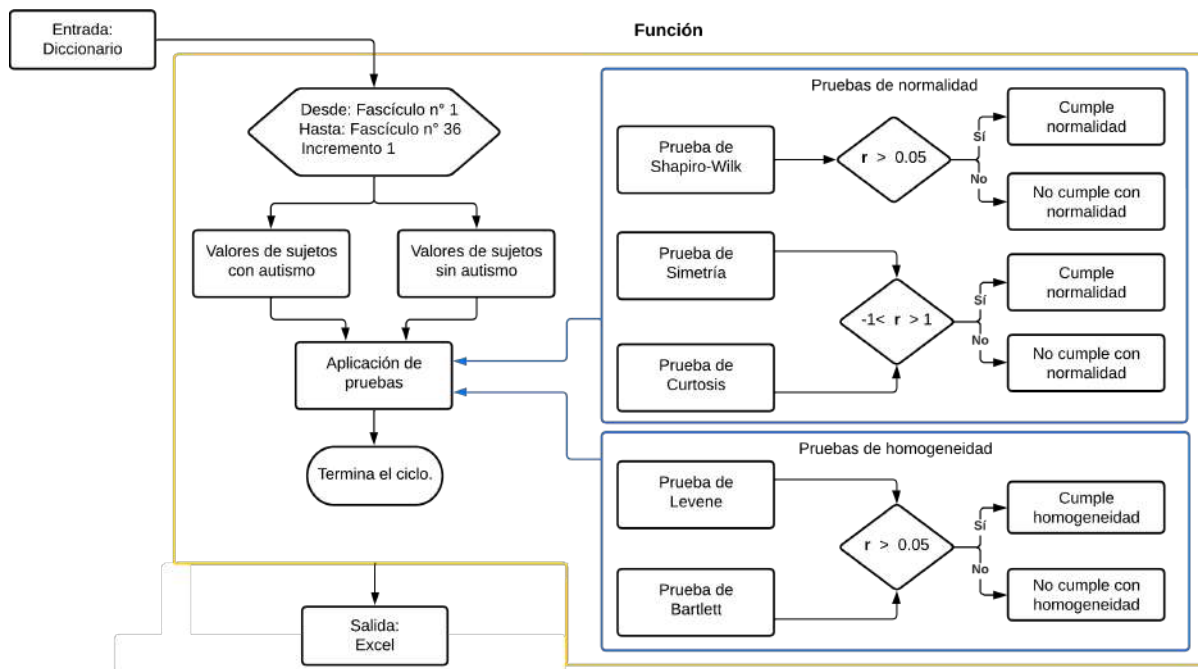
## 4.8 Test estadísticos

Para determinar qué tipo de pruebas inferenciales aplicar, es necesario conocer la normalidad y homogeneidad de las muestras a través de pruebas estadísticas descriptivas. Si cumplen con la normalidad y homogeneidad, se pueden aplicar pruebas paramétricas, de lo contrario, pruebas no paramétricas.

### 4.8.1 Pruebas de estadística descriptiva

En primera instancia, se crearon 6 diccionarios en Python, uno por cada una de las tres medidas que se desean evaluar a partir de los resultados del método de segmentación S1 y el método S2. Las llaves del diccionario corresponden a las etiquetas de los 36 fascículos segmentados. Los valores asociados a cada etiqueta son listas compuestas por dos sub-listas. La primera contiene los valores de los sujetos sin autismo, mientras que la segunda sub-lista contiene los valores de los sujetos con autismo.

Se realizaron las pruebas de normalidad de Shapiro-Wilk, curtosis y asimetría por fascículo y por muestra, es decir, tanto para sujetos con autismo como para sujetos sin autismo. Además, se llevaron a cabo las pruebas de homogeneidad de Levene y Bartlett entre muestras. Estas pruebas se organizaron dentro de una función a la cual solo se le requería ingresar el diccionario como entrada, ver Figura 4.26.

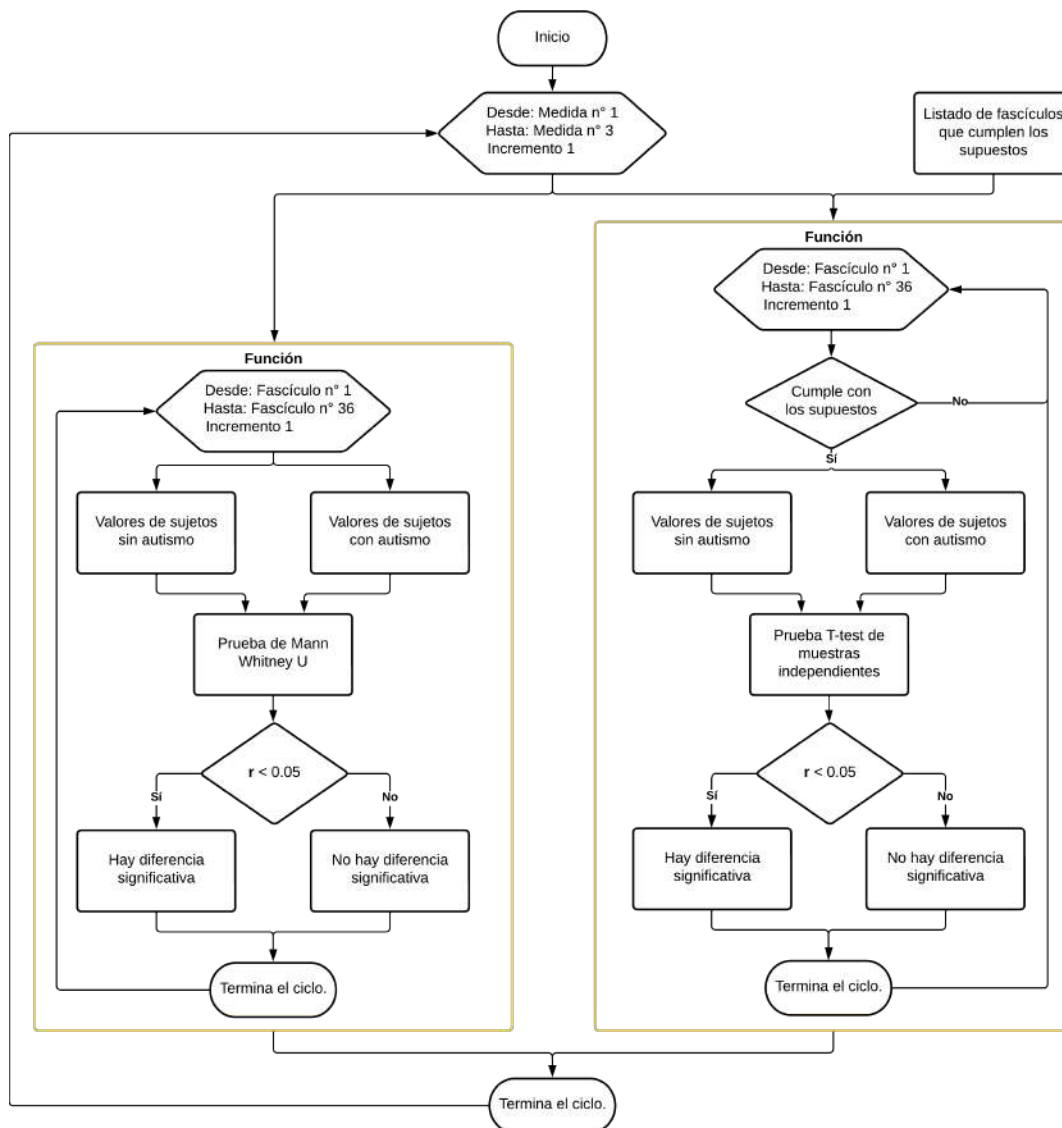


**Fig. 4.26:** Implementación de pruebas estadísticas descriptivas. Consiste en una función que abarca ambos tipos de pruebas, y las aplica a cada fascículo, 'r' es el valor de resultado y 0.05 el valor de significancia elegido. Recibe como entrada un diccionario con los valores de ADC, FA o GFA según los resultados de S1 o S2.

## 4.8.2 Pruebas de estadística inferencial

Se utilizaron los mismos 6 diccionarios mencionados anteriormente complementados con los resultados de las pruebas estadísticas descriptivas, y se estableció un nivel de significancia de 0.05.

En el caso de T-test de muestras independientes, se empleó un condicional respaldado por la información recopilada. En cambio, cada fascículo se evaluó mediante el test de Mann-Whitney U sin distinción. La implementación se muestra en el diagrama de la Figura 4.27.



**Fig. 4.27:** Aplicación de pruebas estadísticas inferenciales paramétricas y no paramétricas. Se lleva a cabo un ciclo iterativo para los diccionarios que contienen los valores de ADC, FA y GFA, ya sea para los resultados de S1 o S2. Por un lado, se ejecuta la prueba no paramétrica de Mann-Whitney U, y por otro, la función que aplica la prueba paramétrica de T-test para muestras independientes, considerar que 'r' corresponde al valor p.

## 4.9 Machine Learning

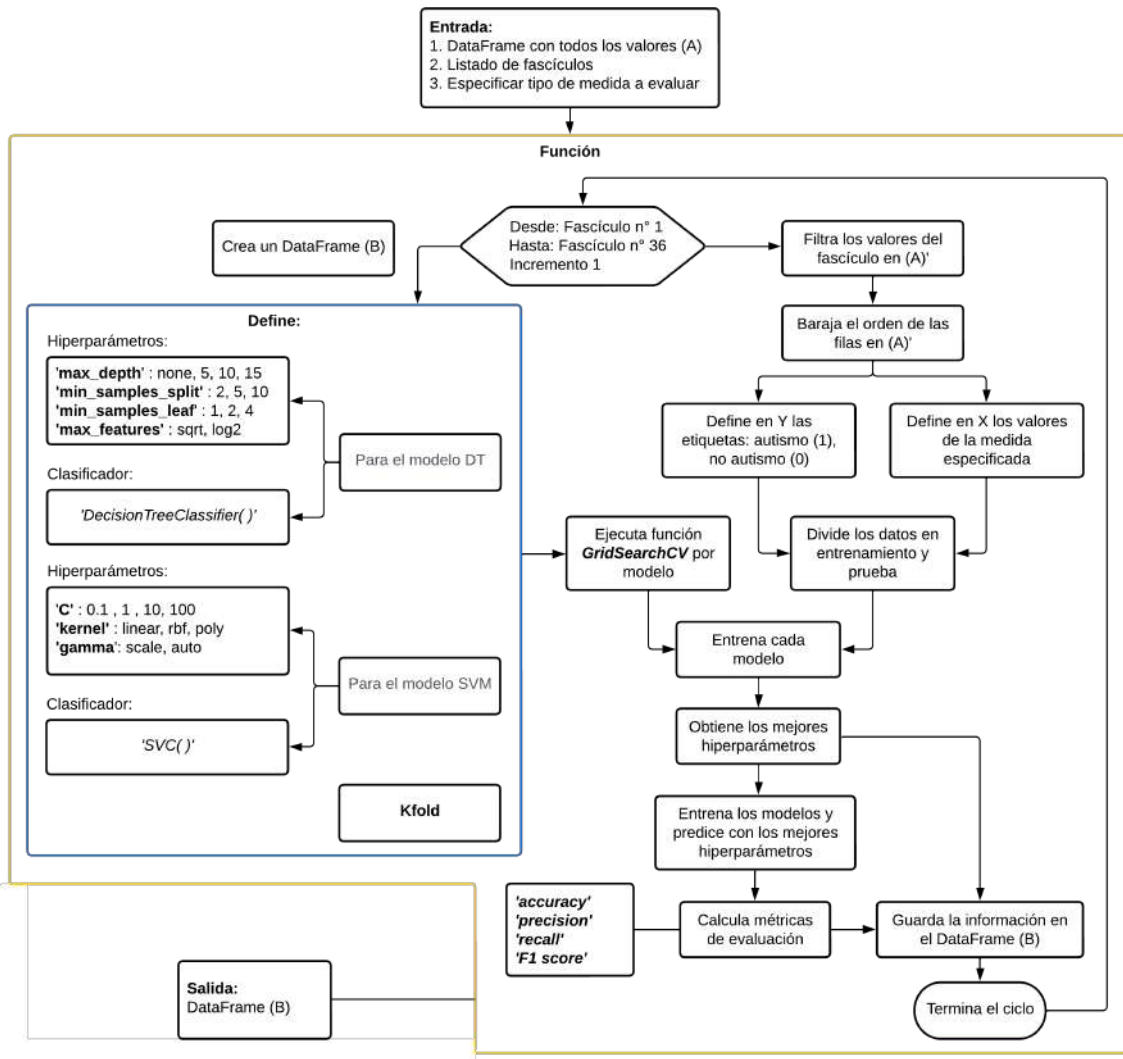
Se implementaron algoritmos de Machine Learning utilizando los resultados de ADC, FA y GFA obtenidos para cada fascículo mediante el método de segmentación S2. Estos datos se organizaron en un solo archivo Excel, donde se incluyeron las tres medidas por fascículo, así como una etiqueta que indica si pertenecen a sujetos con autismo o sin autismo.

Con el fin de optimizar el script, se diseñó una función que toma como entrada el DataFrame del Excel mencionado anteriormente, una lista con los nombres de los 36 fascículos y la especificación de la medida que se desea evaluar. Esta función aplica dos modelos de aprendizaje supervisado, Máquinas de Vectores de Soporte (SVM) y Árboles de Decisión (DT). Dado que el tamaño del conjunto de datos por fascículo es de 37, considerado bajo, se realiza una división de entrenamiento y prueba del 60 % frente al 40 % respectivamente, siendo el orden de aleatoriedad dirigido por el valor de *'random\_state'* que cambia por medio de un contador.

Además, se utiliza la técnica de validación cruzada Grid Search para la optimización de hiperparámetros del modelo. Esta técnica consiste en seleccionar diferentes valores para cada parámetro y generar una cuadrícula con todas las posibles combinaciones. Para evaluar el rendimiento del modelo en cada combinación de parámetros, se utiliza la técnica de K-fold, con K igual a 5. En cada iteración, se seleccionan distintos subconjuntos de entrenamiento y prueba, evitando así el sesgo debido a la selección de un solo conjunto. Finalmente, se selecciona la combinación de parámetros que produce el mejor rendimiento promedio en términos de la métrica *'accuracy'*.

Para SVM, se exploraron los hiperparámetros *'C'*, *'kernel'* y *'gamma'*, mientras que para DT se utilizaron los hiperparámetros de profundidad máxima (*'max\_depth'*), mínimo de muestras para dividir (*'min\_samples\_split'*), mínimo de muestras por hoja (*'min\_samples\_leaf'*) y máximas características (*'max\_features'*). Una vez definidas las mejores combinaciones para cada modelo, se volvieron a entrenar y realizaron las predicciones. Finalmente, se calcularon las métricas de evaluación *'accuracy'*, *'precision'*, *'recall'* y *'F1 Score'* para cada uno. El diagrama de implementación y los valores que se probaron en cada hiperparámetro se muestran en la Figura 4.28.

La función se utilizó para analizar los datos de ADC, FA y GFA, así como las combinaciones de ADC-FA, ADC-GFA, FA-GFA y ADC-FA-GFA, en cada fascículo por separado. Los resultados de los mejores hiperparámetros y las métricas de evaluación se registraron en un DataFrame por tipo/s de medida/s evaluada/s. En este DataFrame, cada columna representa un fascículo específico, lo que permite una fácil comparación de los resultados entre los diferentes modelos y medidas.



**Fig. 4.28:** Aplicación de modelos de clasificación de Machine Learning. Función que utiliza la técnica de validación cruzada Grid Search para obtener los mejores hiperparámetros en los modelos de Support Vector Machine y Árboles de decisión. Una vez obtenidos, se aplican ambos modelos a los valores de ADC, FA o GFA de cada fascículo. Los resultados se almacenan en un DataFrame que incluye los mejores hiperparámetros utilizados por modelo en cada fascículo, junto con las métricas de evaluación correspondientes.

## 4.10 Discusión

Durante la implementación del pipeline, se encontraron varios factores que requirieron consideración y ajustes. En primer lugar, se tuvieron que realizar cambios en los algoritmos de transformación, y desarrollar nuevos, para abordar el corte presente en las tractografías. Estos cambios permitieron superar esta limitación y garantizar un procesamiento adecuado de los datos.



Cabe mencionar que la decisión de realizar el remuestreo de las fibras de tractografías, después de la transformación espacial, fue principalmente debido a consideraciones de tiempo. Cuantos más puntos tenía una tractografía, más tiempo tomaba realizar la transformación correspondiente.

Por otro lado, uno de los desafíos que surgieron fue la elección de los hiperparámetros en los modelos de Machine Learning, pero no por la complejidad del algoritmo, más bien, la función ejecutada resultó ser computacionalmente pesada y requería de un ordenador más potente para una continua reproducción.

Se observó que los tiempos de ejecución de los scripts variaban según las funciones involucradas. La conversión de formato de imagen y la segmentación 'S1' fueron los procesos más rápidos, tomando un par de segundos por sujeto. Por otro lado, la transformación espacial a Talairach, la segmentación de fascículos mediante el método 'S2', y el reconocimiento y aplicación de mejores hiperparámetros en modelos de clasificación, fueron los más lentos. Sin embargo, es importante tener en cuenta que esta información es relativa y depende del procesador utilizado y de la presencia de funciones paralelas.

Finalmente, es necesario destacar que si bien en condiciones ideales el pipeline de principio a fin, no debería tardar más de una hora por sujeto, estas son condiciones donde los datos estén listos, ordenados, dispuesto para su procesamiento, acompañada la ejecución por un ordenador más eficiente. Sin embargo, en la realidad, el proceso de adaptación de los scripts y su ejecución inicial tomó semanas, e incluso meses. Pero a medida que se adquiría más conocimiento, se perfeccionaban los códigos y se contaba con un mejor ordenador, el tiempo de ejecución se redujo significativamente. De esta manera, lo que antes llevaba semanas en completarse, con pausas para evaluar el rendimiento, podía realizarse en tan solo unas horas para todos los sujetos involucrados en el estudio.

## Capítulo 5. Resultados

---

### 5.1 Introducción

A continuación, se presentan los resultados de los dos métodos de segmentación en términos de la cantidad de fibras segmentadas y su representatividad, la verificación de los procesos aplicados, el análisis de la variabilidad de los valores promedio de ADC, FA y GFA por fascículo con dichos métodos, las pruebas estadísticas para identificar fascículos con diferencias significativas y los resultados de los modelos de clasificación de ML para determinar los fascículos con buen rendimiento. Estos resultados brindarán información clave sobre la efectividad de los métodos de segmentación, la calidad de los procesos, la variación en los parámetros de los fascículos y su relación con el autismo.

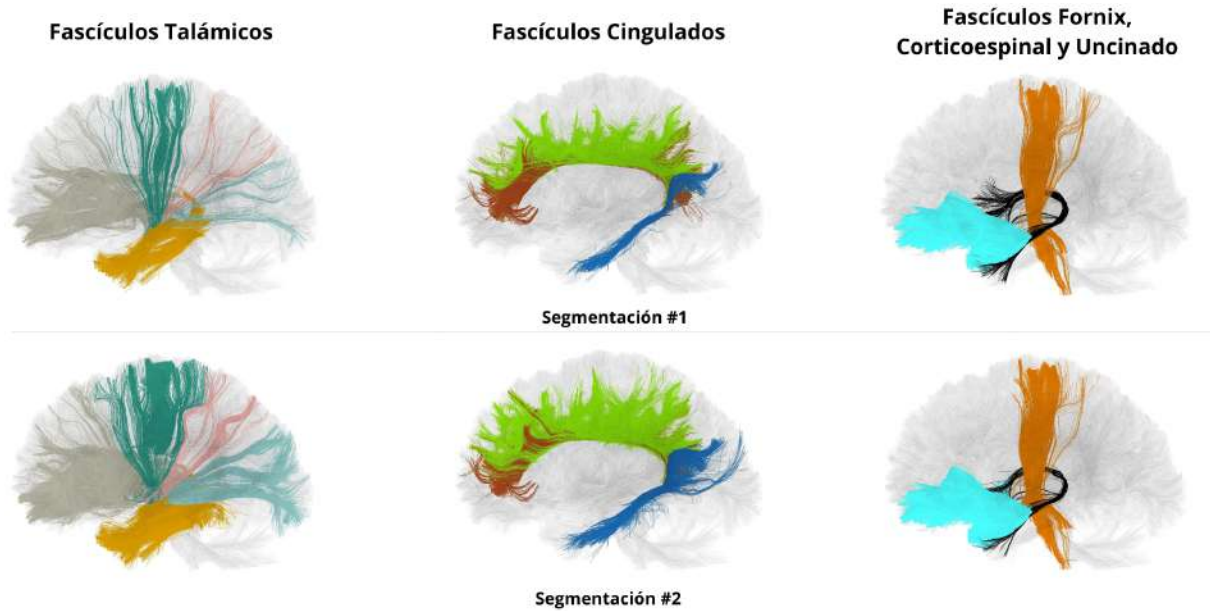
### 5.2 Métodos de segmentación

Con el objetivo de comprender los resultados y las diferencias entre los métodos de segmentación, se realizó una comparación basada en la cantidad de fibras segmentadas para los mismos fascículos y sujetos, un ejemplo de esto se visualiza en el sujeto 'mb' a través de las Figuras 5.29 y 5.30.

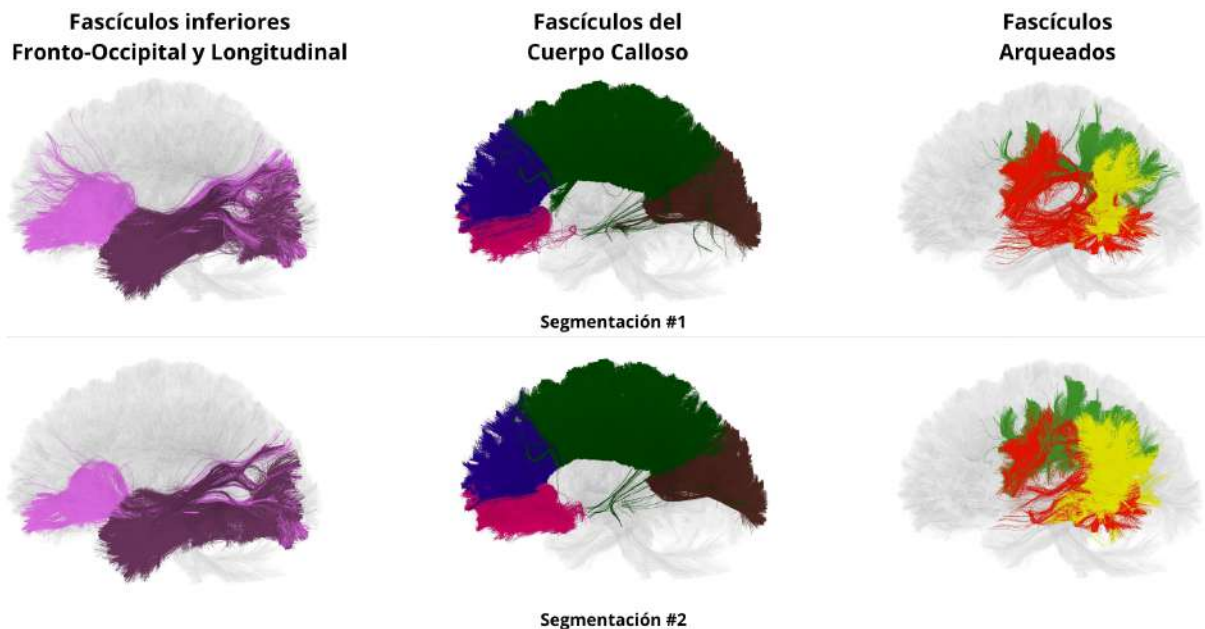
Para su análisis, se dividió el conjunto de 36 fascículos, en 3 grupos de 12 cada uno. El Grupo 1 está conformado por fascículos con un promedio de fibras entre los sujetos inferior a 7 mil, ver Figura 5.31, el Grupo 2 abarca aquellos cuyos valores alcanzan hasta 15 mil fibras aproximadamente, ver Figura 5.32, mientras que al Grupo 3 pertenecen fascículos con una cantidad de hasta 60 mil fibras, ver Figura 5.33.

En el presente estudio, la cantidad de fibras por fascículo no se considera una métrica para predecir el diagnóstico de autismo, por lo que no se distingue entre controles y pacientes en el análisis de gráficos de caja para inferir la dispersión y forma de los datos obtenidos por cada modelo.

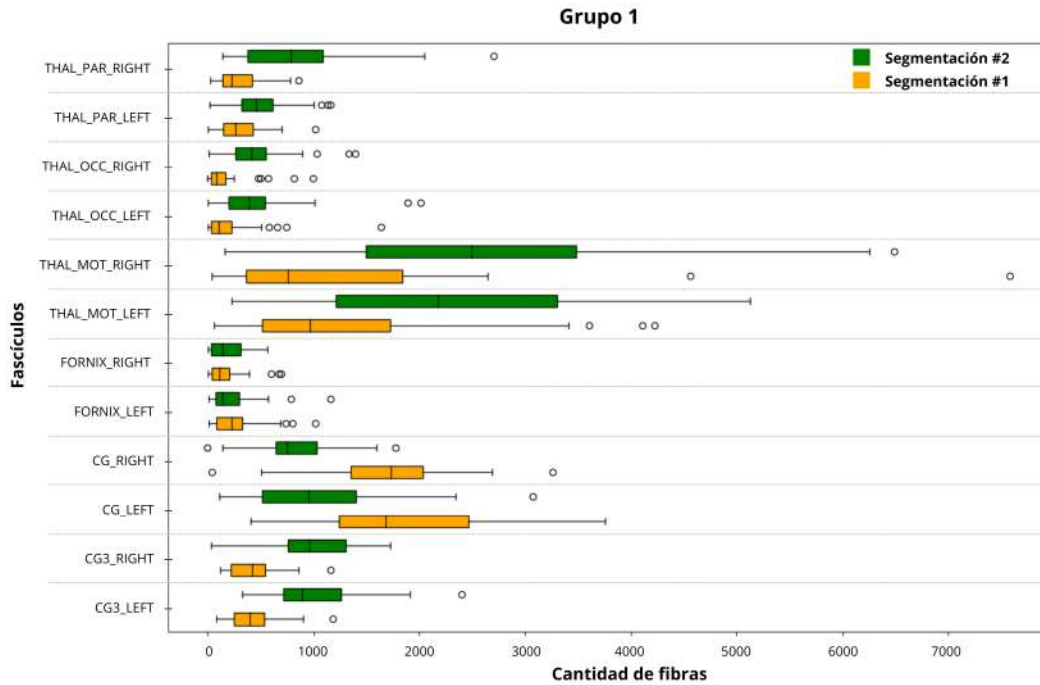
Finalmente, a pesar de que los dos métodos de segmentación utilizan metodologías y presentan desempeños diferentes, ambos tuvieron dificultades para segmentar en su totalidad los 36 fascículos indicados en el atlas para el sujeto 'tm'. Específicamente, para este sujeto, el primer método de segmentación no pudo encontrar fibras para el fascículo Talámico Occipital Derecho, mientras que el segundo método no asignó fibras al fascículo Cingulado de Fibras Largas Derecho.



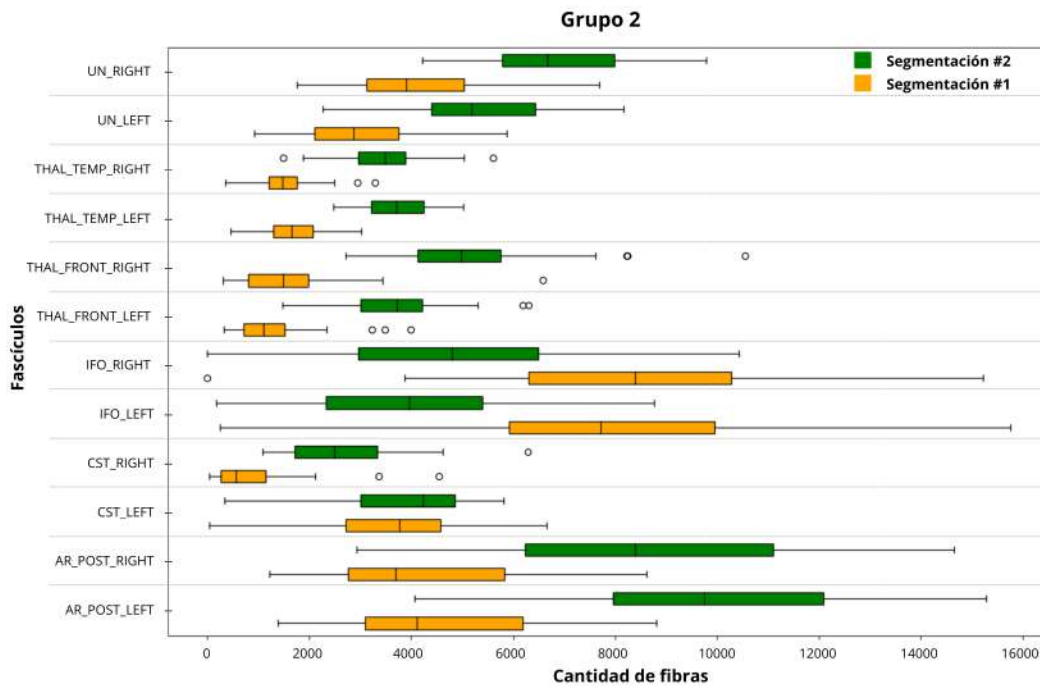
**Fig. 5.29:** Fascículos segmentados del sujeto mb100243 de la base de datos. Se aprecian las diferencias de cantidad de fibras en los fascículos Talámicos y Cingulados izquierdo, además del Fornix, Corticoespinal y Uncinado izquierdos. Para reconocer apropiadamente cada fascículo, observar Figura 3.13.



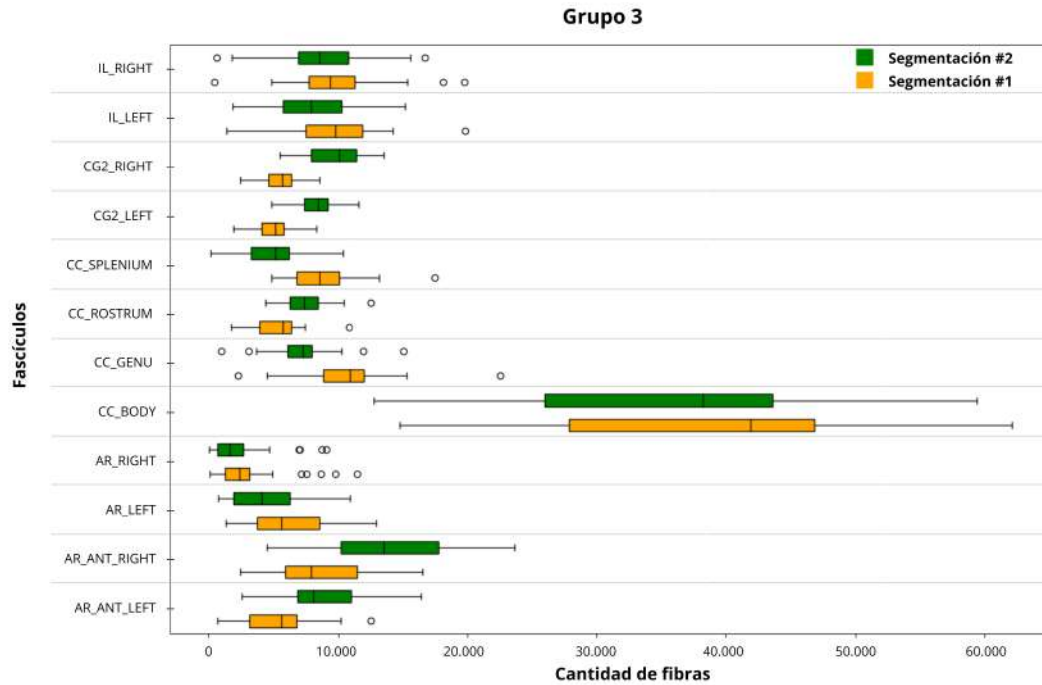
**Fig. 5.30:** Fascículos segmentados del sujeto mb100243 de la base de datos. Se aprecian las diferencias de cantidad de fibras en los fascículos del Cuerpo Calloso, Arqueados izquierdos, y los fascículos inferiores longitudinal y fronto-occipital izquierdos. Para reconocer apropiadamente cada fascículo, ver Figura 3.13.



**Fig. 5.31:** Cantidad de fibras por fascículo para el Grupo 1. El color amarillo corresponde a los valores obtenidos por el método de segmentación 'S1', mientras que el color verde representa los obtenidos por el método 'S2'.



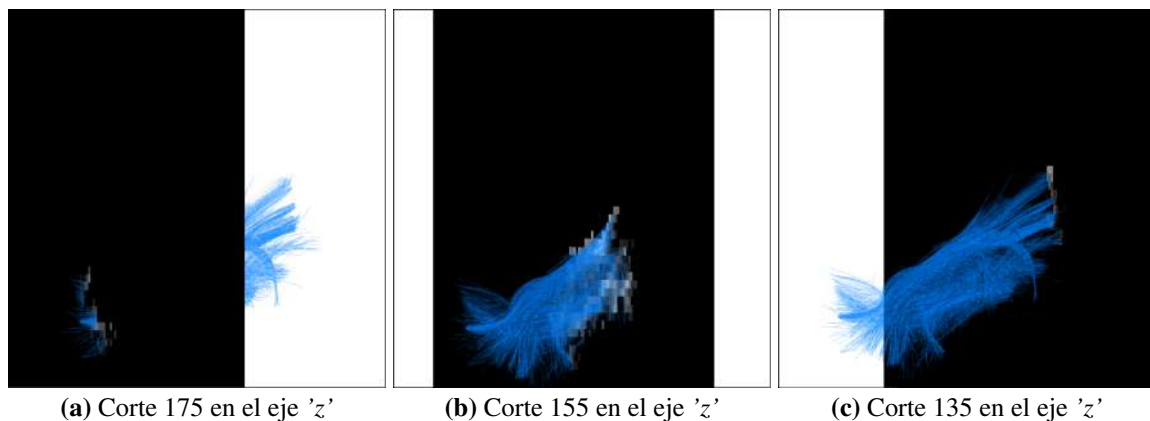
**Fig. 5.32:** Cantidad de fibras por fascículo para el Grupo 2. El color amarillo corresponde a los valores obtenidos por el método de segmentación 'S1', mientras que el color verde representa los obtenidos por el método 'S2'.



**Fig. 5.33:** Cantidad de fibras por fascículo para el Grupo 3. El color amarillo corresponde a los valores obtenidos por el método de segmentación 'S1', mientras que el color verde representa los obtenidos por el método 'S2'.

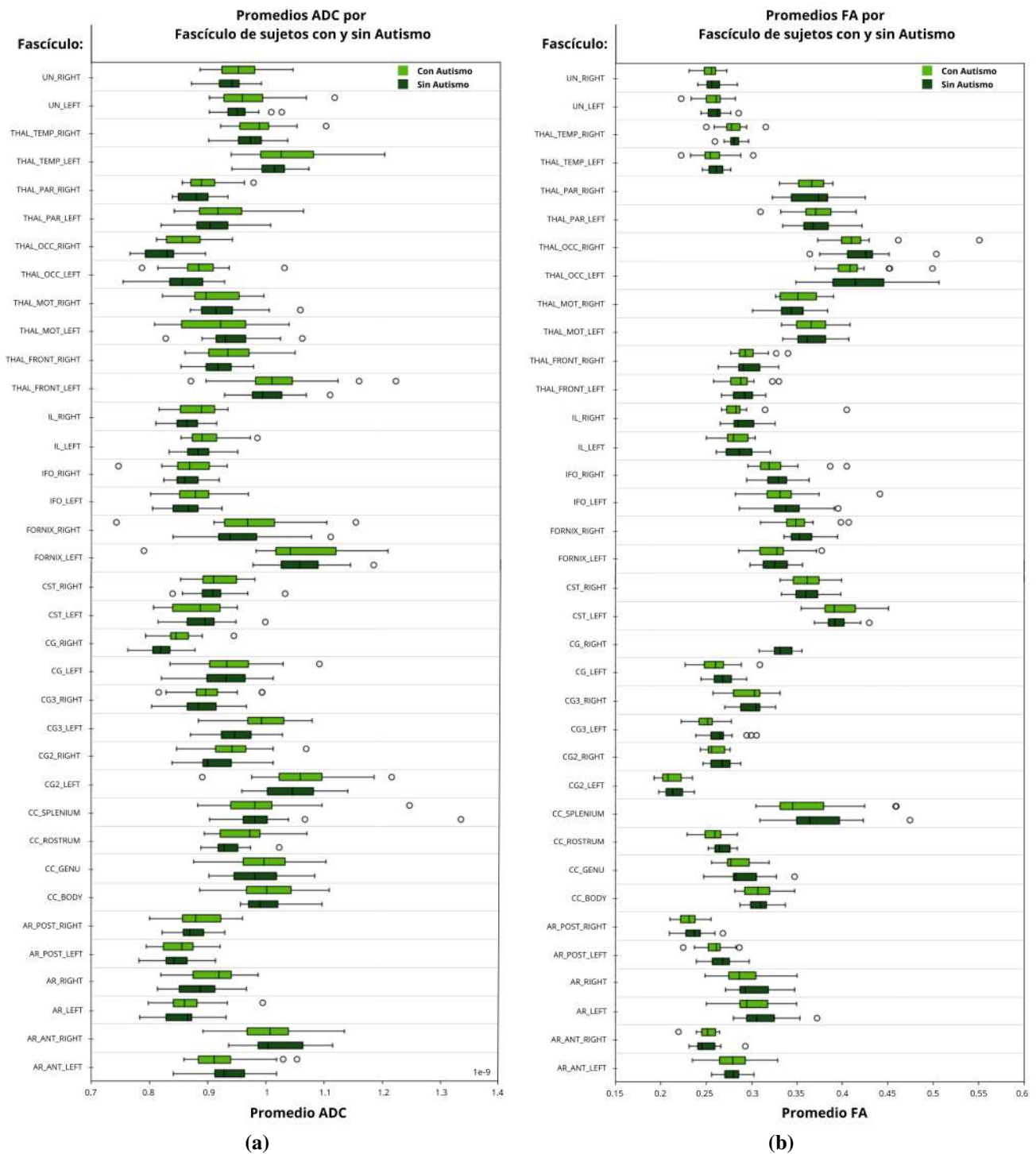
### 5.3 Aplicación de máscaras

Para verificar el funcionamiento de la aplicación de los fascículos como máscaras en los mapas de intensidad de las medidas ADC, FA y GFA, se convirtieron las matrices obtenidas en imágenes NIFTI sin alterar la orientación y dirección originales de las dMRI. Un ejemplo de esto se aprecia en la Figura 5.34.

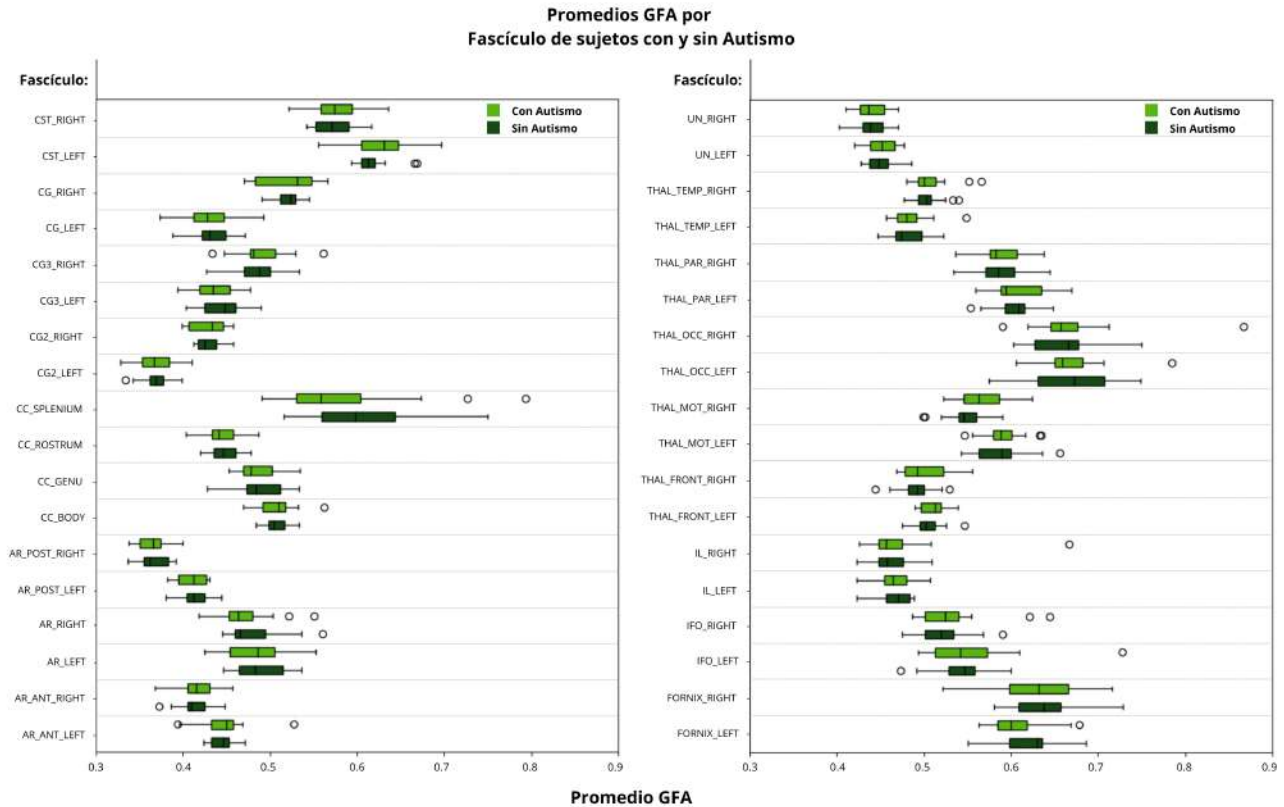


**Fig. 5.34:** Vista lateral izquierda del plano sagital del fascículo Talámico Motor izquierdo del sujeto 'mb' con opacidad al 30 % superpuesto en el resultado de haberlo aplicado como máscara para FA. En estos se pueden apreciar los vóxeles filtrados por los que pasa el fascículo.

## 5.4 Promedios de difusividad y anisotropía



**Fig. 5.35:** Valores promedio de (a) ADC y (b) FA por fascículo, distinguiendo entre sujetos con y sin autismo. Las mediciones corresponden a los fascículos obtenidos mediante el método de segmentación 'S2'.



**Fig. 5.36:** Valores promedio GFA por fascículo, distinguiendo entre sujetos con y sin autismo. Las mediciones corresponden a los fascículos obtenidos mediante el método de segmentación 'S2'.

## 5.5 Pruebas estadísticas

De las pruebas descriptivas, se observa que un 83 % de los fascículos segmentados con el método S1 tienen una distribución normal en la medida ADC, en comparación con el 86 % de los segmentados con el método S2. En el caso de la medida FA, esta proporción es de 83 %/78 %, mientras que para la medida GFA es de 67 %/61 %. Respectivamente, ver Tablas A.5, A.6 y A.7. Por otro lado, se observa que estas proporciones en cuanto a las muestras con varianzas homogéneas son del 83 %/83 % para ADC, 94 %/92 % para FA, y 81 %/83 % para la medida GFA, ver Tabla A.8.

Como resultado, se sabe que las proporciones entre los fascículos segmentados con ambos métodos que cumplen los supuestos para aplicar pruebas paramétricas en la medida ADC son de 25 fascículos, versus 26, para la medida FA, 29 versus 27, y en el caso de la medida GFA, 21 versus 22. Ver Tabla A.9.

Conociendo qué fascículos no cumplen con los supuestos en cada medida, se aplicaron las pruebas de estadística inferencial correspondientes. Los resultados se presentan en la Tabla 5.2.



Fascículo	P-VALUE											
	ADC				FA				GFA			
	S1		S2		S1		S2		S1		S2	
	T-TEST	MWU	T-TEST	MWU	T-Test	MWU	T-Test	MWU	T-Test	MWU	T-Test	MWU
AR_ANT_LEFT	0,994	0,988	0,599	0,338	0,397	0,616	0,977	0,915	0,616	0,727	-	0,494
AR_ANT_RIGHT	0,652	0,843	0,467	0,553	0,887	0,659	-	0,494	0,435	0,37	0,538	0,457
AR_LEFT	-	0,773	-	0,494	-	0,533	0,241	0,254	0,92	0,616	0,7	0,595
AR_RIGHT	0,244	0,338	0,118	0,118	0,19	0,186	0,122	0,167	-	0,186	-	0,338
AR_POST_LEFT	0,729	0,659	0,403	0,514	0,447	0,616	0,16	0,242	0,722	0,867	-	0,773
AR_POST_RIGHT	0,74	0,727	-	0,574	0,608	0,475	0,199	0,197	-	0,75	0,992	0,964
CC_BODY	0,96	0,75	0,979	0,82	0,558	0,595	0,771	0,727	0,84	0,843	0,892	0,988
CC_GENU	0,599	0,475	0,521	0,457	0,569	0,421	0,442	0,323	0,757	0,988	0,98	0,773
<b>CC_ROSTRUM</b>	-	0,218	<b>0,046</b>	<b>0,062</b>	<b>0,095</b>	<b>0,092</b>	<b>0,025</b>	<b>0,04</b>	-	0,37	0,67	0,682
CC_SPLENIUM	0,626	0,75	-	0,82	0,201	0,111	-	0,254	-	0,167	-	0,176
CG2_LEFT	0,511	0,457	0,504	0,514	0,454	0,553	0,377	0,475	-	0,75	0,999	0,82
<b>CG2_RIGHT</b>	-	<b>0,098</b>	0,113	<b>0,075</b>	<b>0,037</b>	<b>0,054</b>	<b>0,093</b>	0,141	-	0,457	-	0,891
<b>CG3_LEFT</b>	0,119	<b>0,098</b>	<b>0,028</b>	<b>0,03</b>	0,168	0,254	<b>0,007</b>	<b>0,011</b>	-	0,533	0,229	0,294
CG3_RIGHT	0,322	0,403	0,416	0,386	0,216	0,218	0,505	0,494	0,771	0,82	0,715	0,915
CG_LEFT	0,43	0,796	0,596	0,82	<b>0,099</b>	0,125	0,17	<b>0,098</b>	-	0,796	0,661	0,553
<b>CG_RIGHT</b>	-	<b>0,04</b>	<b>0,004</b>	<b>0,004</b>	-	<b>0,066</b>	<b>0,081</b>	<b>0,091</b>	-	0,82	-	0,669
CST_LEFT	-	0,595	0,508	0,704	0,941	0,915	0,374	0,616	-	0,891	-	0,309
CST_RIGHT	0,621	0,421	-	0,773	0,739	0,616	0,945	0,867	0,783	0,988	0,541	0,616
FORNIX_LEFT	-	0,533	-	0,82	0,838	0,682	0,868	0,704	0,853	0,438	0,37	0,294
FORNIX_RIGHT	0,443	0,281	0,521	0,338	0,429	0,403	0,255	0,338	0,821	0,75	0,648	0,75
IFO_LEFT	0,403	0,438	0,173	0,176	-	1	-	0,457	-	0,403	-	0,514
IFO_RIGHT	-	0,386	-	0,354	-	0,638	-	0,457	-	0,553	-	0,891
IL_LEFT	-	0,457	0,259	0,323	0,785	0,988	0,261	0,457	-	0,533	0,761	0,659
IL_RIGHT	0,187	0,197	0,128	0,141	-	0,37	-	0,218	-	0,659	-	0,843
THAL_FRONT_LEFT	0,649	0,638	-	0,403	0,828	0,659	0,682	0,438	0,311	0,309	0,28	0,37
THAL_FRONT_RIGHT	0,212	0,338	0,253	0,354	0,89	0,796	-	0,891	0,129	0,118	0,144	0,354
THAL_MOT_LEFT	0,146	0,218	0,295	0,494	0,406	0,638	0,835	0,82	0,344	0,254	0,721	0,727
<b>THAL_MOT_RIGHT</b>	0,504	0,616	-	0,533	0,413	0,37	0,159	0,281	0,175	0,254	<b>0,047</b>	<b>0,066</b>
THAL_OCC_LEFT	0,824	0,796	0,117	0,176	0,598	0,438	-	0,574	0,384	0,37	-	0,796
<b>THAL_OCC_RIGHT</b>	<b>0,05</b>	<b>0,038</b>	<b>0,007</b>	<b>0,011</b>	0,524	0,558	-	0,167	0,59	0,764	-	0,964
THAL_PAR_LEFT	-	0,891	0,223	0,254	0,995	0,796	0,817	0,82	-	0,574	0,597	0,988
THAL_PAR_RIGHT	0,461	0,638	0,135	0,23	-	0,75	0,582	0,574	-	0,574	0,674	0,843
THAL_TEMP_LEFT	-	0,294	-	0,267	-	0,659	-	0,23	0,807	0,796	-	0,891
THAL_TEMP_RIGHT	-	0,281	0,309	0,438	0,19	0,167	0,726	0,682	0,556	0,796	-	0,796
UN_LEFT	0,43	0,574	-	0,595	0,617	0,553	0,529	0,638	0,684	0,867	0,822	0,75
UN_RIGHT	0,258	0,37	0,159	0,421	0,239	0,386	0,212	0,421	0,904	0,988	0,999	0,964

**Tabla 5.2:** Comparación estadística de las medidas ADC, FA y GFA entre sujetos con y sin autismo utilizando las pruebas T-Test de muestras independientes y Mann Whitney U, basada en los valores obtenidos a partir de los métodos de segmentación 'S1' y 'S2'. Considerar que, '-' marca los fascículos no aptos para aplicar la prueba paramétrica T-Test. El color azul representa los p-values menores al nivel de significancia 0.05, y por ende, indica los fascículos en los que se encontró una diferencia estadísticamente significativa. Mientras que el color celeste, representa valores cercanos, pero no menores al nivel de significancia.



## 5.6 Clasificación por Aprendizaje automático

	ADC		FA		GFA		A-F		A-G		F-G		A-F-G	
	DT	SVM	DT	SVM	DT	SVM	DT	SVM	DT	SVM	DT	SVM	DT	SVM
AR_ANT_LEFT	-	-	-	<b>x</b>	<b>x</b>	-	-	-	<b>x</b>	-	-	-	<b>x</b>	-
AR_ANT_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AR_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AR_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AR_POST_LEFT	-	-	-	-	-	<b>x</b>	-	-	-	-	-	-	-	-
AR_POST_RIGHT	-	-	-	-	<b>x</b>	-	-	-	<b>x</b>	-	-	-	-	-
CC_BODY	-	-	-	-	-	-	-	-	-	-	-	-	<b>x</b>	-
CC_GENU	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CC_ROSTRUM	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CC_SPLENIUM	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CG2_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CG2_RIGHT	-	-	-	-	-	<b>x</b>	-	-	-	-	<b>x</b>	-	-	-
CG3_LEFT	-	<b>x</b>	-	-	-	-	-	-	-	-	-	-	-	-
CG3_RIGHT	-	-	-	-	<b>x</b>	-	-	-	<b>x</b>	-	-	-	-	-
CG_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CG_RIGHT	-	<b>x</b>	-	<b>x</b>	<b>x</b>	-	-	-	<b>x</b>	-	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
CST_LEFT	-	-	<b>x</b>	-	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
CST_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FORNIX_LEFT	-	-	-	-	-	-	-	-	-	<b>x</b>	-	-	-	-
FORNIX_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IFO_LEFT	-	-	-	-	<b>x</b>	-	-	<b>x</b>	<b>x</b>	-	-	-	-	-
IFO_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IL_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IL_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_FRONT_LEFT	-	-	<b>x</b>	-	-	-	<b>x</b>	-	-	-	-	-	-	-
THAL_FRONT_RIGHT	-	-	-	-	<b>x</b>	-	-	-	<b>x</b>	-	-	<b>x</b>	-	<b>x</b>
THAL_MOT_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_MOT_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_OCC_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_OCC_RIGHT	-	-	-	<b>x</b>	-	-	-	-	-	-	-	-	-	-
THAL_PAR_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_PAR_RIGHT	-	<b>x</b>	-	-	-	-	-	-	-	-	-	-	-	-
THAL_TEMP_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
THAL_TEMP_RIGHT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
UN_LEFT	-	-	-	-	-	-	-	-	-	-	-	-	-	-
UN_RIGHT	-	-	<b>x</b>	-	<b>x</b>	-	<b>x</b>	-	<b>x</b>	-	<b>x</b>	-	-	-

**Tabla 5.3:** Resultados de los modelos de clasificación utilizando DT y SVM, con ADC (A), FA (F) y GFA (G), así como sus combinaciones, en cada fascículo por separado. Los fascículos que no lograron una precisión mayor al 66 % se muestran en gris. Aquellos fascículos que permitieron clasificar correctamente a los sujetos por algún modelo con una precisión superior al 66 % se marcan con una 'x', y los que en un modelo, con cierta medida o combinación, obtuvo un promedio de métricas de evaluación cercano al 80 % se resaltan en rojo.

Si bien la Tabla 5.3 es un resumen de los resultados de la aplicación de modelos de clasificación, se pueden encontrar detalles más específicos sobre las métricas de evaluación, ya sea *'accuracy'*, *'precision'*, *'recall'* y/o *'F1 Score'*, en el Anexo, específicamente en la Tabla A.10 para la medida ADC, Tabla A.11 para FA, Tabla A.12 para GFA, Tabla A.13 para la combinación de ADC y FA, Tabla A.14 para la combinación de ADC y GFA, Tabla A.15 para la combinación de FA y GFA, y Tabla A.16 para la combinación de las tres medidas.

Mientras que los mejores hiperparámetros calculados para cada prueba realizada en los fascículos destacados de la Tabla 5.3, se aprecian en la Tabla A.17 para las pruebas con DT, y en la Tabla A.18 para las pruebas de clasificación realizadas con SVM.

## 5.7 Discusión

En cuanto a la segmentación de los fascículos, no parece haber una relación directa entre la falta de fascículos segmentados en el sujeto 'tm' y la cantidad de fibras que posee, mucho menos su diagnóstico. Es decir, 'tm' no es el sujeto con menos o más fibras, ya que posee aproximadamente 778 mil fibras. Aunque esta cifra está por debajo del promedio de 924 mil, hay sujetos con menos de 700 mil fibras que se segmentaron sin problemas.

Con respecto al rendimiento entre los dos métodos, no se puede afirmar que haya una característica específica causante de la diferencia entre 'S1' y 'S2'. Sin embargo, en general el método 'S2' segmenta un mayor porcentaje de fibras de manera adecuada en comparación con el primer método. De los 37 sujetos, 'S2' logra una mejor segmentación en 33 de ellos. Para más detalles, consultar las Tablas A.1, A.2, A.3, y A.4, recordando que en algunos casos, un 1 % puede equivaler hasta a 10 mil fibras.

Los fascículos que muestran una mayor diferencia en la cantidad de fibras segmentadas por 'S2' son los Talámicos, seguidos por los Arqueados Posteriores, Cingulados de fibras temporales, Cingulados de fibras cortas, Uncinados y Corticoespinales. En cambio, los fascículos que presentan resultados similares en ambos métodos son los Fornix, los Inferiores Fronto-Occipitales y los de Cuerpos Callosos. Por otro lado, los Cingulados de fibras largas y los Inferiores Longitudinales son los fascículos que se segmentan en menor medida. Cabe destacar que el método 'S1' muestra una clara diferencia entre los fascículos Corticoespinales izquierdo y derecho, como se puede apreciar en el gráfico de la Figura 5.32, mientras que el método 'S2' soluciona esta deficiencia.

Gracias al gráfico mencionado anteriormente, y los gráficos de las Figuras 5.31 y 5.33, podemos observar la distribución de los valores de cada fascículo segmentado por ambos métodos en los 37 sujetos. Los valores atípicos pueden deberse a aquellos sujetos con una mayor proporción de fibras segmentadas, y en la mayoría de los casos, también son aquellos sujetos con mayor cantidad de fibras en la tractografía.

Posteriormente, se utilizaron los valores obtenidos de ADC, FA y GFA para graficar los respectivos resultados de cada fascículo obtenido mediante la segmentación 'S2'. Esto se debe a la similitud, o mejora de esta, en la distribución de valores y varianzas de la cantidad de fibras, lo cual hace que esta representación sea más significativa al contar con más datos. A través de las gráficas de las Figuras 5.35 y 5.36, se puede observar qué fascículos presentan valores más altos o más bajos en las diferentes medidas.

Un valor de ADC más bajo implica una menor difusión de agua, y por ende, una mayor restricción del movimiento de las moléculas de agua, mientras que uno más alto implica una mayor difusión de agua, y una menor densidad de fibras o una disminución en la integridad estructural del tejido, lo cual puede ser relevante en algunos casos, como en la detección de daño o patología.

En el caso de las medidas de FA, un valor alto, cercano a 1, implica una mayor organización y coherencia direccional de la difusión del agua, mientras que un valor más bajo implica una difusión más isotrópica y una menor organización direccional en ese tejido, además de sugerir una menor organización microestructural, o la presencia de cambios patológicos en el tejido.

Así mismo, un valor alto de GFA significa una mayor coherencia y una mayor direccionalidad en la difusión del agua, lo que tiende a ser positivo, mientras que un valor menor, implica una menor coherencia y una menor direccionalidad en la difusión del agua, lo cual puede indicar cambios patológicos.

En relación con las pruebas estadísticas, se aplicaron a ambos métodos para determinar cuál proporciona información más relevante. Sin embargo, se estima que los resultados de la segmentación 'S2' serán los que permitan revelar correctamente esta predicción en esta etapa de análisis.

De las Tablas A.5, A.6 y A.7, se puede apreciar sutilmente que los sujetos con autismo presentan una mayor tasa de anormalidad de datos en comparación con los sujetos sin autismo. Luego, al comparar los resultados correspondientes a los fascículos segmentados en ambos métodos, se observa que para ADC, 23 de los 36 fascículos comparten resultados de pruebas descriptivas, mientras que para FA son 34 y para GFA son 25. Es importante destacar que los fascículos que comparten su resultado de normalidad en diferentes medidas no son siempre los mismos. Esto significa que si un fascículo cumple con la normalidad y la homogeneidad en ADC, no implica necesariamente que también lo cumpla en FA o GFA.

En cuanto a los resultados de las pruebas inferenciales, en la mayoría de los casos en los que se encontró una diferencia estadísticamente significativa para algún fascículo obtenido con 'S1', esa diferencia se incrementó aún más con los valores obtenidos para los fascículos de 'S2'. Además, se encontraron más fascículos relevantes en el segundo método.

Mientras que los obtenidos por los modelos de clasificación de Machine Learning, a pesar de que se esperaba que no fueran concluyentes debido al tamaño de la muestra disponible, ya que los 37 sujetos parecían ser insuficientes. Mediante la adaptación de los porcentajes de prueba y entrenamiento, así como la realización de combinaciones de métricas, se lograron obtener algunos resultados interesantes.

Según los resultados resumidos en la tabla 5.3, no todos los fascículos significativos proporcionan una buena clasificación para ambos modelos. En cambio, se observa que el modelo SVM es el que muestra mayor precisión en la predicción para ADC, mientras que el modelo DT muestra mejores resultados para GFA. En el caso de FA, la proporción de resultados es similar para ambos modelos.

En relación con las combinaciones de medidas, la combinación de FA y ADC es la menos concluyente en términos de clasificación. Sin embargo, se destaca que la combinación de ADC con GFA arrojó los mejores resultados en todo el análisis realizado, obteniendo un porcentaje de '*accuracy*' mayor al 66 % en al menos uno de los modelos aplicados en 9 fascículos. Se destacan de ellos el fascículo Fornix izquierdo que obtuvo un '*accuracy*' de 73 % con el modelo SVM, y el fascículo Corticoespinal izquierdo que respectivamente obtuvo un 73 % con SVM y un 80 % con el modelo DT.

Cabe mencionar que no mostrar un buen rendimiento en ambos modelos de clasificación simultáneamente no indica un error, sino más bien sugiere que esa muestra de datos es más adecuada para ser analizada por un método en particular. Además, se observa que los modelos basados en árboles de decisiones son los que presentan una menor variabilidad en sus hiperparámetros, como se puede apreciar entre las Tablas A.17 y A.18.

Finalmente, es destacable que, sorprendentemente, a pesar de que la medida GFA no mostró diferencias significativas en las pruebas estadísticas, en las pruebas de clasificación mediante Machine Learning, ofreció indicios de ser una métrica más efectiva para la clasificación.

## Capítulo 6. Conclusiones

---

### 6.1 Discusión

La presente investigación implicó varios desafíos, incluyendo la necesidad de abordar múltiples áreas interdisciplinarias, como estadística, álgebra, imágenes médicas, neurociencia y reconocimiento de patrones, entre otras. Este enfoque multidisciplinario fue fundamental para lograr el objetivo propuesto.

Durante el desarrollo, se encontraron contratiempos, que son parte natural del proceso de aprendizaje. Uno de ellos fue lidiar con casos no óptimos al aplicar diferentes algoritmos, pero se logró superar cada desafío mediante la adaptación, la resolución de problemas y el razonamiento ingenieril.

Factores a considerar, son la importancia de contar con los recursos apropiados y estar en un entorno ideal. Tener acceso a un buen computador puede marcar la diferencia en el tiempo de procesamiento, la ejecución de los algoritmos y, en última instancia, los resultados obtenidos. Además, mantener una comunicación efectiva con otros estudiantes en un entorno de estudio común resulta beneficioso para obtener una comprensión más amplia y enriquecer los puntos de vista al abordar problemáticas específicas.

Por otro lado, en cuanto a los valores obtenidos, al comparar los resultados de las pruebas estadísticas y la clasificación por Machine Learning, se observa cierta similitud y diferencia en diferentes casos.

En el caso del fascículo Cingulado de fibras cortas derecho, las medidas ADC y FA demostraron diferencias estadísticas, mientras que los modelos de clasificación mostraron un buen rendimiento en las medidas evaluadas de GFA y FA-GFA. Para el fascículo Talámico Occipital derecho, las pruebas estadísticas destacaron en ADC, mientras que la clasificación por ML fue más efectiva en FA. Por el contrario, el fascículo Cingulado de Fibras temporales derecho, encontró una buena diferencia estadísticamente significativa en FA y ADC, y el modelo de clasificación obtuvo un buen rendimiento en ADC con SVM.

En cuanto al fascículo Cingulado de Fibras largas derecho, se encontraron diferencias estadísticas en ADC y se acercó a una diferencia significativa en FA, pero los modelos de clasificación mostraron un buen rendimiento en ADC, FA, GFA, ADC con GFA, en las tres métricas juntas y en FA con GFA, que obtuvieron el mejor rendimiento donde en promedio de sus métricas de evaluación, superaban el 75%. Curiosamente, el fascículo corticoespinal izquierdo parece tener un mejor rendimiento en la clasificación por ML, a pesar de no mostrar diferencias estadísticas previamente.

Se reconoce que contar con una base de datos más amplia permitiría obtener conclusiones más sólidas. No se debe olvidar que el espectro autista es muy variado y cada sujeto presenta comportamientos característicos dentro de su grado específico. Aunque en este estudio no se ha explorado esta información, su consideración podría proporcionar valiosos indicios y mejorar la comprensión de los datos disponibles.

Es importante tener en cuenta que las diferencias observadas en las pruebas mencionadas anteriormente podrían atribuirse en el peor de los casos al azar. Sin embargo, se destaca que los fascículos Cingulados y Corticoespinal han mostrado valores destacados, al igual que algunos Talámicos, como se mencionó previamente, lo cual coincide con los hallazgos de otros estudios publicados mencionados en el Marco Teórico.

## 6.2 Conclusiones

Con éxito, se lograron cumplir todos los objetivos inicialmente propuestos. Se desarrolló y aplicó un pipeline organizado y documentado para el análisis de fibras cerebrales mediante adaptaciones en los scripts existentes de procesamiento y la creación de nuevos scripts para funciones faltantes.

Específicamente, se implementaron algoritmos para alinear los datos de tractografía en el espacio Talairach, remuestrear sus fibras y segmentar los fascículos basados en un atlas. Además, se obtuvieron medidas promedio de FA, GFA y ADC para los fascículos de cada sujeto. Mediante la aplicación de tests estadísticos y algoritmos de clasificación, se identificaron alteraciones en los fascículos de la materia blanca en pacientes con autismo, logrando en algunos casos predecir el diagnóstico utilizando algoritmos de Machine Learning, con una precisión mayor al 70 %.

Al evaluar los resultados, se obtuvieron hallazgos significativos sobre las características distintivas de ciertos fascículos entre personas con y sin autismo. Aunque los resultados podrían ser aún más concluyentes con una base de datos más amplia y específica, este estudio respalda la importancia de prestar una mayor atención en los fascículos cingulados, talámicos y corticoespinales.

### 6.3 Trabajos Futuros

A pesar de los avances logrados en este estudio sobre el análisis de la materia blanca en sujetos con autismo y sin autismo, existen varias propuestas que pueden ser exploradas en futuras investigaciones.

En primer lugar, se sugiere aplicar más modelos de clasificación de Machine Learning en todos los fascículos, y utilizar validación cruzada, esta vez, para obtener un promedio basado en distintos `random_states` de los fascículos destacados previamente. Lo anterior permitirá evaluar y comparar diferentes resultados variando entre prueba y entrenamiento en los mismos y nuevos modelos de clasificación.

Seguido a esto, cubriendo uno de los aspectos más importantes para obtener resultados más robustos y generalizables, se propone ampliar el tamaño de la muestra de sujetos de estudio. Ya que esto proporcionará una mayor diversidad de datos y permitirá obtener conclusiones más sólidas sobre la relación entre los cambios en la materia blanca y el autismo.

Se propone trabajar con información más detallada sobre los sujetos de investigación y su nivel de autismo en el espectro. El nivel de autismo está estrechamente relacionado con el comportamiento del sujeto, y se cree que podría estar asociado con características específicas en ciertas áreas del cerebro. Analizar la relación entre estos comportamientos y los cambios en los fascículos o áreas de la materia blanca brindaría una perspectiva más completa.

Además, se recomienda colaborar con profesionales médicos especializados en autismo. Su experiencia clínica y conocimientos complementarios pueden aportar una segunda interpretación de los hallazgos y garantizar una comprensión más precisa de los resultados.

Finalmente, si bien los datos de tractografía han sido una herramienta valiosa en este estudio, explorar otras técnicas de imagenología, como la resonancia magnética funcional (fMRI) o la resonancia magnética estructural (sMRI), proporcionarían una visión más completa del impacto de los cambios en la materia blanca, permitiendo un análisis más extenso de las diferencias entre las submuestras de sujetos con y sin autismo.

## Capítulo 5. Glosario

AD	Difusividad Axial , en inglés, Axial diffusivity
ADC	Coefficiente de Difusión Aparente , en inglés, Apparent Diffusion Coefficient
dMRI	Imágenes de Resonancia Magnética por Difusión , en inglés, Diffusion Magnetic Resonance Imaging
DT	Tensor de Difusión , en inglés, Diffusion Tensor
DTI	Imágenes de Tensor de Difusión , en inglés, Diffusion Tensor Imaging
DWI	Imágenes Ponderadas por Difusión , en inglés, Diffusion Weighted Imaging
FA	Anisotropía Fraccional , en inglés, Fractional Anisotropy
GFA	Anisotropía Fraccional Generalizada , en inglés, Generalized Fractional Anisotropy
GM	Sustancia o Materia Gris , en inglés, Gray Matter
HARDI	Imágenes de Difusión de Alta Resolución Angular , en inglés, High Angular Resolution Diffusion Imaging
MD	Difusividad Media , en inglés, Medium diffusivity
ML	Aprendizaje Automático , en inglés, Machine Learning
MNI	Instituto Neurológico de Montreal , en inglés, Montreal Neurological Institute
MRI	Imágenes de Resonancia Magnética , en inglés, Magnetic Resonance Imaging
ODF	Función de Distribución de Orientación , en inglés, Orientation Distribution Function
QBI	Imágenes Q-Ball , en inglés, Q-Ball Imaging
RD	Difusividad Radial , en inglés, Radial diffusivity
RM	Resonancia Magnética , en inglés, Magnetic Resonance
SNC	Sistema Nervioso Central , en inglés, Central Nervous System
SVM	Máquinas de Vectores de Soporte , en inglés, Support Vector Machine
TBSS	Estadísticas Espaciales Basadas en Tractos , en inglés, Tract Based Spatial Statistics
TEA	Trastorno del Espectro Autista , en inglés, Autism Spectrum Disorder
WM	Sustancia o Materia Blanca , en inglés, White Matter



## Capítulo 6. Referencias

---

- [1] L. Thau, V. Reddy, and P. Singh, “Anatomy, central nervous system,” in *StatPearls*. Treasure Island (FL): StatPearls Publishing, Jan. 2022.
- [2] R. S. Snell, *Neuroanatomía clínica*, 6th ed. Ed. Médica Panamericana, 2007.
- [3] K. F. Swaiman, S. Ashwal, D. M. Ferriero, R. S. Finkel, A. L. Gropman, P. L. Pearl, and M. Shevell, *Swaiman’s pediatric neurology*, 6th ed. Philadelphia, PA: Elsevier - Health Sciences Division, May 2017.
- [4] S. H. Ameis and M. Catani, “Altered white matter connectivity as a neural substrate for social impairment in autism spectrum disorder,” *Cortex*, vol. 62, pp. 158–181, Jan. 2015.
- [5] H. Johansen-Berg and T. E. J. Behrens, Eds., *Diffusion MRI. From Quantitative Measurement to In vivo Neuroanatomy*, 2nd ed. San Diego, CA: Academic Press, Apr. 2014.
- [6] J. Katz, M.-A. d’Albis, J. Boisgontier, C. Poupon, J.-F. Mangin, P. Guevara, D. Duclap, N. Hamdani, J. Petit, D. Monnet, P. Le Corvoisier, M. Leboyer, R. Delorme, and J. Houenou, “Similar white matter but opposite grey matter changes in schizophrenia and high-functioning autism,” *Acta Psychiatr. Scand.*, vol. 134, no. 1, pp. 31–39, Jul. 2016.
- [7] P. Guevara, D. Duclap, C. Poupon, L. Marrakchi-Kacem, P. Fillard, D. Le Bihan, M. Leboyer, J. Houenou, and J.-F. Mangin, “Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas,” *NeuroImage*, vol. 61, no. 4, pp. 1083–1099, Jul. 2012.
- [8] D. Casanova Martínez, “Vista lateral de la distribución de la sustancia blanca (fibras) en el hemisferio derecho de un cerebro humano,” <http://www.blog.illustraciencia.info/2019/10/daniel-casanova-martinez-vista-lateral.html>, 2019.
- [9] M. Jensen. (2015) Introduction to neurology. [Online]. Available: <https://introtoneuro.com/>
- [10] S. Zauderer. (2022) 69+ autism statistics: How many people have autism? [Online]. Available: <https://www.crossrivertherapy.com/autism-statistics#:~:text=According%20to%20the%20CDC%2C%20around,States%20according%20to%20the%20CDC.>
- [11] A. Berger, “Magnetic resonance imaging,” *BMJ*, vol. 324, no. 7328, p. 35, Jan. 2002.

- [12] R. W. Brown, Y. C. Norman Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging*, 2nd ed., R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, Eds. Chichester, England: Wiley-Blackwell, Jun. 2014.
- [13] W. Herring, *Radiología Básica*, 3rd ed. Elsevier, Mar. 2016.
- [14] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from picture to proton*, 2nd ed. Cambridge, England: Cambridge University Press, Oct. 2009.
- [15] V. Baliyan, C. J. Das, R. Sharma, and A. K. Gupta, “Diffusion weighted imaging: Technique and applications,” *World J. Radiol.*, vol. 8, no. 9, pp. 785–798, Sep. 2016.
- [16] T. A. G. M. Huisman, “Diffusion-weighted and diffusion tensor imaging of the brain, made easy,” *Cancer Imaging*, vol. 10 Spec no A, no. 1A, pp. S163–71, Oct. 2010.
- [17] T. Martín Noguerol and J. P. Martínez Barbero, “RM-Difusión avanzada y biomarcadores en el sistema nervioso central: un nuevo enfoque,” *Radiologia*, vol. 59, no. 4, pp. 273–285, Jul. 2017.
- [18] S. Sunaert, W. van Hecke, and L. Emsell, Eds., *Diffusion tensor imaging*, 1st ed. New York, NY: Springer, Dec. 2015.
- [19] D. S. Tuch, “Q-ball imaging,” *Magn. Reson. Med.*, vol. 52, no. 6, pp. 1358–1372, Dec. 2004.
- [20] S. Rao, J. G. Ibrahim, J. Cheng, P.-T. Yap, and H. Zhu, “SR-HARDI: Spatially regularizing high angular resolution diffusion imaging,” *J. Comput. Graph. Stat.*, vol. 25, no. 4, pp. 1195–1211, 2016.
- [21] D. Christiaens and J. D. Tournier, “Modeling fiber orientations using diffusion MRI,” in *Advances in Magnetic Resonance Technology and Applications*. Elsevier, 2020, pp. 509–532.
- [22] E. Ordóñez Rubiano, J. Marín Muñoz, J. Hernández, M. Bello Dávila, J. Ordóñez, K. Saavedra, N. Montes Reyes, D. Angulo, and E. Ordóñez Mora, “Principios de tractografía cerebral,” *Rev. repert. med. cir.*, vol. 28, no. 1, Jan. 2019.
- [23] M. Porcu, L. Cocco, J. Puig, L. Mannelli, Q. Yang, J. S. Suri, G. Defazio, and L. Saba, “Global fractional anisotropy: Effect on resting-state neural activity and brain networking in healthy participants,” *Neuroscience*, vol. 472, pp. 103–115, Sep. 2021.
- [24] J. N. Alcaide. (2019) Sistema límbico y emociones. [Online]. Available: <https://impulsaneuropsicologia.com/sistema-limbico-y-emociones/>

- [25] N. A. Lazar, *The statistical analysis of functional MRI data*, 2008th ed., ser. Statistics for Biology and Health. New York, NY: Springer, Dec. 2008.
- [26] W. Chau and A. R. McIntosh, “The talairach coordinate of a point in the MNI space: how to interpret it,” *Neuroimage*, vol. 25, no. 2, pp. 408–416, Apr. 2005.
- [27] J. CARLIN. (2017) The mni brain and the talairach atlas. [Online]. Available: <https://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach>
- [28] K. L. Weiss, H. Pan, J. Storrs, W. Strub, J. L. Weiss, L. Jia, and O. P. Eldevik, “Clinical brain MR imaging prescriptions in talairach space: technologist- and computer-driven methods,” *AJNR Am. J. Neuroradiol.*, vol. 24, no. 5, pp. 922–929, May 2003.
- [29] J. P. Porto and M. Merino. (2013) Definición de pipeline. [Online]. Available: <https://definicion.de/pipeline/>
- [30] D. S. Andrews, J. K. Lee, D. J. Harvey, E. Waizbard-Bartov, M. Solomon, S. J. Rogers, C. W. Nordahl, and D. G. Amaral, “A longitudinal study of white matter development in relation to changes in autism severity across early childhood,” *Biol. Psychiatry*, vol. 89, no. 5, pp. 424–432, Mar. 2021.
- [31] M.-A. d’Albis, P. Guevara, M. Guevara, C. Laidi, J. Boisgontier, S. Sarrazin, D. Duclap, R. Delorme, F. Bolognani, C. Czech, C. Bouquet, M. Ly-Le Moal, S. Holiga, A. Amestoy, I. Scheid, A. Gaman, M. Leboyer, C. Poupon, J.-F. Mangin, and J. Houenou, “Local structural connectivity is associated with social cognition in autism spectrum disorder,” *Brain*, vol. 141, no. 12, pp. 3472–3481, Dec. 2018.
- [32] S.-J. Hong, B. Hyung, C. Paquola, and B. C. Bernhardt, “The superficial white matter in autism and its role in connectivity anomalies and symptom severity,” *Cereb. Cortex*, vol. 29, no. 10, pp. 4415–4425, Sep. 2019.
- [33] R. Azucas. (2023) Lóbulo temporal. [Online]. Available: <https://www.kenhub.com/es/library/anatomia-es/lobulo-temporal>
- [34] M. Laguna. (2023) Lóbulo de la ínsula. [Online]. Available: <https://www.kenhub.com/es/library/anatomia-es/lobulo-de-la-insula>
- [35] D. C. Lay, *Algebra Lineal Y Sus Aplicaciones*, 3rd ed. Pearson Educación, 2007.
- [36] H. C. Guzmán. (2022) Matrices de transformación 3d. [Online]. Available: <https://docs.hektorprofe.net/graficos-3d/16-matrices-de-transformacion-3d/>

- [37] R. Johnson and P. Kuby, *Estadística elemental. Lo esencial. 10e*, 2008.
- [38] R. Ortiz. (2017) Estadística básica agronómica. [Online]. Available: [https://rstudio-pubs-static.s3.amazonaws.com/332739\\_9407da5ce6a147b0bd4898775a494c26.html](https://rstudio-pubs-static.s3.amazonaws.com/332739_9407da5ce6a147b0bd4898775a494c26.html)
- [39] J. C. Correa, R. Iral, and L. Rojas, “Estudio de potencia de pruebas de homogeneidad de varianza,” *Revista Colombiana de Estadística*, vol. 29, no. 1, p. 57–76, Jan. 2006. [Online]. Available: <https://revistas.unal.edu.co/index.php/estad/article/view/29169>
- [40] R. B. Mayorga Ponce, D. C. Graciano Ventura, A. Martínez Hernández, P. M. Moctezuma Jiménez, B. Pérez Galindo, and A. Roldan Carpio, “Cuadro comparativo de análisis paramétrico y no paramétrico,” *Educación y Salud Boletín Científico Instituto de Ciencias de la Salud Universidad Autónoma del Estado de Hidalgo*, vol. 10, no. 20, pp. 90–93, jun. 2022. [Online]. Available: <https://repository.uaeh.edu.mx/revistas/index.php/ICSA/article/view/9143>
- [41] J. Rodrigo. (2016) T-test. [Online]. Available: [https://www.cienciadedatos.net/documentos/12\\_t-test](https://www.cienciadedatos.net/documentos/12_t-test)
- [42] J. A. Rodrigo. (2017) Test de wilcoxon mann whitney como alternativa al t-test. [Online]. Available: [https://www.cienciadedatos.net/documentos/17\\_mann%E2%80%93whitney\\_u\\_test](https://www.cienciadedatos.net/documentos/17_mann%E2%80%93whitney_u_test)
- [43] A. Geron, *Hands-on machine learning with scikit-learn, keras, and TensorFlow*, 2nd ed. Sebastopol, CA: O’Reilly Media, Oct. 2019.
- [44] Anatomis glossary. [Online]. Available: [https://brainvisa.info/anatomist-5.0/user\\_doc/glossary.html](https://brainvisa.info/anatomist-5.0/user_doc/glossary.html)
- [45] Y. Cointepas, J.-F. Mangin, L. Garnero, J.-B. Poline, and H. Benali, “BrainVISA: Software platform for visualization and analysis of multi-modality brain data,” *Neuroimage*, vol. 13, no. 6, p. 98, Jun. 2001.
- [46] C. Román, C. Hernández, M. Figueroa, J. Houenou, C. Poupon, J.-F. Mangin, and P. Guevara, “Superficial white matter bundle atlas based on hierarchical fiber clustering over probabilistic tractography data,” *Neuroimage*, vol. 262, no. 119550, p. 119550, Nov. 2022.
- [47] A. Vázquez, N. López-López, N. Labra, M. Figueroa, C. Poupon, J.-F. Mangin, C. Hernández, and P. Guevara, “Parallel optimization of fiber bundle segmentation for massive tractography datasets,” in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 178–181.

- [48] S. Navarrete, “Informe de memoria de título Ingeniería Civil Biomédica, Universidad de Concepción, 2023: Uso de autoencoder para la segmentación y simulación de fibras cerebrales.”
- [49] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. van der Walt, M. Descoteaux, I. Nimmo-Smith, and Dipy Contributors, “Dipy, a library for the analysis of diffusion MRI data,” *Front. Neuroinform.*, vol. 8, p. 8, Feb. 2014.
- [50] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: fundamental algorithms for scientific computing in python,” *Nat. Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020.
- [51] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, “Scikit-learn: Machine learning in python,” *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

## Anexo

Sujeto	as	bw	cb	ch	ct	db	df	ef	fa	fg
Cant. Fibras	940551	985020	926240	955667	860247	746313	841087	912931	1006840	1087360
% S2	19 %	21 %	16 %	23 %	19 %	25 %	21 %	23 %	16 %	18 %
% S1	16 %	19 %	14 %	19 %	18 %	21 %	17 %	19 %	13 %	17 %
Diferencia	3 %	2 %	2 %	4 %	1 %	4 %	4 %	4 %	3 %	1 %

**Tabla A.1:** Parte 1 de la comparación de porcentajes de fibras segmentadas de 'S1' y 'S2' según la cantidad de fibras de la tractografía original.

Sujeto	gb	hc	ks	lk	mb	nl	pb	rd	sc
Cant. Fibras	985471	872280	864309	1024622	927849	849484	922973	862667	877225
% S2	19 %	24 %	20 %	23 %	23 %	24 %	23 %	24 %	22 %
% S1	20 %	21 %	17 %	20 %	19 %	19 %	18 %	21 %	19 %
Diferencia	-1 %	3 %	3 %	3 %	4 %	5 %	5 %	3 %	3 %

**Tabla A.2:** Parte 2 de la comparación de porcentajes de fibras segmentadas de 'S1' y 'S2' según la cantidad de fibras de la tractografía original.

Sujeto	tm	cl4	cl1	cn	hl	ig	ip	jg	jp
Cant. Fibras	777916	969268	989190	1018088	914469	847919	896261	1053387	958126
% S2	15 %	26 %	15 %	22 %	22 %	17 %	24 %	22 %	18 %
% S1	9 %	23 %	15 %	18 %	19 %	19 %	19 %	21 %	19 %
Diferencia	6 %	3 %	0 %	4 %	3 %	-2 %	5 %	1 %	-1 %

**Tabla A.3:** Parte 3 de la comparación de porcentajes de fibras segmentadas de 'S1' y 'S2' según la cantidad de fibras de la tractografía original.

Sujeto	lf	lj	ls	oj	ra	ru	sb3	sb0	yh
Cant. Fibras	994919	943514	910532	681053	994015	921121	852948	985079	1047253
% S2	23 %	20 %	23 %	19 %	21 %	17 %	19 %	20 %	22 %
% S1	19 %	18 %	21 %	15 %	17 %	18 %	18 %	16 %	19 %
Diferencia	4 %	2 %	2 %	4 %	4 %	-1 %	1 %	4 %	3 %

**Tabla A.4:** Parte 4 de la comparación de porcentajes de fibras segmentadas de 'S1' y 'S2' según la cantidad de fibras de la tractografía original.

Fascículo	PRUEBAS DE NORMALIDAD EN VALORES ADC											
	SHAPIRO-WILK				ASIMETRÍA				CURTOSIS			
	S1		S2		S1		S2		S1		S2	
	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA
AR_ANT_LEFT	0,391	0,097	0,755	<b>0,019</b>	0,076	0,767	-0,174	0,999	-0,977	-0,075	-0,344	0,108
AR_ANT_RIGHT	0,613	0,559	0,359	0,549	0,565	-0,068	0,25	-0,028	-0,251	-0,405	-0,974	-0,439
AR_LEFT	0,746	<b>0,002</b>	0,864	0,084	-0,426	<b>1,836</b>	-0,15	<b>1,17</b>	-0,083	<b>3,853</b>	-0,489	<b>1,717</b>
AR_RIGHT	0,363	0,811	0,917	0,626	0,339	-0,03	-0,007	-0,226	-0,869	-0,78	-0,698	-0,867
AR_POST_LEFT	0,379	0,25	0,838	0,627	-0,054	-0,119	0,241	0,114	0,412	<b>-1,165</b>	0,064	-0,913
AR_POST_RIGHT	0,989	0,143	0,98	0,605	0,187	-0,173	0,164	-0,012	-0,101	<b>-1,106</b>	-0,455	<b>-1,087</b>
CC_BODY	0,054	0,938	<b>0,037</b>	0,919	0,945	-0,105	0,997	-0,118	-0,026	-0,665	0,107	-0,591
CC_GENU	0,572	0,562	0,517	0,889	0,48	-0,261	0,386	-0,237	-0,682	-0,738	-0,896	-0,658
CC_ROSTRUM	<b>0,013</b>	0,596	0,168	0,234	<b>1,191</b>	-0,046	0,999	0,133	<b>1,187</b>	-0,851	<b>1,174</b>	-0,619
CC_SPLENIUM	0,708	0,918	$\approx 0$	<b>0,005</b>	-0,276	0,186	<b>2,812</b>	<b>1,725</b>	-0,186	-0,388	<b>8,068</b>	<b>3,669</b>
CG2_LEFT	0,804	0,997	0,519	0,921	0,395	-0,043	0,234	-0,015	-0,652	0,075	<b>-1,021</b>	0,424
CG2_RIGHT	0,487	0,19	0,251	0,643	0,406	0,957	0,598	0,613	-0,439	0,92	-0,304	0,85
CG3_LEFT	0,952	0,734	0,94	0,074	0,333	-0,12	0,024	-0,33	-0,294	-0,754	-0,708	-0,837
CG3_RIGHT	0,288	0,376	0,929	0,343	-0,207	0,731	0,111	0,367	0,454	<b>1,053</b>	-0,312	0,015
CG_LEFT	0,798	0,254	0,713	0,747	-0,417	0,958	-0,446	0,633	-0,356	0,869	-0,181	0,346
CG_RIGHT	0,257	0,131	0,595	0,184	-0,164	<b>1,004</b>	-0,185	0,945	0,307	<b>1,502</b>	0,26	<b>1,319</b>
CST_LEFT	0,398	0,339	0,78	0,123	0,372	-0,145	0,544	-0,184	<b>1,124</b>	<b>-1,095</b>	0,457	<b>-1,349</b>
CST_RIGHT	0,222	0,667	0,107	0,367	0,571	-0,552	<b>1,03</b>	0,098	0,327	-0,158	<b>1,943</b>	<b>-1,019</b>
FORNIX_LEFT	0,106	0,208	0,397	0,077	-0,407	-0,617	0,689	-0,945	<b>-1,169</b>	<b>1,762</b>	0	<b>1,974</b>
FORNIX_RIGHT	0,189	0,584	0,131	0,23	0,672	-0,202	0,574	-0,329	-0,3	<b>1,046</b>	-0,346	<b>1,191</b>
IFO_LEFT	0,971	0,866	0,948	0,726	0,009	0,363	-0,07	0,262	-0,173	0,349	-0,506	-0,094
IFO_RIGHT	0,663	0,05	0,685	0,2	0,196	<b>-1,123</b>	0,356	-0,94	-0,704	<b>1,524</b>	-0,49	<b>1,034</b>
IL_LEFT	0,401	<b>0,026</b>	0,849	0,084	0,488	<b>1,137</b>	0,504	0,914	0,158	0,8	-0,172	0,138
IL_RIGHT	0,228	0,27	0,813	0,285	0,901	-0,197	0,109	-0,375	0,785	<b>-1,218</b>	-0,519	<b>-1,086</b>
THAL_FRONT_LEFT	0,393	0,47	0,602	0,266	0,557	0,322	0,612	0,598	-0,424	-0,555	-0,051	0,681
THAL_FRONT_RIGHT	0,44	0,502	0,895	0,587	0,617	0,207	0,052	0,456	-0,069	-0,899	-0,736	-0,305
THAL_MOT_LEFT	0,195	0,259	0,461	0,328	0,897	-0,06	0,424	-0,112	0,361	-0,929	0,612	<b>-1,125</b>
THAL_MOT_RIGHT	0,147	0,804	<b>0,019</b>	0,567	0,877	0,051	<b>1,344</b>	0,081	0,299	-0,905	<b>1,631</b>	<b>-1,013</b>
THAL_OCC_LEFT	0,265	0,43	0,62	0,166	0,919	0,512	-0,459	0,719	<b>1,124</b>	0,259	-0,553	<b>1,541</b>
THAL_OCC_RIGHT	0,512	0,832	0,514	0,286	0,786	-0,143	0,101	0,55	0,695	-0,511	-0,863	-0,566
THAL_PAR_LEFT	0,31	<b>0,042</b>	0,878	0,431	0,185	<b>1,215</b>	0,237	0,594	<b>-1,022</b>	<b>2,197</b>	0,2	0,18
THAL_PAR_RIGHT	0,871	0,462	0,2	0,105	-0,146	0,618	0,114	0,833	0,421	0,098	<b>-1,131</b>	-0,106
THAL_TEMP_LEFT	0,684	0,815	0,968	0,576	0,282	0,314	-0,23	0,628	0,737	-0,203	-0,139	-0,245
THAL_TEMP_RIGHT	0,701	0,424	0,926	0,253	0,574	0,547	-0,053	0,905	0,192	-0,014	0,063	0,95
UN_LEFT	0,71	0,222	0,349	<b>0,016</b>	0,474	0,821	0,771	<b>1,247</b>	-0,159	0,372	0,204	<b>1,233</b>
UN_RIGHT	0,808	0,502	0,679	0,101	0,086	0,447	-0,256	0,603	-0,588	-0,502	-0,41	-0,738

**Tabla A.5:** Pruebas de normalidad de Shapiro-Wilk, asimetría y curtosis en valores ADC por fascículo obtenido según los métodos de segmentación 'S1' y 'S2', para las submuestras de sujetos sin autismo (SA) y con autismo (CA). El color rojo representa los valores que indican una distribución no normal. En caso de Shapiro-Wilk,

$>0.05$ , y en el caso de asimetría y curtosis,  $>1$  y  $-1<$ .

Fascículo	PRUEBAS DE NORMALIDAD EN VALORES FA											
	SHAPIRO-WILK				ASIMETRÍA				CURTOSIS			
	S1		S2		S1		S2		S1		S2	
	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA
AR_ANT_LEFT	0,497	0,729	0,61	0,687	-0,239	-0,449	0,096	0,1	-0,378	-0,007	-0,399	0,252
AR_ANT_RIGHT	0,173	0,864	0,051	<b>0,043</b>	0,947	0,186	<b>1,179</b>	<b>-1,089</b>	0,526	0,11	<b>1,319</b>	<b>1,206</b>
AR_LEFT	<b>0,007</b>	0,561	0,154	0,444	<b>1,536</b>	-0,553	0,958	0,218	<b>2,859</b>	0,211	0,376	-0,266
AR_RIGHT	0,267	0,253	0,096	0,818	0,549	0,771	0,748	0,535	-0,534	0,878	-0,461	0,14
AR_POST_LEFT	0,506	0,312	0,577	0,858	0,481	-0,824	0,29	-0,309	-0,383	0,603	-0,734	0,223
AR_POST_RIGHT	0,521	0,886	0,804	0,839	-0,273	0,263	0,058	0,217	-0,729	-0,142	-0,463	-0,686
CC_BODY	0,453	0,64	0,839	0,658	0,293	0,144	0,369	0,385	-0,791	<b>-1,065</b>	-0,377	-0,686
CC_GENU	0,818	<b>0,02</b>	0,68	<b>0,032</b>	0,207	0,667	0,49	0,622	-0,26	-0,976	-0,086	-0,856
CC_ROSTRUM	0,419	0,515	0,193	0,981	-0,637	0,454	0,261	-0,052	-0,225	-0,436	<b>-1,252</b>	-0,278
CC_SPLENIUM	0,336	0,552	0,472	<b>0,016</b>	-0,276	0,741	0,767	<b>1,08</b>	-0,397	0,45	0,436	0,244
CG2_LEFT	0,346	0,734	0,448	0,373	0,675	0,156	0,402	0,177	-0,141	-0,656	-0,934	<b>-1,119</b>
CG2_RIGHT	0,573	0,643	0,482	0,072	0,16	0,242	-0,058	0,189	-0,844	-0,567	<b>-1,125</b>	<b>-1,385</b>
CG3_LEFT	<b>0,042</b>	0,417	0,203	0,773	0,767	0,507	0,635	0,09	-0,595	-0,241	-0,162	0,03
CG3_RIGHT	0,205	0,986	0,112	0,267	-0,632	-0,11	-0,463	-0,355	-0,655	-0,27	-0,869	-0,783
CG_LEFT	0,236	0,661	0,84	0,703	0,254	0,427	0,15	0,625	<b>-1,186</b>	-0,681	-0,92	0,322
CG_RIGHT	0,212	<b>0,016</b>	0,353	0,427	0,503	<b>1,482</b>	-0,133	-0,193	-0,917	<b>2,964</b>	-0,987	<b>-1,037</b>
CST_LEFT	0,217	0,897	0,371	0,622	0,848	0,431	0,638	0,369	0,83	0,039	0,085	-0,72
CST_RIGHT	0,999	0,236	0,555	0,465	-0,165	0,341	0,377	0,239	-0,359	<b>-1,153</b>	-0,768	-0,789
FORNIX_LEFT	0,22	0,732	0,563	0,286	-0,169	0,29	-0,016	0,575	<b>-1,276</b>	-0,467	<b>-1,106</b>	0,291
FORNIX_RIGHT	0,872	0,744	0,152	0,136	0,101	0,044	0,686	0,515	-0,767	-0,73	-0,585	0,438
IFO_LEFT	0,778	$\approx 0$	0,808	<b>0,015</b>	-0,314	<b>2,006</b>	0,241	<b>1,476</b>	-0,735	<b>5,821</b>	-0,364	<b>2,958</b>
IFO_RIGHT	0,274	$\approx 0$	0,732	<b>0,004</b>	0,238	<b>2,736</b>	-0,078	<b>1,468</b>	-0,912	<b>8,179</b>	-0,535	<b>1,628</b>
IL_LEFT	0,605	0,902	0,22	0,565	0,478	-0,112	0,272	-0,241	-0,511	-0,009	<b>-1,255</b>	-0,803
IL_RIGHT	0,197	$\approx 0$	0,098	$\approx 0$	0,326	<b>2,847</b>	0,663	<b>3,156</b>	<b>-1,166</b>	<b>8,422</b>	-0,597	<b>9,657</b>
THAL_FRONT_LEFT	0,286	0,228	0,384	0,357	0,626	0,418	-0,026	0,63	0,191	-0,924	<b>-1,121</b>	0,362
THAL_FRONT_RIGHT	0,325	0,307	0,236	<b>0,023</b>	0,505	0,389	0,386	<b>1,154</b>	-0,765	-0,232	-0,475	0,586
THAL_MOT_LEFT	0,67	0,643	0,394	0,645	-0,412	0,331	0,446	0,337	-0,569	-0,796	-0,796	-0,91
THAL_MOT_RIGHT	0,868	0,416	0,973	0,07	0,474	0,582	0,014	0,296	0,464	-0,003	-0,044	<b>-1,259</b>
THAL_OCC_LEFT	0,678	0,415	0,62	<b>0,022</b>	-0,313	-0,723	0,476	<b>1,281</b>	-0,393	0,305	-0,258	<b>1,953</b>
THAL_OCC_RIGHT	0,833	0,701	0,269	$\approx 0$	0,129	0,46	0,462	<b>2,364</b>	-0,815	0,206	<b>1,048</b>	<b>6,216</b>
THAL_PAR_LEFT	<b>0,003</b>	0,999	0,083	0,539	<b>1,223</b>	0,073	0,704	-0,272	0,456	0,062	-0,54	0,043
THAL_PAR_RIGHT	0,546	0,741	0,633	0,47	0,093	-0,183	0,118	-0,365	0,196	-0,621	-0,473	-0,972
THAL_TEMP_LEFT	0,165	0,884	0,603	0,599	0,59	0,049	0,097	0,576	-0,498	-0,752	<b>-1,087</b>	0,4
THAL_TEMP_RIGHT	0,369	0,853	0,934	0,581	-0,412	-0,19	-0,337	0,269	-0,586	-0,77	0,288	0,88
UN_LEFT	0,327	0,635	0,604	0,44	0,357	0,413	0,377	-0,463	-0,769	0,425	-0,422	0,132
UN_RIGHT	0,334	0,38	0,304	0,863	0,321	-0,214	0,472	-0,287	-0,825	<b>-1,077</b>	-0,833	-0,544

**Tabla A.6:** Pruebas de normalidad de Shapiro-Wilk, asimetría y curtosis en valores FA por fascículo obtenido según los métodos de segmentación 'S1' y 'S2', para las submuestras de sujetos sin autismo (SA) y con autismo (CA). El color rojo representa los valores que indican una distribución no normal. En caso de Shapiro-Wilk,  $>0.05$ , y en el caso de asimetría y curtosis,  $>1$  y  $-1<$ .



Fascículo	PRUEBAS DE NORMALIDAD EN VALORES GFA											
	SHAPIRO-WILK				ASIMETRÍA				CURTOSIS			
	S1		S2		S1		S2		S1		S2	
	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA	SA	CA
AR_ANT_LEFT	0,83	0,088	0,518	<b>0,024</b>	0,356	-0,268	0,225	0,51	0,013	0,77	-0,828	<b>1,769</b>
AR_ANT_RIGHT	0,327	0,847	0,899	0,937	0,456	0,051	-0,036	-0,254	-0,669	0,001	-0,188	-0,258
AR_LEFT	0,365	0,651	0,202	0,784	0,507	-0,131	0,221	0,284	-0,314	-0,751	<b>-1,291</b>	-0,757
AR_RIGHT	0,65	<b>0,014</b>	<b>0,003</b>	0,163	0,593	<b>1,438</b>	<b>1,393</b>	0,973	0,179	<b>2,514</b>	<b>1,116</b>	0,794
AR_POST_LEFT	0,347	<b>0,046</b>	0,426	<b>0,026</b>	-0,31	-0,377	-0,17	-0,348	-0,646	<b>-1,298</b>	-0,487	<b>-1,401</b>
AR_POST_RIGHT	0,766	0,404	0,256	0,234	0,293	0,596	0,018	0,499	-0,88	-0,228	<b>-1,142</b>	-0,717
CC_BODY	0,13	0,988	0,686	0,721	-0,028	0,054	0,242	0,458	<b>-1,146</b>	-0,373	-0,816	0,529
CC_GENU	0,279	0,578	0,674	0,187	-0,424	0,149	-0,175	0,518	-0,487	-0,812	-0,747	-0,764
CC_ROSTRUM	0,321	0,061	0,328	0,949	0,551	<b>1,189</b>	-0,11	0,001	0,474	<b>2,205</b>	-0,937	0,097
CC_SPLENIUM	0,96	0,057	0,138	<b>0,007</b>	-0,115	<b>1,07</b>	0,747	<b>1,35</b>	-0,262	<b>1,804</b>	-0,236	<b>1,102</b>
CG2_LEFT	0,269	0,497	0,918	0,657	-0,023	0,116	-0,322	0,15	0,573	<b>-1,042</b>	0,107	-0,899
CG2_RIGHT	<b>0,018</b>	0,082	0,203	<b>0,034</b>	0,245	-0,187	0,588	-0,114	<b>-1,562</b>	<b>-1,368</b>	-0,669	<b>-1,542</b>
CG3_LEFT	0,713	0,581	0,863	0,877	-0,055	0,599	0,017	0,108	<b>-1,055</b>	0,587	-0,849	-0,742
CG3_RIGHT	0,956	0,906	0,987	0,758	0,131	0,221	-0,313	0,349	0,25	-0,393	-0,248	0,147
CG_LEFT	0,575	0,771	0,968	0,988	0,162	-0,015	-0,199	0,182	0,666	-0,925	-0,153	0,15
CG_RIGHT	0,995	<b>0,017</b>	0,499	<b>0,015</b>	-0,002	<b>1,263</b>	-0,447	-0,237	-0,749	<b>1,951</b>	-0,496	<b>-1,595</b>
CST_LEFT	0,073	0,905	<b>0,003</b>	0,974	<b>1,17</b>	0,147	<b>1,332</b>	-0,113	<b>2,072</b>	0,365	<b>1,215</b>	-0,407
CST_RIGHT	0,88	0,451	0,257	0,537	-0,077	0,328	0,272	0,258	-0,454	-0,974	<b>-1,101</b>	-0,746
FORNIX_LEFT	0,099	0,777	0,674	0,087	-0,692	0,397	-0,258	0,678	-0,305	-0,128	-0,535	-0,565
FORNIX_RIGHT	0,655	0,975	0,646	0,921	0,537	-0,007	0,603	-0,333	-0,2	-0,079	-0,131	-0,385
IFO_LEFT	0,05	$\approx 0$	0,46	<b>0,002</b>	<b>-1,083</b>	<b>2,672</b>	-0,335	<b>1,821</b>	0,791	<b>7,647</b>	-0,018	<b>3,753</b>
IFO_RIGHT	0,857	$\approx 0$	0,756	<b>0,001</b>	0,396	<b>2,951</b>	0,586	<b>1,56</b>	-0,565	<b>8,834</b>	0,099	<b>1,888</b>
IL_LEFT	0,974	0,199	0,12	0,949	0,164	0,93	-0,85	-0,048	0,305	0,527	0,102	-0,477
IL_RIGHT	0,051	$\approx 0$	0,717	$\approx 0$	-0,725	<b>3,22</b>	0,339	<b>3,141</b>	-0,066	<b>10,23</b>	-0,099	<b>9,726</b>
THAL_FRONT_LEFT	0,576	0,61	0,821	0,164	0,317	0,253	0,479	0,372	0,217	-0,718	0,269	-0,943
THAL_FRONT_RIGHT	<b>0,047</b>	0,65	0,958	0,171	0,799	0,293	-0,293	0,432	0,435	-0,853	0,069	-0,982
THAL_MOT_LEFT	0,407	0,535	0,369	0,544	-0,009	0,008	0,701	0,168	-0,645	-0,874	0,099	-0,201
THAL_MOT_RIGHT	0,983	0,665	0,478	0,47	0,223	0,31	-0,277	0,374	-0,025	-0,767	-0,304	-0,727
THAL_OCC_LEFT	0,694	0,971	0,83	<b>0,003</b>	-0,142	-0,126	-0,087	<b>1,63</b>	-0,52	-0,814	-0,868	<b>3,855</b>
THAL_OCC_RIGHT	0,08	0,432	0,203	$\approx 0$	-0,264	0,375	0,356	<b>2,36</b>	<b>-1,32</b>	-0,825	0,195	<b>6,62</b>
THAL_PAR_LEFT	<b>0,006</b>	0,114	0,796	0,075	<b>1,058</b>	$\approx 1$	-0,185	0,559	0,11	0,893	-0,24	-0,87
THAL_PAR_RIGHT	0,067	0,9	0,998	0,301	<b>-1,051</b>	-0,062	0,03	0,016	<b>1,941</b>	0,065	-0,504	-0,041
THAL_TEMP_LEFT	0,157	0,252	0,499	<b>0,013</b>	0,785	0,459	0,177	<b>1,486</b>	0,087	-0,37	-0,884	<b>2,251</b>
THAL_TEMP_RIGHT	0,689	0,309	0,331	<b>0,003</b>	0,149	-0,391	0,569	<b>1,5</b>	-0,257	-0,589	0,081	<b>1,766</b>
UN_LEFT	0,77	0,126	0,466	0,284	0,451	0,953	0,52	-0,131	-0,431	0,921	-0,691	<b>-1,011</b>
UN_RIGHT	0,37	0,504	0,933	0,605	0,471	0,351	-0,168	0,19	0,217	-0,455	-0,637	-0,777

**Tabla A.7:** Pruebas de normalidad de Shapiro-Wilk, asimetría y curtosis en valores GFA por fascículo obtenido según los métodos de segmentación 'S1' y 'S2', para las submuestras de sujetos sin autismo (SA) y con autismo (CA). El color rojo representa los valores que indican una distribución no normal. En caso de Shapiro-Wilk,

$>0.05$ , y en el caso de asimetría y curtosis,  $>1$  y  $-1<$ .

Fascículo	PRUEBAS DE HOMOGENEIDAD DE VARIANZAS EN VALORES ADC, FA Y GFA											
	ADC				FA				GFA			
	S1		S2		S1		S2		S1		S2	
	B	L	B	L	B	L	B	L	B	L	B	L
AR_ANT_LEFT	0,594	0,417	0,655	0,427	0,307	0,234	0,099	<b>0,03</b>	0,657	0,238	0,164	<b>0,002</b>
AR_ANT_RIGHT	0,566	0,441	0,636	0,487	0,972	0,857	0,351	0,209	0,855	0,52	0,51	0,513
AR_LEFT	0,828	0,22	0,877	0,608	0,113	0,16	0,912	0,942	0,079	0,051	0,375	0,319
AR_RIGHT	0,171	0,196	0,531	0,591	0,463	0,29	0,557	0,606	0,685	0,172	0,859	0,929
AR_POST_LEFT	0,193	0,346	0,367	0,402	0,798	0,897	0,437	0,539	0,407	0,718	0,752	0,721
AR_POST_RIGHT	<b>0,026</b>	0,063	<b>0,01</b>	<b>0,032</b>	0,682	0,696	0,538	0,345	0,992	0,845	0,978	0,883
CC_BODY	0,122	0,097	0,163	0,098	0,075	0,161	0,144	0,185	0,268	0,2	0,205	0,086
CC_GENU	0,736	0,633	0,57	0,495	0,325	0,183	0,544	0,343	0,409	0,389	0,663	0,523
CC_ROSTRUM	0,196	0,544	0,082	0,081	0,717	0,53	0,366	0,276	0,755	0,841	0,726	0,564
CC_SPLENIUM	0,169	0,111	0,774	0,583	0,782	0,62	0,84	0,667	0,824	0,962	0,948	0,644
CG2_LEFT	0,383	0,231	0,595	0,278	0,35	0,495	0,429	0,487	<b>0,021</b>	0,051	0,064	0,078
CG2_RIGHT	0,123	<b>0,023</b>	0,806	0,652	0,484	0,536	0,666	0,568	0,099	0,241	<b>0,008</b>	<b>0,048</b>
CG3_LEFT	0,201	0,196	0,294	0,166	0,837	0,61	0,332	0,16	0,906	0,943	0,91	0,98
CG3_RIGHT	0,251	0,093	0,818	0,597	0,295	0,268	0,434	0,393	0,32	0,287	0,721	0,607
CG_LEFT	0,684	0,464	0,497	0,26	0,306	0,214	0,379	0,255	<b>0,003</b>	<b>0,009</b>	0,299	0,229
CG_RIGHT	0,244	0,166	0,601	0,338	0,387	0,074	0,092	0,138	<b>0,026</b>	<b>0,001</b>	$\approx 0$	<b>0,001</b>
CST_LEFT	0,695	0,836	0,232	0,597	0,411	0,335	0,081	0,057	0,383	0,551	<b>0,026</b>	<b>0,048</b>
CST_RIGHT	0,385	0,22	0,739	0,745	0,529	0,785	0,851	0,824	0,32	0,386	0,238	0,123
FORNIX_LEFT	0,078	<b>0,002</b>	0,157	<b>0,028</b>	0,16	0,06	0,72	0,333	0,879	0,979	0,752	0,741
FORNIX_RIGHT	0,585	0,284	0,591	0,481	0,239	0,179	0,504	0,148	0,305	0,181	0,201	0,254
IFO_LEFT	0,387	0,224	0,315	0,176	0,641	<b>0,033</b>	0,951	0,531	0,152	$\approx 0$	0,252	<b>0,049</b>
IFO_RIGHT	0,307	0,184	0,067	<b>0,025</b>	0,924	0,106	0,45	0,093	0,576	<b>0,006</b>	0,41	0,14
IL_LEFT	0,504	0,201	0,523	0,377	0,865	0,83	0,306	0,528	<b>0,029</b>	<b>0,001</b>	0,432	0,394
IL_RIGHT	0,155	0,321	0,213	0,252	0,956	0,05	0,89	<b>0,04</b>	0,44	$\approx 0$	0,466	$\approx 0$
THAL_FRONT_LEFT	0,259	0,181	0,161	<b>0,018</b>	0,318	0,27	0,649	0,259	0,253	0,385	0,943	0,637
THAL_FRONT_RIGHT	0,067	0,129	0,156	0,098	0,92	0,736	0,835	0,973	0,097	0,227	0,243	0,364
THAL_MOT_LEFT	0,469	0,668	0,102	0,34	0,355	0,36	0,779	0,945	0,277	0,264	0,3	0,344
THAL_MOT_RIGHT	0,669	0,904	0,512	0,833	0,859	0,894	0,287	0,662	0,269	0,389	0,413	0,439
THAL_OCC_LEFT	0,546	0,352	0,799	0,742	0,522	0,279	0,138	0,218	0,464	0,538	0,083	0,255
THAL_OCC_RIGHT	0,072	0,09	0,764	0,757	0,844	0,934	0,995	0,435	0,521	0,551	0,725	0,094
THAL_PAR_LEFT	0,311	0,056	0,374	0,361	0,376	0,595	0,895	0,85	0,246	0,321	0,354	0,315
THAL_PAR_RIGHT	0,092	0,069	0,677	0,6	0,692	0,761	0,149	0,084	0,521	0,545	0,571	0,518
THAL_TEMP_LEFT	0,088	<b>0,049</b>	<b>0,019</b>	<b>0,003</b>	<b>0,014</b>	<b>0,009</b>	0,141	<b>0,008</b>	0,708	0,807	0,568	0,916
THAL_TEMP_RIGHT	0,17	<b>0,044</b>	0,466	0,237	0,75	0,903	0,173	0,053	0,29	0,148	0,507	0,274
UN_LEFT	0,252	0,183	0,075	<b>0,025</b>	0,341	0,114	0,386	0,207	0,124	0,059	0,842	0,927
UN_RIGHT	0,303	0,233	0,135	0,108	0,957	0,762	0,401	0,436	0,739	0,731	0,721	0,781

**Tabla A.8:** Pruebas de homogeneidad de Bartlett (B) y Levene (L), en valores ADC, FA y GFA obtenidos de los fascículos segmentados con 'S1' y 'S2'. El color rojo destaca los valores que indican no homogeneidad de varianzas al ser menores que el valor de significancia 0.05.

Fascículo	ADC						FA						GFA					
	S1			S2			S1			S2			S1			S2		
	DN	VH	PP	DN	VH	PP	DN	VH	PP	DN	VH	PP	DN	VH	PP	DN	VH	PP
AR_ANT_LEFT	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	0
AR_ANT_RIGHT	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1
AR_LEFT	0	1	0	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1
AR_RIGHT	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0
AR_POST_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0
AR_POST_RIGHT	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
CC_BODY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CC_GENU	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CC_ROSTRUM	0	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1
CC_SPLENIUM	1	1	1	0	1	0	1	1	1	0	1	0	0	1	0	0	1	0
CG2_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
CG2_RIGHT	1	0	0	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0
CG3_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CG3_RIGHT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CG_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
CG_RIGHT	0	1	0	1	1	1	0	1	0	1	1	1	0	0	0	0	0	0
CST_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0
CST_RIGHT	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
FORNIX_LEFT	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
FORNIX_RIGHT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
IFO_LEFT	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0
IFO_RIGHT	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	1	0	0
IL_LEFT	0	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
IL_RIGHT	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
THAL_FRONT_LEFT	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
THAL_FRONT_RIGHT	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1
THAL_MOT_LEFT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
THAL_MOT_RIGHT	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
THAL_OCC_LEFT	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0
THAL_OCC_RIGHT	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0
THAL_PAR_LEFT	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	1	1
THAL_PAR_RIGHT	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1
THAL_TEMP_LEFT	1	0	0	1	0	0	1	0	0	1	0	0	1	1	1	0	1	0
THAL_TEMP_RIGHT	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0
UN_LEFT	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
UN_RIGHT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

**Tabla A.9:** Resumen de resultados de pruebas estadísticas descriptivas para cada fascículo obtenido con distintos métodos de segmentación (1 = cumple, 0 = no cumple). Si para las pruebas de normalidad, la prueba Shapiro-Wilk y al menos una más, indicaban normalidad, se indica que cumple con normalidad. En cuanto a las pruebas de homogeneidad, si ambas pruebas indican homogeneidad, se dice que cumple con homogeneidad. En cuanto a la columna 'PP', indica si se cumplen ambos supuestos. Es decir, si ambas columnas anteriores marcan '1', cumplirá los supuestos para aplicar pruebas paramétricas. En azul se destacan las que sí cumplen.

	Resultados de modelos de clasificación aplicados a valores ADC							
	Decision Trees				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
AR_ANT_LEFT	0,4	0,2	0,5	0,286	0,467	0,5	0,5	0,464
AR_ANT_RIGHT	0,333	0,167	0,5	0,25	0,333	0,167	0,5	0,25
AR_LEFT	0,467	0,233	0,5	0,318	0,267	0,25	0,277	0,253
AR_RIGHT	0,467	0,233	0,5	0,318	0,533	0,583	0,554	0,498
AR_POST_LEFT	0,467	0,233	0,5	0,318	0,467	0,233	0,5	0,318
AR_POST_RIGHT	0,4	0,2	0,5	0,286	0,267	0,261	0,306	0,253
CC_BODY	0,467	0,233	0,5	0,318	0,4	0,4	0,411	0,389
CC_GENU	0,467	0,233	0,5	0,318	0,333	0,333	0,339	0,33
CC_ROSTRUM	0,467	0,233	0,5	0,318	0,6	0,6	0,589	0,583
CC_SPLENIUM	0,4	0,2	0,5	0,286	0,533	0,556	0,556	0,533
CG2_LEFT	0,467	0,233	0,5	0,318	0,533	0,55	0,545	0,525
CG2_RIGHT	0,467	0,233	0,5	0,318	0,6	0,6	0,589	0,583
<b>CG3_LEFT</b>	0,4	0,2	0,5	0,286	<b>0,667</b>	<b>0,773</b>	<b>0,722</b>	<b>0,661</b>
CG3_RIGHT	0,4	0,2	0,5	0,286	0,4	0,2	0,5	0,286
CG_LEFT	0,333	0,167	0,5	0,25	0,4	<b>0,679</b>	0,55	0,354
<b>CG_RIGHT</b>	0,4	0,2	0,5	0,286	<b>0,8</b>	<b>0,833</b>	<b>0,833</b>	<b>0,8</b>
CST_LEFT	0,467	0,233	0,5	0,318	0,267	0,268	0,268	0,267
CST_RIGHT	0,467	0,233	0,5	0,318	0,467	0,233	0,5	0,318
FORNIX_LEFT	0,4	0,2	0,5	0,286	0,333	0,179	0,417	0,25
FORNIX_RIGHT	0,467	0,233	0,5	0,318	0,467	0,233	0,5	0,318
IFO_LEFT	0,467	0,233	0,5	0,318	0,533	0,528	0,527	0,525
IFO_RIGHT	0,467	0,233	0,5	0,318	0,467	0,233	0,5	0,318
IL_LEFT	0,467	0,233	0,5	0,318	0,467	0,233	0,5	0,318
IL_RIGHT	0,467	0,233	0,5	0,318	0,6	0,598	0,598	0,598
THAL_FRONT_LEFT	0,467	0,233	0,5	0,318	0,333	0,33	0,33	0,33
THAL_FRONT_RIGHT	0,467	0,233	0,5	0,318	0,4	0,389	0,393	0,389
THAL_MOT_LEFT	0,467	0,233	0,5	0,318	0,6	0,598	0,598	0,598
THAL_MOT_RIGHT	0,467	0,233	0,5	0,318	0,533	0,519	0,509	0,444
THAL_OCC_LEFT	0,467	0,233	0,5	0,318	0,467	0,472	0,473	0,464
THAL_OCC_RIGHT	0,4	0,2	0,5	0,286	0,533	<b>0,731</b>	<b>0,611</b>	0,498
THAL_PAR_LEFT	0,4	0,2	0,5	0,286	0,4	0,393	0,389	0,389
<b>THAL_PAR_RIGHT</b>	0,467	0,233	0,5	0,318	<b>0,667</b>	<b>0,693</b>	<b>0,652</b>	<b>0,641</b>
THAL_TEMP_LEFT	0,467	0,233	0,5	0,318	0,467	0,477	0,482	0,444
THAL_TEMP_RIGHT	0,4	0,2	0,5	0,286	0,4	0,2	0,5	0,286
UN_LEFT	0,467	0,233	0,5	0,318	0,533	0,536	0,536	0,533
UN_RIGHT	0,267	0,133	0,5	0,211	0,4	<b>0,654</b>	0,591	0,389

**Tabla A.10:** Resultados de modelos de clasificación aplicados a valores ADC, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.

	Resultados de modelos de clasificación aplicados a valores FA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
<b>AR_ANT_LEFT</b>	<b>0,667</b>	<b>0,65</b>	<b>0,639</b>	<b>0,641</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>
AR_ANT_RIGHT	0,467	0,5	0,5	0,464	0,467	0,5	0,5	0,464
AR_LEFT	0,333	0,33	0,33	0,33	0,333	0,333	0,339	0,33
AR_RIGHT	0,533	0,528	0,527	0,525	0,467	0,477	0,482	0,444
AR_POST_LEFT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,333	0,333	0,339	0,33
AR_POST_RIGHT	0,4	0,432	0,444	0,389	0,4	0,432	0,444	0,389
CC_BODY	0,333	0,333	0,339	0,33	0,467	0,472	0,473	0,464
CC_GENU	0,467	0,25	0,438	0,318	0,467	0,233	0,5	0,318
CC_ROSTRUM	0,467	0,45	0,455	0,444	0,4	0,4	0,411	0,389
CC_SPLENIUM	<b>0,6</b>	<b>0,607</b>	<b>0,611</b>	0,598	0,533	0,5	0,5	0,498
CG2_LEFT	0,467	0,481	0,491	0,4	0,467	0,472	0,473	0,464
CG2_RIGHT	<b>0,6</b>	<b>0,6</b>	0,589	0,583	0,533	<b>0,75</b>	0,563	0,444
CG3_LEFT	0,533	0,5	0,5	0,498	0,4	0,2	0,5	0,286
CG3_RIGHT	0,4	0,442	0,472	0,354	0,467	0,542	0,528	0,444
CG_LEFT	0,467	<b>0,692</b>	<b>0,6</b>	0,444	0,333	0,167	0,5	0,25
<b>CG_RIGHT</b>	0,467	0,444	0,444	0,444	<b>0,6</b>	<b>0,75</b>	<b>0,667</b>	0,583
<b>CST_LEFT</b>	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>	0,467	0,481	0,491	0,4
CST_RIGHT	0,4	0,352	0,384	0,354	0,133	0,134	0,134	0,133
FORNIX_LEFT	0,333	0,306	0,306	0,306	0,267	0,278	0,278	0,267
FORNIX_RIGHT	0,333	0,333	0,339	0,33	0,467	0,481	0,491	0,4
IFO_LEFT	0,4	0,4	0,411	0,389	0,467	0,464	0,464	0,464
IFO_RIGHT	<b>0,6</b>	0,598	0,598	0,598	0,467	0,233	0,5	0,318
IL_LEFT	0,333	0,3	0,321	0,306	0,467	0,45	0,455	0,444
IL_RIGHT	<b>0,6</b>	0,598	0,598	0,598	0,533	0,583	0,554	0,498
<b>THAL_FRONT_LEFT</b>	<b>0,667</b>	<b>0,7</b>	<b>0,679</b>	<b>0,661</b>	0,333	0,192	0,357	0,25
THAL_FRONT_RIGHT	0,333	0,33	0,33	0,33	0,333	0,333	0,339	0,33
THAL_MOT_LEFT	0,4	0,375	0,42	0,354	0,467	0,233	0,5	0,318
THAL_MOT_RIGHT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,533	0,523	0,518	0,498
THAL_OCC_LEFT	0,4	0,4	0,411	0,389	0,333	0,333	0,339	0,33
<b>THAL_OCC_RIGHT</b>	0,467	0,444	0,444	0,444	<b>0,6</b>	<b>0,75</b>	<b>0,667</b>	0,583
THAL_PAR_LEFT	0,4	0,417	0,417	0,4	<b>0,6</b>	0,3	0,5	0,375
THAL_PAR_RIGHT	0,267	0,268	0,268	0,267	0,333	0,33	0,33	0,33
THAL_TEMP_LEFT	0,333	0,333	0,339	0,33	0,467	0,464	0,464	0,464
THAL_TEMP_RIGHT	0,4	0,442	0,472	0,354	0,4	0,2	0,5	0,286
UN_LEFT	0,533	0,528	0,527	0,525	0,4	0,375	0,42	0,354
<b>UN_RIGHT</b>	0,533	<b>0,682</b>	<b>0,682</b>	0,533	0,4	<b>0,654</b>	0,591	0,389

**Tabla A.11:** Resultados de modelos de clasificación aplicados a valores FA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.

	Resultados de modelos de clasificación aplicados a valores GFA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
<b>AR_ANT_LEFT</b>	<b>0,667</b>	<b>0,694</b>	<b>0,694</b>	<b>0,667</b>	<b>0,6</b>	<b>0,65</b>	<b>0,639</b>	<b>0,598</b>
AR_ANT_RIGHT	0,467	0,455	0,45	0,444	0,333	0,167	0,5	0,25
AR_LEFT	0,467	0,464	0,464	0,464	<b>0,6</b>	<b>0,611</b>	<b>0,607</b>	0,598
AR_RIGHT	0,4	0,352	0,384	0,354	0,4	0,214	0,429	0,286
<b>AR_POST_LEFT</b>	0,533	0,583	0,554	0,498	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>
<b>AR_POST_RIGHT</b>	<b>0,667</b>	<b>0,821</b>	0,583	0,534	0,4	0,2	0,5	0,286
CC_BODY	0,333	0,3	0,321	0,306	0,4	0,389	0,393	0,389
CC_GENU	<b>0,6</b>	<b>0,611</b>	<b>0,607</b>	0,598	0,467	0,477	0,482	0,444
CC_ROSTRUM	0,467	0,464	0,464	0,464	0,467	0,477	0,482	0,444
CC_SPLENIUM	0,4	0,417	0,417	0,4	0,467	0,444	0,444	0,444
CG2_LEFT	0,4	0,352	0,384	0,354	0,467	0,233	0,5	0,318
<b>CG2_RIGHT</b>	<b>0,6</b>	<b>0,769</b>	<b>0,625</b>	0,55	<b>0,8</b>	<b>0,85</b>	<b>0,813</b>	<b>0,796</b>
CG3_LEFT	0,333	0,35	0,361	0,33	0,4	0,2	0,5	0,286
<b>CG3_RIGHT</b>	<b>0,6</b>	<b>0,75</b>	<b>0,667</b>	0,583	<b>0,6</b>	<b>0,65</b>	<b>0,639</b>	0,598
CG_LEFT	0,333	0,361	0,35	0,33	0,333	0,167	0,5	0,25
<b>CG_RIGHT</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,65</b>	<b>0,639</b>	<b>0,641</b>
<b>CST_LEFT</b>	<b>0,8</b>	<b>0,85</b>	<b>0,813</b>	<b>0,796</b>	<b>0,8</b>	<b>0,85</b>	<b>0,813</b>	<b>0,796</b>
CST_RIGHT	0,267	0,25	0,259	0,253	0,267	0,25	0,277	0,253
FORNIX_LEFT	0,533	0,527	0,528	0,525	0,333	0,179	0,417	0,25
FORNIX_RIGHT	0,333	0,307	0,348	0,306	0,333	0,167	0,5	0,25
<b>IFO_LEFT</b>	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>	0,4	0,402	0,402	0,4
IFO_RIGHT	0,533	0,519	0,509	0,444	0,467	0,233	0,5	0,318
IL_LEFT	0,333	0,33	0,33	0,33	0,467	0,233	0,5	0,318
IL_RIGHT	0,467	0,472	0,473	0,464	0,267	0,278	0,278	0,267
THAL_FRONT_LEFT	0,467	0,45	0,455	0,444	0,467	0,464	0,464	0,464
<b>THAL_FRONT_RIGHT</b>	<b>0,667</b>	<b>0,667</b>	<b>0,661</b>	<b>0,661</b>	0,467	0,464	0,464	0,464
THAL_MOT_LEFT	0,333	0,307	0,348	0,306	0,467	0,472	0,473	0,464
THAL_MOT_RIGHT	0,467	0,477	0,482	0,444	0,4	0,432	0,444	0,389
THAL_OCC_LEFT	0,533	0,528	0,527	0,525	<b>0,6</b>	<b>0,648</b>	<b>0,616</b>	0,583
THAL_OCC_RIGHT	0,533	0,527	0,528	0,525	0,467	0,233	0,5	0,318
THAL_PAR_LEFT	0,467	0,473	0,472	0,464	0,533	0,5	0,5	0,498
THAL_PAR_RIGHT	0,467	0,464	0,464	0,464	0,467	0,473	0,472	0,464
THAL_TEMP_LEFT	0,467	0,472	0,473	0,464	0,4	0,411	0,4	0,389
THAL_TEMP_RIGHT	0,467	0,542	0,528	0,444	0,4	0,2	0,5	0,286
UN_LEFT	0,533	0,536	0,536	0,533	0,533	0,528	0,527	0,525
<b>UN_RIGHT</b>	0,533	<b>0,682</b>	<b>0,682</b>	0,533	0,467	0,233	0,5	0,318

**Tabla A.12:** Resultados de modelos de clasificación aplicados a valores GFA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.

	Resultados de modelos de clasificación aplicados a valores ADC y FA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
AR_ANT_LEFT	<b>0,667</b>	<b>0,65</b>	<b>0,639</b>	<b>0,641</b>	0,4	0,2	0,5	0,286
AR_ANT_RIGHT	0,467	0,5	0,5	0,464	0,333	0,167	0,5	0,25
AR_LEFT	0,333	0,33	0,33	0,33	0,467	0,233	0,5	0,318
AR_RIGHT	0,533	0,528	0,527	0,525	0,4	0,214	0,429	0,286
AR_POST_LEFT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,467	0,233	0,5	0,318
AR_POST_RIGHT	0,4	0,432	0,444	0,389	0,4	0,432	0,444	0,389
CC_BODY	0,333	0,333	0,339	0,33	0,467	0,472	0,473	0,464
CC_GENU	0,467	0,25	0,438	0,318	0,467	0,233	0,5	0,318
CC_ROSTRUM	0,467	0,45	0,455	0,444	0,4	0,4	0,411	0,389
CC_SPLENIUM	<b>0,6</b>	<b>0,607</b>	<b>0,611</b>	0,598	0,533	0,5	0,5	0,498
CG2_LEFT	0,467	0,481	0,491	0,4	0,467	0,233	0,5	0,318
CG2_RIGHT	<b>0,6</b>	<b>0,6</b>	0,589	0,583	0,533	0,55	0,545	0,525
CG3_LEFT	0,533	0,5	0,5	0,498	0,467	<b>0,714</b>	0,556	0,4
CG3_RIGHT	0,4	0,442	0,472	0,354	0,4	0,2	0,5	0,286
CG_LEFT	0,467	<b>0,692</b>	<b>0,6</b>	0,444	0,333	0,167	0,5	0,25
CG_RIGHT	0,467	0,444	0,444	0,444	0,467	<b>0,714</b>	0,556	0,4
CST_LEFT	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>	<b>0,667</b>	<b>0,7</b>	<b>0,679</b>	<b>0,661</b>
CST_RIGHT	0,4	0,352	0,384	0,354	0,133	0,134	0,134	0,133
FORNIX_LEFT	0,333	0,306	0,306	0,306	0,267	0,278	0,278	0,267
FORNIX_RIGHT	0,333	0,333	0,339	0,33	0,467	0,233	0,5	0,318
IFO_LEFT	0,4	0,4	0,411	0,389	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>
IFO_RIGHT	<b>0,6</b>	0,598	0,598	0,598	0,467	0,233	0,5	0,318
IL_LEFT	0,333	0,3	0,321	0,306	0,467	0,233	0,5	0,318
IL_RIGHT	<b>0,6</b>	0,598	0,598	0,598	0,533	0,583	0,554	0,498
THAL_FRONT_LEFT	<b>0,667</b>	<b>0,7</b>	<b>0,679</b>	<b>0,661</b>	0,467	0,233	0,5	0,318
THAL_FRONT_RIGHT	0,333	0,33	0,33	0,33	0,4	0,4	0,411	0,389
THAL_MOT_LEFT	0,4	0,375	0,42	0,354	0,467	0,233	0,5	0,318
THAL_MOT_RIGHT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,467	0,233	0,5	0,318
THAL_OCC_LEFT	0,4	0,4	0,411	0,389	0,533	0,519	0,509	0,444
THAL_OCC_RIGHT	0,467	0,444	0,444	0,444	0,4	0,2	0,5	0,286
THAL_PAR_LEFT	0,4	0,417	0,417	0,4	0,4	0,2	0,5	0,286
THAL_PAR_RIGHT	0,267	0,268	0,268	0,267	0,467	0,233	0,5	0,318
THAL_TEMP_LEFT	0,333	0,333	0,339	0,33	0,467	0,45	0,455	0,444
THAL_TEMP_RIGHT	0,4	0,442	0,472	0,354	0,4	0,2	0,5	0,286
UN_LEFT	0,533	0,528	0,527	0,525	0,467	0,233	0,5	0,318
UN_RIGHT	0,533	<b>0,682</b>	<b>0,682</b>	0,533	0,267	0,133	0,5	0,211

**Tabla A.13:** Resultados de modelos de clasificación aplicados a valores ADC y FA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.



	Resultados de modelos de clasificación aplicados a valores ADC y GFA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
<b>AR_ANT_LEFT</b>	<b>0,667</b>	<b>0,694</b>	<b>0,694</b>	<b>0,667</b>	0,4	0,2	0,5	0,286
AR_ANT_RIGHT	0,467	0,455	0,45	0,444	0,333	0,167	0,5	0,25
AR_LEFT	0,467	0,464	0,464	0,464	0,467	0,233	0,5	0,318
AR_RIGHT	0,4	0,352	0,384	0,354	0,4	0,214	0,429	0,286
AR_POST_LEFT	0,533	0,583	0,554	0,498	0,2	0,136	0,214	0,167
<b>AR_POST_RIGHT</b>	<b>0,667</b>	<b>0,821</b>	0,583	0,534	0,4	0,2	0,5	0,286
CC_BODY	0,333	0,3	0,321	0,306	0,533	0,528	0,527	0,525
CC_GENU	<b>0,6</b>	<b>0,611</b>	<b>0,607</b>	0,598	0,467	0,233	0,5	0,318
CC_ROSTRUM	0,467	0,464	0,464	0,464	0,467	0,233	0,5	0,318
CC_SPLENIUM	0,4	0,417	0,417	0,4	0,467	0,444	0,444	0,444
CG2_LEFT	0,4	0,352	0,384	0,354	0,467	0,233	0,5	0,318
<b>CG2_RIGHT</b>	<b>0,6</b>	<b>0,769</b>	<b>0,625</b>	0,55	0,467	0,233	0,5	0,318
CG3_LEFT	0,333	0,35	0,361	0,33	0,4	0,2	0,5	0,286
<b>CG3_RIGHT</b>	<b>0,6</b>	<b>0,75</b>	<b>0,667</b>	0,583	0,4	0,2	0,5	0,286
CG_LEFT	0,333	0,361	0,35	0,33	0,333	0,167	0,5	0,25
<b>CG_RIGHT</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>	0,4	0,2	0,5	0,286
<b>CST_LEFT</b>	<b>0,8</b>	<b>0,85</b>	<b>0,813</b>	<b>0,796</b>	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>
CST_RIGHT	0,267	0,25	0,259	0,253	0,467	0,233	0,5	0,318
<b>FORNIX_LEFT</b>	0,533	0,527	0,528	0,525	<b>0,733</b>	<b>0,739</b>	<b>0,694</b>	<b>0,7</b>
FORNIX_RIGHT	0,333	0,307	0,348	0,306	0,467	0,233	0,5	0,318
<b>IFO_LEFT</b>	<b>0,667</b>	<b>0,67</b>	<b>0,67</b>	<b>0,667</b>	<b>0,6</b>	<b>0,611</b>	<b>0,607</b>	0,598
IFO_RIGHT	0,533	0,519	0,509	0,444	0,467	0,233	0,5	0,318
IL_LEFT	0,333	0,33	0,33	0,33	0,533	0,55	0,545	0,525
IL_RIGHT	0,467	0,472	0,473	0,464	0,467	0,233	0,5	0,318
THAL_FRONT_LEFT	0,467	0,45	0,455	0,444	0,467	0,233	0,5	0,318
<b>THAL_FRONT_RIGHT</b>	<b>0,667</b>	<b>0,667</b>	<b>0,661</b>	<b>0,661</b>	0,533	0,55	0,545	0,525
THAL_MOT_LEFT	0,333	0,307	0,348	0,306	0,467	0,233	0,5	0,318
THAL_MOT_RIGHT	0,467	0,477	0,482	0,444	0,533	0,536	0,536	0,533
THAL_OCC_LEFT	0,533	0,528	0,527	0,525	0,533	0,519	0,509	0,444
THAL_OCC_RIGHT	0,533	0,527	0,528	0,525	0,4	0,2	0,5	0,286
THAL_PAR_LEFT	0,467	0,473	0,472	0,464	0,4	0,2	0,5	0,286
THAL_PAR_RIGHT	0,467	0,464	0,464	0,464	0,467	0,233	0,5	0,318
THAL_TEMP_LEFT	0,467	0,472	0,473	0,464	0,467	0,472	0,473	0,464
THAL_TEMP_RIGHT	0,467	0,542	0,528	0,444	0,4	0,2	0,5	0,286
UN_LEFT	0,533	0,536	0,536	0,533	0,467	0,233	0,5	0,318
<b>UN_RIGHT</b>	0,533	<b>0,682</b>	<b>0,682</b>	0,533	0,267	0,133	0,5	0,211

**Tabla A.14:** Resultados de modelos de clasificación aplicados a valores ADC y GFA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.



	Resultados de modelos de clasificación aplicados a valores FA y GFA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
AR_ANT_LEFT	<b>0,6</b>	0,583	0,583	0,583	0,4	0,2	0,5	0,286
AR_ANT_RIGHT	0,467	0,455	0,45	0,444	0,333	0,167	0,5	0,25
<b>AR_LEFT</b>	0,4	0,4	0,411	0,389	<b>0,6</b>	<b>0,769</b>	<b>0,625</b>	0,55
AR_RIGHT	0,4	0,214	0,429	0,286	0,533	0,583	0,554	0,498
AR_POST_LEFT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,467	0,233	0,5	0,318
AR_POST_RIGHT	0,4	0,35	0,361	0,354	0,4	0,393	0,389	0,389
CC_BODY	0,533	0,536	0,536	0,533	0,467	0,472	0,473	0,464
CC_GENU	0,467	0,417	0,446	0,4	0,467	0,481	0,491	0,4
CC_ROSTRUM	0,4	0,402	0,402	0,4	0,4	0,4	0,411	0,389
CC_SPLENIUM	0,467	0,5	0,5	0,464	0,467	0,444	0,444	0,444
CG2_LEFT	0,267	0,268	0,268	0,267	0,333	0,333	0,339	0,33
<b>CG2_RIGHT</b>	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>	0,467	0,477	0,482	0,444
CG3_LEFT	0,467	0,398	0,417	0,4	0,333	0,333	0,389	0,306
CG3_RIGHT	0,467	0,5	0,5	0,464	0,4	0,2	0,5	0,286
CG_LEFT	0,4	0,5	0,5	0,389	0,4	0,5	0,5	0,389
<b>CG_RIGHT</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>	<b>0,867</b>	<b>0,861</b>	<b>0,861</b>	<b>0,861</b>
<b>CST_LEFT</b>	<b>0,733</b>	<b>0,75</b>	<b>0,741</b>	<b>0,732</b>	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>
CST_RIGHT	0,267	0,25	0,259	0,253	0,333	0,333	0,339	0,33
FORNIX_LEFT	0,4	0,393	0,389	0,389	0,467	0,444	0,444	0,444
FORNIX_RIGHT	0,267	0,268	0,268	0,267	0,267	0,268	0,268	0,267
IFO_LEFT	0,333	0,307	0,348	0,306	0,333	0,333	0,339	0,33
IFO_RIGHT	0,533	0,55	0,545	0,525	0,467	0,233	0,5	0,318
IL_LEFT	0,267	0,268	0,268	0,267	0,533	0,55	0,545	0,525
IL_RIGHT	0,4	0,4	0,411	0,389	0,333	0,307	0,348	0,306
THAL_FRONT_LEFT	0,467	0,45	0,455	0,444	0,467	0,233	0,5	0,318
<b>THAL_FRONT_RIGHT</b>	0,467	0,45	0,455	0,444	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>
THAL_MOT_LEFT	0,333	0,307	0,348	0,306	<b>0,6</b>	<b>0,648</b>	<b>0,616</b>	0,583
THAL_MOT_RIGHT	0,467	0,472	0,473	0,464	0,467	0,472	0,473	0,464
THAL_OCC_LEFT	0,333	0,33	0,33	0,33	0,467	0,45	0,455	0,444
THAL_OCC_RIGHT	<b>0,6</b>	<b>0,607</b>	<b>0,611</b>	0,598	0,467	0,5	0,5	0,464
THAL_PAR_LEFT	0,467	0,473	0,472	0,464	0,4	0,2	0,5	0,286
THAL_PAR_RIGHT	0,533	0,523	0,518	0,498	0,4	0,4	0,411	0,389
THAL_TEMP_LEFT	0,467	0,472	0,473	0,464	0,467	0,45	0,455	0,444
THAL_TEMP_RIGHT	0,533	0,556	0,556	0,533	0,467	0,473	0,472	0,464
UN_LEFT	0,4	0,402	0,402	0,4	0,467	0,233	0,5	0,318
<b>UN_RIGHT</b>	<b>0,733</b>	<b>0,75</b>	<b>0,818</b>	<b>0,722</b>	0,267	0,133	0,5	0,211

**Tabla A.15:** Resultados de modelos de clasificación aplicados a valores FA y GFA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.

	Resultados de modelos de clasificación aplicados a valores ADC, FA y GFA							
	Decision Trees				Support Vector Machine			
	Accuracy	Presicion	Recall	F1 Score	Accuracy	Presicion	Recall	F1 Score
<b>AR_ANT_LEFT</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>	0,4	0,2	0,5	0,286
AR_ANT_RIGHT	0,467	0,455	0,45	0,444	0,333	0,167	0,5	0,25
AR_LEFT	0,467	0,464	0,464	0,464	0,467	0,233	0,5	0,318
AR_RIGHT	<b>0,6</b>	<b>0,625</b>	0,58	0,55	0,4	0,214	0,429	0,286
AR_POST_LEFT	0,2	0,136	0,214	0,167	0,467	0,233	0,5	0,318
AR_POST_RIGHT	0,4	0,432	0,444	0,389	0,333	0,333	0,389	0,306
<b>CC_BODY</b>	<b>0,667</b>	<b>0,693</b>	<b>0,652</b>	<b>0,641</b>	0,467	0,472	0,473	0,464
CC_GENU	<b>0,6</b>	<b>0,6</b>	0,589	0,583	0,467	0,233	0,5	0,318
CC_ROSTRUM	0,467	0,472	0,473	0,464	0,4	0,4	0,411	0,389
CC_SPLENIUM	0,4	0,417	0,417	0,4	0,467	0,444	0,444	0,444
CG2_LEFT	0,467	0,45	0,455	0,444	0,467	0,233	0,5	0,318
CG2_RIGHT	0,533	0,536	0,536	0,533	0,267	0,25	0,277	0,253
CG3_LEFT	<b>0,6</b>	0,583	0,583	0,583	0,4	0,2	0,5	0,286
CG3_RIGHT	0,467	0,542	0,528	0,444	0,4	0,2	0,5	0,286
CG_LEFT	0,4	0,411	0,4	0,389	0,333	0,167	0,5	0,25
<b>CG_RIGHT</b>	<b>0,667</b>	<b>0,661</b>	<b>0,667</b>	<b>0,661</b>	<b>0,6</b>	<b>0,75</b>	<b>0,667</b>	0,583
<b>CST_LEFT</b>	<b>0,733</b>	<b>0,75</b>	<b>0,741</b>	<b>0,732</b>	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>
CST_RIGHT	0,267	0,25	0,259	0,253	0,267	0,25	0,277	0,253
FORNIX_LEFT	0,4	0,432	0,444	0,389	0,4	0,393	0,389	0,389
FORNIX_RIGHT	0,467	0,477	0,482	0,444	0,467	0,233	0,5	0,318
IFO_LEFT	0,4	0,4	0,411	0,389	0,533	0,536	0,536	0,533
IFO_RIGHT	<b>0,6</b>	<b>0,6</b>	0,589	0,583	0,467	0,233	0,5	0,318
IL_LEFT	0,267	0,268	0,268	0,267	0,533	0,55	0,545	0,525
IL_RIGHT	0,467	0,45	0,455	0,444	0,533	0,583	0,554	0,498
THAL_FRONT_LEFT	0,467	0,45	0,455	0,444	0,467	0,233	0,5	0,318
<b>THAL_FRONT_RIGHT</b>	0,4	0,402	0,402	0,4	<b>0,733</b>	<b>0,818</b>	<b>0,75</b>	<b>0,722</b>
THAL_MOT_LEFT	0,333	0,3	0,321	0,306	0,467	0,233	0,5	0,318
THAL_MOT_RIGHT	0,467	0,472	0,473	0,464	0,467	0,472	0,473	0,464
THAL_OCC_LEFT	0,333	0,33	0,33	0,33	0,467	0,417	0,446	0,4
THAL_OCC_RIGHT	0,533	0,527	0,528	0,525	0,4	0,2	0,5	0,286
THAL_PAR_LEFT	0,533	<b>0,731</b>	<b>0,611</b>	0,498	0,4	0,2	0,5	0,286
THAL_PAR_RIGHT	0,467	0,464	0,464	0,464	0,4	0,4	0,411	0,389
THAL_TEMP_LEFT	0,333	0,333	0,339	0,33	0,533	0,536	0,536	0,533
THAL_TEMP_RIGHT	0,333	0,179	0,417	0,25	0,467	0,473	0,472	0,464
UN_LEFT	0,467	0,464	0,464	0,464	0,467	0,233	0,5	0,318
UN_RIGHT	0,467	<b>0,667</b>	<b>0,636</b>	0,464	0,267	0,133	0,5	0,211

**Tabla A.16:** Resultados de modelos de clasificación aplicados a valores ADC, FA y GFA, por fascículo. Aquellos valores que son mayores o iguales a 0.6, es decir, 60 %, se marcan con celeste, los valores mayores a 2/3, es decir, el 66 % se marcan en azul, y aquellos valores iguales o mayores a un 75 % se marcan en color rojo.

Hiperparámetros de DT				
	MD	MF	MSL	MSS
FA:				
CST_LEFT	None	sqrt	1	10
THAL_FRONT_LEFT	None	sqrt	1	10
UN_RIGHT	None	sqrt	1	2
GFA:				
AR_ANT_LEFT	None	sqrt	1	5
AR_POST_RIGHT	None	sqrt	2	2
CG3_RIGHT	None	sqrt	2	2
CG_RIGHT	None	sqrt	1	5
CST_LEFT	None	sqrt	1	2
IFO_LEFT	None	sqrt	1	2
THAL_FRONT_LEFT	None	sqrt	1	10
UN_RIGHT	None	sqrt	1	10
A-F:				
CST_LEFT	None	sqrt	1	10
THAL_FRONT_LEFT	None	sqrt	1	10
UN_RIGHT	None	sqrt	1	2
A-G:				
AR_ANT_LEFT	None	sqrt	1	5
AR_POST_RIGHT	None	sqrt	2	2
CG3_RIGHT	None	sqrt	2	2
CG_RIGHT	None	sqrt	1	5
CST_LEFT	None	sqrt	1	2
UN_RIGHT	None	sqrt	1	10
F-G:				
CG2_RIGHT	None	sqrt	2	5
CG_RIGHT	None	log2	1	5
CST_LEFT	None	log2	1	2
UN_RIGHT	15	sqrt	1	2
A-F-G:				
AR_ANT_LEFT	10	sqrt	1	10
CC_BODY	None	log2	2	5
CG_RIGHT	None	log2	1	5
CST_LEFT	None	log2	2	5

**Tabla A.17:** Hiperparámetros destacados del modelo Decision Trees para los fascículos relevantes para cada medida. Sea MD = profundidad máxima, MF = máximas características, MSL = mínimo de muestras por hoja y MSS = mínimo de muestras para dividir.

	Hiperparámetros de SVM		
	C	G	K
ADC:			
CG3_LEFT	100	scale	rbf
CG_RIGHT	1	scale	poly
THAL_PAR_RIGHT	100	scale	rbf
FA:			
AR_ANT_LEFT	100	scale	rbf
CG_RIGHT	1	scale	poly
THAL_OCC_RIGHT	10	scale	rbf
GFA:			
AR_POST_RIGHT	0,1	scale	linear
CG2_RIGHT	1	scale	rbf
CST_LEFT	100	scale	rbf
AF:			
CST_LEFT	0,1	scale	linear
IFO_LEFT	0,1	scale	linear
AG:			
CST_LEFT	0,1	scale	linear
FORNIX_LEFT	10	scale	poly
FG:			
CG_RIGHT	100	scale	rbf
CST_LEFT	0,1	scale	linear
THAL_FRONT_RIGHT	100	scale	poly
AFG:			
CG_RIGHT	1	scale	poly
CST_LEFT	0,1	scale	linear
THAL_FRONT_RIGHT	10	scale	poly

**Tabla A.18:** Hiperparámetros destacados del modelo Support Vector Machine para los fascículos relevantes de cada medida. Sea G = gamma, y K = kernel.

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

<b>Departamento</b>	: Departamento de Ingeniería Eléctrica
<b>Carrera</b>	: Ingeniería Civil Biomédica
<b>Nombre del memorista</b>	: Natalia Valentina Vidal Barraza
<b>Título de la memoria</b>	: Análisis de Imágenes de Resonancia Magnética por Difusión de Adolescentes con Autismo
<b>Fecha de la presentación oral</b>	: 21 de Agosto 2023
<b>Profesor(es) Guía</b>	: Pamela Guevara A.
<b>Profesor(es) Revisor(es)</b>	: Cecilia Hernandez, Jaime Pinto
<b>Concepto</b>	:
<b>Calificación</b>	:

**Resumen**

En las últimas décadas, el uso de neuroimágenes ha crecido significativamente, proporcionando diversas técnicas e investigaciones para el estudio, diagnóstico y seguimiento de trastornos cerebrales. Este estudio se enfoca en Imágenes de Resonancia Magnética por Difusión de adolescentes con y sin autismo. A través de un análisis detallado, se busca identificar diferencias en la materia blanca entre los grupos. Esto podría mejorar la comprensión y tratamiento del autismo, permitiendo diagnósticos precisos y terapias tempranas.

Se desarrollaron algoritmos para procesar datos de dMRI y de tractografías. Estos algoritmos transforman los datos, segmentan fibras en 36 fascículos, obtienen valores medios en mapas de intensidad y aplican pruebas estadísticas y Machine Learning para identificar diferencias.

Aunque los resultados de las pruebas estadísticas no coinciden del todo con las conclusiones de clasificación, se encontró información relevante y diferencias en áreas cercanas al cuerpo caloso en sujetos con autismo.

Se sugiere realizar más estudios en el futuro, ampliando la base de datos y considerando aspectos específicos del espectro autista, como comunicación, comportamientos repetitivos y sensibilidades sensoriales. Esto podría brindar resultados más concluyentes y precisos para comprender mejor el autismo.