



Proyecto Tópicos Selectos de Ciencia de Datos: Deep Learning

Métodos de DL para Análisis de Tractografía Mediante Datos de dMRI.

Autores: David Alejandro Muro Campa, Gabriel Alejandro Aguilar Farrera

Maestría en Cómputo Estadístico, CIMAT Sede Monterrey

11 de diciembre de 2023

Resumen: La tractografía por resonancia magnética ponderada en difusión, combinada con la selección de trayectorias basada en regiones (bundles), es una combinación única de herramientas que permite la delineación y análisis in vivo de tractos anatómicamente conocidos. Sin embargo, estos procesos son complejos y computacionalmente intensivos. En este proyecto se utiliza como modelo base a “TractSeg”, estado del arte para segmentar directamente los tractos mediante las direcciones predominantes de la difusión del agua. La base de datos que con la que se trabajó corresponde a la de Renaud 2023, la cual contiene 21 bundles (ground truth) de un sujeto que fueron segmentadas por expertos. En este trabajo, exploramos el problema de segmentación cuando se requieren bundles que no formaron parte del modelo preentrenado TractSeg y cuando se tiene únicamente un paciente para entrenamiento. Los resultados aquí expuestos nos llevan a concluir que el problema de segmentación con un solo paciente es un reto muy desafiante y bastante común que puede tratarse con la metodología aquí mostrada pues se obtuvieron *dice scores* aceptables para la complejidad del problema.

1. Introducción

1.1. Antecedentes

La tractografía de materia blanca es un procedimiento que se usa para poner en manifiesto los tractos neuronales, así como para determinar la conectividad cerebral. Utiliza técnicas especiales de imagen por resonancia magnética (MRI por sus siglas en inglés) y de procesamiento de imágenes.

Determinar la conectividad cerebral es uno de los principales objetivos médicos del análisis moderno de imágenes de resonancia magnética dado que este conocimiento se relaciona con el estado de algunas enfermedades fisiológicas del cerebro (como la apoplejía o embolia cerebral) y también puede ser usada para investigar desórdenes neurológicos (por ejemplo problemas del lenguaje) y en general el desarrollo cerebral.

Ahora bien, la segmentación de tractos de fibras de materia blanca permite un análisis más detallado de los tractos individuales de materia blanca. Ayuda a caracterizar el cerebro sano e identificar áreas que contienen morfología anormal en cerebros enfermos (Bells

et al., 2011; Yendiki et al., 2011; Yeatman et al., 2012; Dayan et al., 2016). Tales análisis de tractometría dependen inherentemente de la calidad de la segmentación del tracto subyacente, y por lo tanto requieren una delimitación precisa.

Actualmente, el enfoque más comúnmente empleado para segmentar los tractos de materia blanca es la *disección virtual*: las trayectorias que corresponden a tractos anatómicamente bien definidos se extraen manualmente de un tractograma utilizando combinaciones de inclusión y exclusión de regiones de interés (ROIs) (Wang et al., 2007; Stieltjes et al., 2013; Thiebaut de Schotten et al., 2011). Este es un proceso muy laborioso que tiene una reproducibilidad limitada debido a la subjetividad en la interacción humana. Por lo tanto, se han propuesto una variedad de enfoques automáticos para la segmentación de materia blanca. Se pueden agrupar en tres categorías: *segmentación basada en ROI*, *segmentación basada en agrupamiento* y *segmentación directa*.

En este trabajo, ocupamos TractSeg [1], el cual es un enfoque basado en redes neuronales para la segmentación directa de tractos de materia blanca que proporciona segmentaciones completas y precisas que no requiere registro, parcelación, tractografía o agrupamiento.

1.2. Objetivos

- Utilizar el modelo TractSeg, el cual es un modelo de segmentación de tractografía por resonancia magnética ponderada en difusión (dMRI) para la identificación precisa de tractos de materia blanca, abordando la segmentación de bundles no cubiertos por modelos preentrenados.
- Evaluar la efectividad del modelo en la a segmentación de bundles no cubiertos inicialmente, utilizando métricas de evaluación como el Dice Score para validar la precisión de la segmentación.

2. Marco Teórico

2.1. TractSeg

En esta sección se presenta el marco de trabajo de TractSeg, con qué datos se entrenó, cuál es la entrada y salida del modelo, así como aspectos más particulares.

TractSeg es una red neuronal completamente convolucional (FCNN) que segmenta directamente los tractos de materia blanca en campos de picos de la función de distribución de orientación de las fibras (fODF). Este modelo fue entrenado con 72 tractos anatómicamente bien descritos en un cohorte de 105 sujetos seleccionados del Proyecto de Conectoma Humano (HCP) (Van Essen et al., 2013). La Fig. (1) muestra el flujo de trabajo de TractSeg. Aunque los valores de las imágenes en crudo podrían haber sido utilizados como entrada, para conjuntos de datos de alta resolución angular, esto habría resultado en una imagen de entrada con un gran número de canales, generando una alta demanda de memoria y una lenta entrada/salida de archivos durante el entrenamiento. Por tal motivo se eligió una representación más condensada de los datos para mitigar este problema: TractSeg espera recibir las tres direcciones principales de las fibras por vóxel como entrada, lo que requiere nueve canales diferentes de entrada (tres por dirección principal).

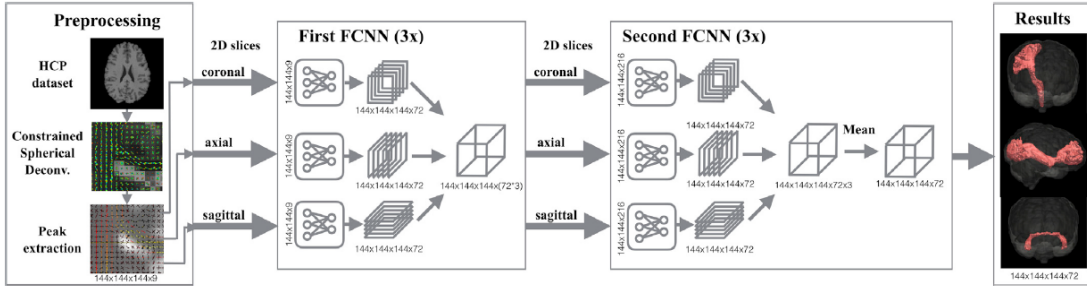


Figura 1: Flujo de trabajo de TractSeg. Se utiliza la Deconvolución Esférica Restringida para extraer las tres direcciones dominantes de difusión en cada vóxel. Luego, un FCNN encoder-decoder en 2D produce una imagen de probabilidad de tracto para cada orientación (coronal, axial, sagital) y para cada tracto. Las imágenes de probabilidad de tracto de las tres orientaciones se concatenan en la dimensión de los canales, lo que resulta en una imagen en 3D con 216 canales. Esta imagen se utiliza como entrada para un segundo FCNN que, nuevamente, se ejecuta tres veces. Los tres resultados por tracto del segundo FCNN se fusionan utilizando la media para generar la segmentación final. La segmentación final es una imagen de 72 canales, donde cada canal contiene las probabilidades de vóxel para un tracto específico.

Si un vóxel contenía solo una dirección de fibra, por ejemplo, vóxeles en el cuerpo calloso, entonces el segundo y tercer pico se establecían en cero. Las imágenes de HCP tienen una resolución espacial de $145 \times 174 \times 145$ vóxeles, pero se recortaron a $144 \times 144 \times 144$ sin eliminar ningún tejido cerebral para que se ajusten al tamaño de entrada de la red.

La arquitectura propuesta de la red neuronal convolucional totalmente convolucional (FCNN) de encoder-decoder en 2D fue inspirada por la arquitectura de U-Net previamente propuesta por Ronneberger et al. (2015). La Figura (2) detalla la configuración de red propuesta.

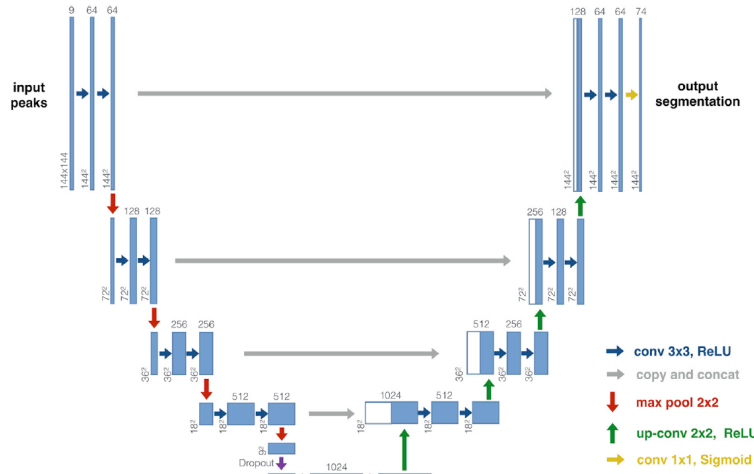


Figura 2: Arquitectura U-Net.

La entrada para la red propuesta es una imagen en 2D de 144×144 píxeles y 9 canales correspondientes a los 3 picos por píxel (cada pico está representado por un vector tridimensional). La salida es una imagen con múltiples canales (72 canales) y dimensiones espaciales de 144×144 píxeles, donde cada canal contiene las probabilidades de píxeles para un tracto. Aunque en principio la arquitectura U-Net permite extensiones a segmentación de imágenes en 3D (Çiçek et al., 2016), proponen que se utilicen cortes en 2D como

entrada para aumentar la eficiencia de memoria y permitir que los datos se procesen a la resolución original del HCP. Para aprovechar la información adicional proporcionada por la tercera dimensión, se muestrean aleatoriamente cortes en 2D en tres orientaciones diferentes durante el entrenamiento: axial, coronal y sagital. Esto significaba que el modelo aprendía a trabajar con las tres orientaciones.

Durante la inferencia, se generaron tres predicciones por píxel por tracto, una para cada orientación, lo que resultó en una imagen con dimensiones de $144 \times 144 \times 144 \times 72 \times 3$ (después de ejecutar el modelo 1443 veces) y, si se concatenan lo largo de las últimas 2 dimensiones, se obtiene $144 \times 144 \times 144 \times 216$ (una imagen en 3D con 216 canales). Se entrenó un segundo FCNN con tres canales de entrada por tracto (es decir, $3 \times 72 = 216$ canales de entrada) para fusionar óptimamente las tres predicciones por píxel. Sin embargo, dado que el segundo FCNN también funciona en cortes en 2D, se pueden muestrear en tres orientaciones diferentes (axial, coronal, sagital) aquí también. El segundo FCNN también se ejecutó tres veces, una vez para cada orientación, dando tres predicciones por píxel por tracto. Simplemente se toma la media para fusionar esos resultados en una segmentación final.

2.2. Entrenamiento

El entrenamiento se realizó utilizando la pérdida de entropía cruzada binaria. Para un objetivo dado t , una salida de la red o y un número n de clases, la pérdida se calculó de la siguiente manera:

$$loss(o, t) = -\frac{1}{n} \sum_{i=0}^n (t[i] \log(o[i]) + (1 - t[i]) \log(o[i])) \quad (1)$$

Se utilizaron unidades lineales rectificadas (ReLU) como no linealidad (Nair y Hinton, 2010). Las funciones de activación sigmoideas solo se usaron en la última capa. La activación sigmoide emparejada con entropía cruzada binaria permite segmentaciones multietiqueta. Se utilizó una tasa de aprendizaje de 0.002 y se implementó Adamax (Kingma y Ba, 2014) como optimizador. El tamaño del lote fue de 56. Se utilizó dropout (Srivastava et al., 2014) con una probabilidad del 0.4. Todos los hiperparámetros se optimizaron en un conjunto de datos de validación independiente del conjunto de datos de prueba final. Los pesos de la red de la época con el puntaje de Dice más alto durante la validación se utilizaron para las pruebas.

Para mejorar la capacidad de generalización de TractSeg, aplicaron una intensa aumentación de datos a las imágenes de picos durante el entrenamiento. Se aplicaron las siguientes transformaciones a cada muestra de entrenamiento, variando aleatoriamente la intensidad

de cada transformación.

Rotación:	$\phi_x \sim U\left[\frac{\pi}{4}, \frac{\pi}{4}\right]; \phi_y \sim U\left[\frac{\pi}{4}, \frac{\pi}{4}\right]; \phi_z \sim U\left[\frac{\pi}{4}, \frac{\pi}{4}\right]$
Deformación elástica:	$(\alpha, \sigma) \sim (U[90, 120], U[9, 11])$
Desplazamiento:	$(\Delta x, \Delta y) \sim (U[-10, 10], U[-10, 10])$
Zoom:	$\lambda \sim U[0, 9, 1, 5]$
Re-muestreo:	$\lambda \sim U[0, 5, 1]$
Ruido Gaussiano:	$(\mu, \sigma) \sim (0, U[0, 05])$
Aumento de contraste:	$X_{\text{aumentado}} = (X \cdot \text{avg}(X)) * \beta + \text{avg}(X); \beta \sim U[0, 7, 1, 3]$
Aumento de brillo:	$X_{\text{aumentado}} = X \cdot \gamma; \gamma \sim U[0, 7, 1, 3]$

El conjunto de entrenamiento consistió de 105 sujetos del Proyecto de Conectoma Humano (HCP, por sus siglas en inglés) para los siguientes experimentos. Las imágenes correspondientes de difusión de resonancia magnética (dMRI) tienen una resolución isotrópica de 1.25 mm y 270 direcciones de gradiente con 3 valores b (1000, 2000, 3000 s/mm²) y 18 imágenes con b = 0 (Sotiropoulos et al., 2013).

3. Metodología

3.1. Descripción de los datos

Los datos corresponden a la base de datos Renauld 2023, los cuales se obtuvieron en el “ISMRM 2015 Tractography Challenge”. Los datos contienen 21 bundles o máscaras (archivos con extensión .nii), tractografía por cada uno de los bundles (archivos con extensión .tck), la imagen pesada por difusión (DWI, por sus siglas en inglés), la tabla de gradientes (archivo con extensión .bvec) y finalmente el archivo de b-valores (archivo con extensión .tck). A continuación en las Figuras (3) y (4) se muestran los bundles con los que se cuentan (en las imágenes que dicen “Der-Izq” se están mostrando dos bundles en una sola imagen):

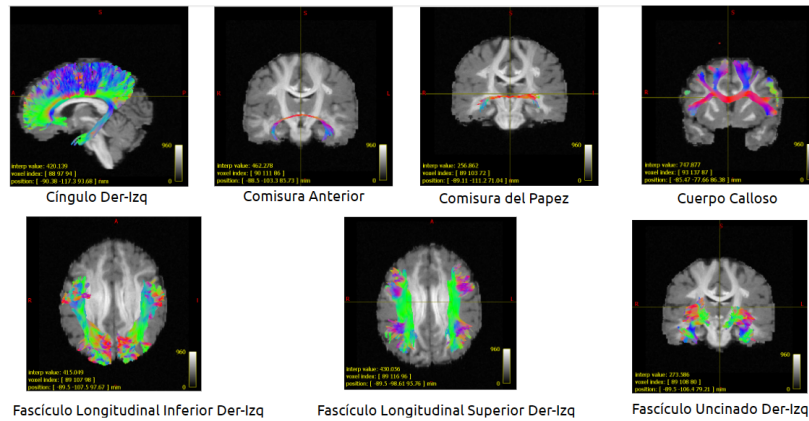


Figura 3: Bundles con los que se cuentan (grupo 1).

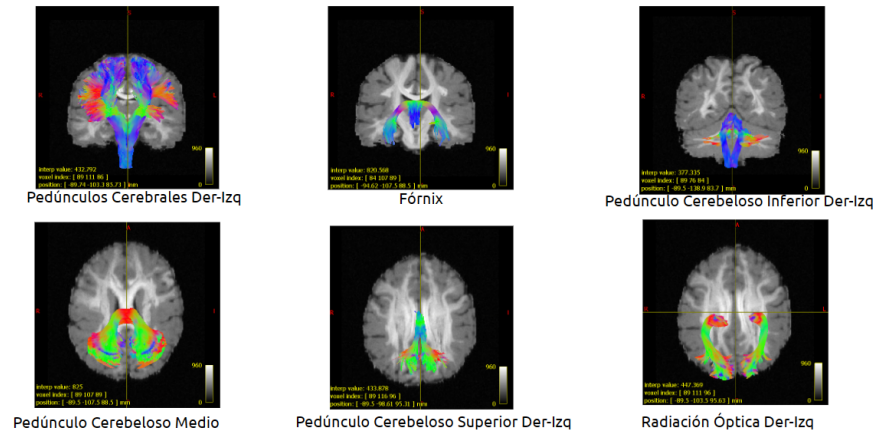
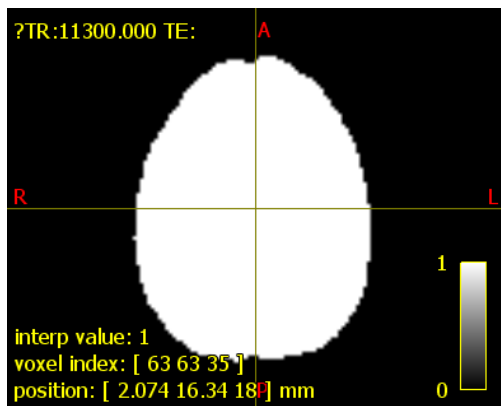
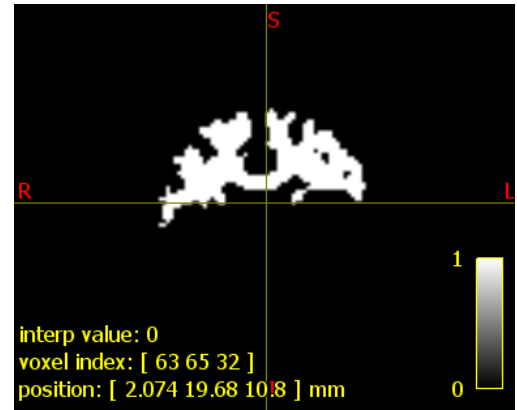


Figura 4: Bundles con los que se cuentan (grupo 2).

En las Figuras (3) y (4) se muestran las tractografías generadas en la sección que corresponden a un bundle en particular. El bundle en sí mismo es una máscara binaria que representa la zona del cerebro de interés. Por ejemplo en la Fig.(5a) se muestra la máscara binaria del cerebro completo (una rebanada en particular), mientras que en la Fig. (5b) se muestra la máscara binaria o bundle correspondiente al cuerpo calloso.



(a) Máscara del cerebro completo (una rebanada en particular)



(b) Máscara de la región del cerebro correspondiente al cuerpo calloso (una rebanada en particular).

Figura 5: Izquierda: Máscara binaria del cerebro compelpeto. Derecha: Máscara binaria del cuerpo calloso.

De igual forma, en la Fig (6) se muestra la vista coronal de la DWI (imagen de resonancia magnética por difusión) de 3 rebanadas en particular del cerebro. En esta imagen se puede apreciar el cuerpo calloso (región más blanca) y la materia gris (región más oscura).

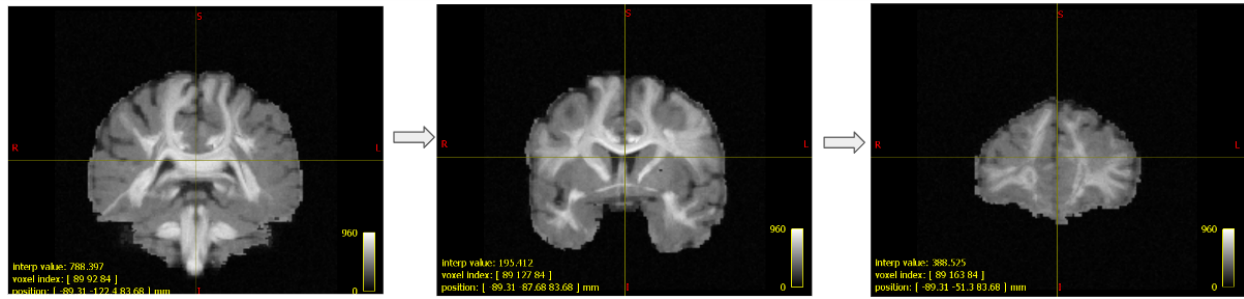


Figura 6: Bundles con los que se cuentan (grupo 2).

El conjunto de datos de desafío de estilo clínico consiste en una adquisición de difusión isotrópica de 2 mm, con 32 direcciones de gradiente, un valor $b=1000 \text{ s/mm}^2$. También contiene una imagen $b=0$ y un volumen opcional $b=0$, con dirección de codificación de fase invertida.

1. Adquisición de Datos de Preentrenamiento:

- Descargar datos del Human Connectome Project (HCP) que serán los datos que utilizaremos para el preentrenamiento.
- Utilizar la imagen de difusión de los datos descargados para calcular los picos utilizando la técnica de “Constrained Spherical Deconvolution” (CSD).
- Una vez obtenidos los picos podemos obtener las segmentaciones de los bundles con TractSeg. TractSeg nos da 72 bundles.

2. Identificación de Bundles:

- Analizar las segmentaciones obtenidas y clasificar los bundles en dos categorías: “bundles existentes” y “bundles novel”.
- Bundles existentes son aquellos bundles de nuestros datos que TractSeg ya segmenta.
- Bundles novel son aquellos bundles de nuestros datos que TractSeg no segmenta.

Bundles Existentes	Bundles Novel
CA	BPS_left
CC	BPS_right
ICP_left	CP
ICP_right	Fornix
ILF_left	SLF_left
ILF_right	SLF_right
MCP	Cingulum_left
OR_left	Cingulum_right
OR_right	-
SCP_left	-
SCP_right	-
UF_left	-
UF_right	-

Cuadro 1: Tabla de Bundles Existentes y Novel.

3. Extracción de Datos para el Preentrenamiento y el Entrenamiento:

- Extraer los picos y las segmentaciones correspondientes a los bundles existentes y novel de los datos de preentrenamiento.
- Organizar los datos para que cada paciente tenga los picos y los bundles existentes, estableciendo así los conjuntos de preentrenamiento. En la siguiente Figura se puede apreciar como ordenamos nuestros datos

```
Pretraining/HCP_preproc/201818/  
    '-> mrtrix_peaks.nii.gz  
    '-> bundle_masks_13.nii.gz  
Pretraining/HCP_preproc/205220/  
    '-> mrtrix_peaks.nii.gz  
    '-> bundle_masks_13.nii.gz  
...
```

Figura 7: Datos de Preentrenamiento.

- Organizar los datos de nuestro único paciente con los picos y los bundles novel como se ve en la siguiente Figura

```
Training/HCP_preproc/Renauld/  
    '-> mrtrix_peaks.nii.gz  
    '-> bundle_masks_8.nii.gz
```

Figura 8: Datos de Preentrenamiento.

3.2. 4. Fase de Preentrenamiento

- Una vez acomodados los datos podemos realizar el preentrenamiento de tal forma que obtenemos un modelo vía TractSeg como modelo base que segmenta únicamente en los 13 bundles existentes. El preentrenamiento se realiza siguiendo la estrategia que se utiliza en [TRACTSEG ORIGINAL]. TractSeg no está optimizado para realizar el entrenamiento de tu propio modelo por lo que se deben seguir una serie de pasos que se enlistan a continuación
 1. Realizar el procesamiento con preprocessing.py en la que hay que adaptar los paths
 2. Adaptar los paths en el archivo my_custom_experiment.py
 3. Crear un archivo config.txt con el nombre de la carpeta donde están tus datos
 4. Adaptar la función get_bundle_names() y get_labels_filename del script datasetspecifutils.py donde establezcas los nombres de los bundles existentes y novel.
 5. Adaptar el archivo subjects.py con los id de los pacientes que vas a utilizar
 6. Adaptar ExpRunner con los paths de tus datos
 7. Ejecutar ExpRunner

5. Fase de Entrenamiento (Fine Tuning):

- Aplicar data augmentation mediante Random Cutout [CHINOS] a nuestro único paciente para generar 63 nuevas instancias de pacientes. En la siguiente Figura apreciamos como se debe organizar la carpeta

```
Training/HCP_preproc/daug_01/  
    '-> mrtrix_peaks.nii.gz  
    '-> bundle_masks_8.nii.gz  
Training/HCP_preproc/daug_02/  
    '-> mrtrix_peaks.nii.gz  
    '-> bundle_masks_8.nii.gz  
...
```

Figura 9: Datos para el Entrenamiento.

- Asignar 48 instancias para el conjunto de entrenamiento y 15 para el conjunto de validación.
- Realizar fine-tuning siguiendo la estrategia de [CHINOS] para ajustar el modelo a la segmentación de los 18 bundles novel. Como su modelo base es TractSeg igual el fine-tuning no está optimizado por lo que se deben hacer una serie de pasos:
 1. Modificar el archivo TractSeg_Fewshot_Pretrain_base.py indicando la fase de fine-tuning, el path del modelo preentrenado y de los datos que usaremos.
 2. Adaptar el archivo TractSeg_Fewshot_Pretrain_ExpRunner.py con los ID de los pacientes que utilizaremos
 3. Ejecutar ExpRunner.

6. Evaluación del Modelo:

- Para testear el modelo lo evaluamos en los picos de nuestro paciente de entrenamiento ya que no tenemos otros datos para realizar un test. La justificación es que el modelo se entrenó con transformaciones de éste y no con esos datos en sí.

4. Resultados

A continuación se muestran los resultados obtenidos en los bundles novel. Como métrica de evaluación usamos el Dice Score ([1], [2]), el cual tiene la siguiente ecuación:

$$Dice = \frac{2|S_g^1 \cap S_t^1|}{|S_g^1| + |S_t^1|} \quad (2)$$

Donde S_g^1 es la segmentación de referencia (ground truth) y S_t^1 es la segmentación de prueba. En la Fig. (10) se pueden observar los valores de dice score para c/u de los bundles novel. Vemos que el bundle con el dice score más alto es el CP o comisura del papez y muy probablemente esto se deba a la geometría simple que tiene dicho bundle (ver Fig. 3). Por otro lado, el bundle con el valor de dice score más bajo es el BPS_right seguido del BPS_left o los pedúnculos cerebrales y esto muy probablemente se deba a su morfología más compleja en comparación al CP (ver Fig. 3),

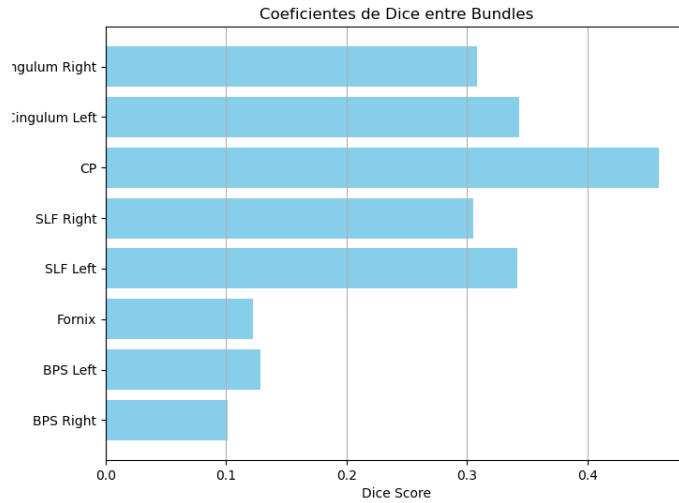


Figura 10: Valor de los Dice Score para c/u de los bundles novel.

En la Fig. (11) se puede ver el bundle de CP original de Renault y el bundle segmentado por TractSeg.

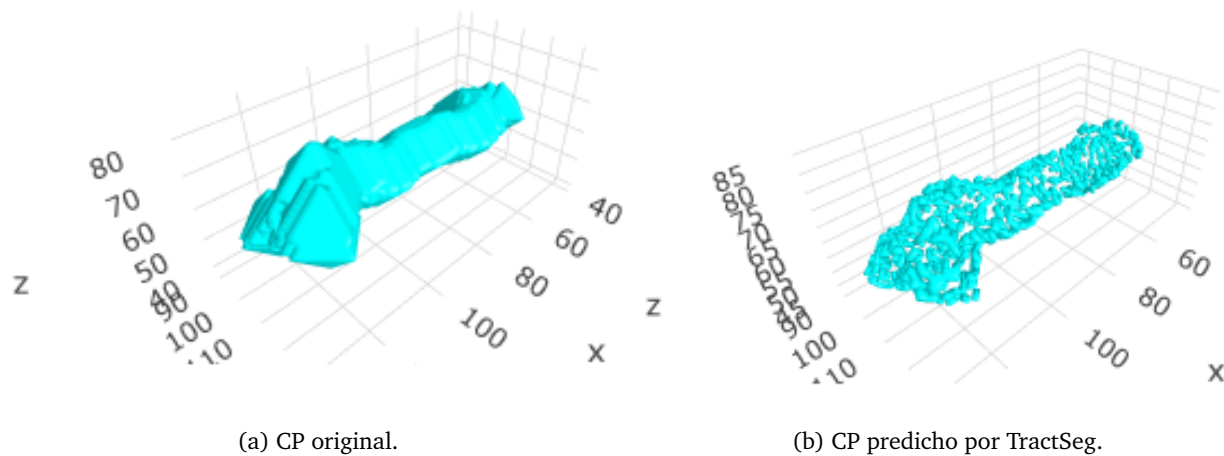
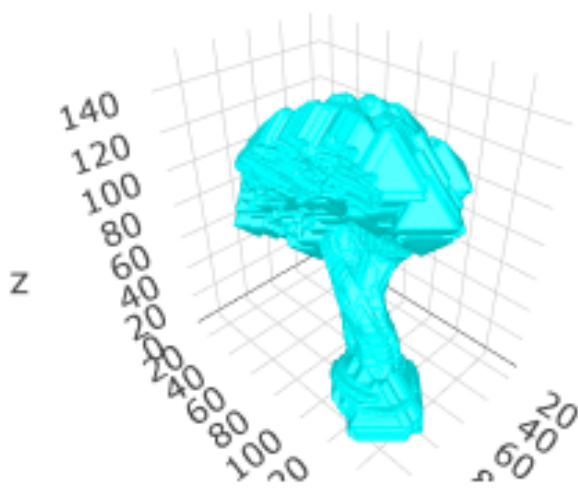
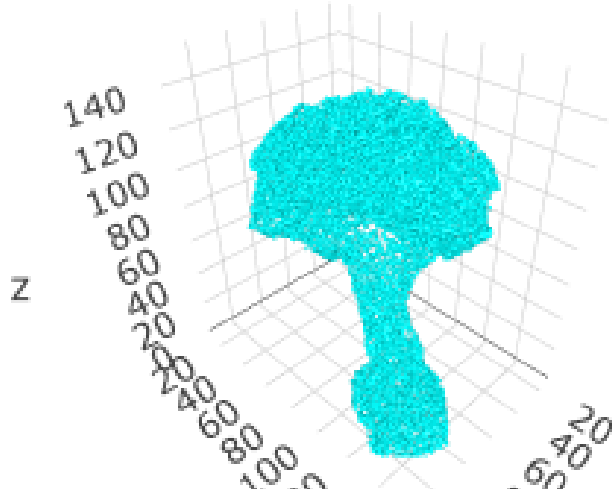


Figura 11: Izquierda: Comisura del Papez original. Derecha: Comisura del papez obtenida por TractSeg. Dice Score = 0.459.

Y finalmente, en la Fig. (12) se puede observar el bundle del pedúnculo cerebral derecho original y el predicho por TractSeg.



(a) BPS right original.



(b) BPS right predicho por TractSeg.

Figura 12: Izquierda: pedúnculo cerebral derecho original. Derecha: pedúnculo cerebral derecho obtenida por TractSeg. Dice Score = 0.101.

Podemos observar que la morfología entre estos dos bundles (Figuras 11 y 12) son muy distintas, por un lado el CP tiene una forma muy sencilla y ocupa un volumen pequeño en comparación al volumen que ocupa el bundle BPS right, además de que la morfología de este último es mucho más complejo y elaborado.

5. Conclusiones

- La metodología, centrada en la identificación y segmentación de "bundles existentes" y "bundles novel", ha demostrado ser una aproximación valiosa, especialmente en situaciones donde la disponibilidad de datos de entrenamiento es limitada, como en el caso de contar con un único paciente.
- La mejora del dice score para datos limitados es aún objeto de investigación pues un problema muy frecuente al que se enfrentan los investigadores y profesionistas en el campo de la neurología y neuroimagen.
- La morfología y tamaño de los bundles influyen mucho en el valor del dice score.

Referencias

- [1] Wasserthal, J., Neher, P., & Maier-Hein, K. H. (2018). TractSeg-Fast and accurate white matter tract segmentation. *NeuroImage*, 183, 239-253.
- [2] Liu, W., Lu, Q., Zhuo, Z., Liu, Y., & Ye, C. (2022, September). One-Shot Segmentation of Novel White Matter Tracts via Extensive Data Augmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 133-142). Cham: Springer Nature Switzerland.