

SEGMENTACIÓN MEDIANTE INTELIGENCIA ARTIFICIAL
PARA LA DETECCIÓN DE LESIONES POR TRAUMA
CRANEOENCEFÁLICO CERRADO A TRAVÉS DE IMÁGENES
POR RESONANCIA MAGNÉTICA

JENIFER CASTAÑO LEDESMA



Escuela de Ingeniería de Sistemas y Computación
Facultad de Ingeniería
Universidad del Valle

Febrero 2026

SEGMENTACIÓN MEDIANTE INTELIGENCIA ARTIFICIAL
PARA LA DETECCIÓN DE LESIONES POR TRAUMA
CRANEOENCEFÁLICO CERRADO A TRAVÉS DE IMÁGENES
POR RESONANCIA MAGNÉTICA

JENIFER CASTAÑO LEDESMA

202067525

JENIFER.CASTANO@CORREOUNIVALLE.EDU.CO

Directora:

DEISY CHAVES SANCHEZ, PH.D

Co-Directores:

MARÍA PATRICIA TRUJILLO URIBE, PH.D

ALEJANDRO HERRERA TRUJILLO, MSC

PEDRO ALEJANDRO AMADOR RODRÍGUEZ, MSC



Escuela de Ingeniería de Sistemas y Computación
Facultad de Ingeniería
Universidad del Valle

Febrero 2026

Llegar al final de este trayecto ha sido posible por la presencia de las personas que caminaron a mi lado. Agradezco a mi familia y amigos por su compañía, amor y confianza.

En especial, a mis padres, por estar siempre presentes, por escucharme, y darme ánimo en los momentos de dificultad. A mi hermano, por su apoyo constante, por motivarme a crecer, aprender y superarme, así como por estar a mi lado tanto en las dificultades como en los triunfos.

Asimismo, agradezco a todos los profesores que contribuyeron a mi formación académica durante toda la carrera, por los conocimientos y experiencias brindadas. De manera especial, a la profesora Deisy, por su acompañamiento, apoyo y orientación durante el desarrollo de este trabajo.

RESUMEN

El Trauma Craneoencefálico Cerrado (TCEC) genera una o múltiples lesiones que afectan las funciones cerebrales de forma transitoria o permanente, siendo una de las principales causas de discapacidad y mortalidad a nivel mundial. La segmentación manual de dichas lesiones a partir de Imágenes por Resonancia Magnética (IRM) es un proceso lento, subjetivo y propenso a errores, lo que limita la obtención de un diagnóstico adecuado. Por ende, resulta necesario automatizar el proceso mediante técnicas de aprendizaje profundo.

En este trabajo de grado se propone una estrategia para la segmentación automática de lesiones asociadas al TCEC en IRM ponderadas en T1. Para ello, se estableció un preprocesamiento que incluye la remoción del cráneo, el alineamiento de todas las imágenes a una plantilla de referencia estándar, el recorte basado en la región de interés, la normalización y el remuestreo. Asimismo, se evaluaron diferentes métodos de normalización con el fin de analizar su efecto en el desempeño del modelo. Posteriormente, se entrenaron y evaluaron dos modelos basados en la arquitectura U-Net: nnU-Net y 3D U-Net. Además, se realizó un experimento con nnU-Net para analizar el desempeño utilizando aprendizaje por transferencia. El rendimiento de ambos se midió usando métricas como el Coeficiente de Dice e Intersección sobre la Unión.

Los resultados muestran que el modelo que obtuvo el mejor desempeño es nnUNet junto con el método de normalización KDE, y que el aprendizaje por transferencia permitió mejorar significativamente las métricas de evaluación. Finalmente, se implementó un framework que integra todo el proceso de segmentación, proporcionando una herramienta para apoyar la práctica clínica y futuras investigaciones relacionadas con el TCEC.

ABSTRACT

Closed Traumatic Brain Injury (cTBI) produces one or multiple lesions that affect brain functions either transiently or permanently, constituting one of the leading causes of disability and mortality worldwide. Manual segmentation of these lesions from Magnetic Resonance Imaging (MRI) is a time-consuming, subjective, and error-prone process, which limits the ability to obtain an accurate diagnosis. Therefore, automating this process through deep learning techniques is necessary.

In this undergraduate thesis, a strategy for automatically segmenting lesions associated with cTBI in T1-weighted MRI is proposed. To this end, a preprocessing pipeline was established, including skull stripping, alignment of all images to a standard reference template, region-of-interest-based cropping, normalization, and resampling. Additionally, different normalization methods were evaluated to assess their impact on model performance. Subsequently, two models based on the U-Net architecture were trained and evaluated: nnU-Net and 3D U-Net. Moreover, an experiment using nnU-Net was conducted to analyze the model's behavior when applying transfer learning. The performance of both models was assessed using metrics such as the Dice Similarity Coefficient and Intersection over Union.

The results show that the model achieving the best performance was nnU-Net combined with the KDE normalization method, and that transfer learning significantly improved the evaluation metrics. Finally, a framework integrating the entire segmentation process was implemented, providing a tool to support clinical practice and future research related to cTBI.

CONTENIDO

1	INTRODUCCIÓN	1
1.1	Planteamiento del problema	1
1.2	Justificación del problema	4
1.3	Objetivos	5
1.3.1	Objetivo General	5
1.3.2	Objetivos Especificos	5
1.4	Alcance de la propuesta	5
1.5	Resultados obtenidos	5
2	MARCO REFERENCIAL	8
2.1	Marco teórico	8
2.1.1	Trauma Craneoencefálico Cerrado	8
2.1.2	Imagen por Resonancia Magnética	9
2.1.3	Atlas cerebral	10
2.1.4	Aprendizaje Automático	11
2.1.5	Aprendizaje Profundo	12
2.1.6	Aprendizaje por Transferencia	13
2.2	Estado del arte	14
2.2.1	Bases de datos	15
2.2.2	Estrategias de búsqueda	15
2.2.3	Proceso de selección	16
2.2.4	Proceso de recolección de datos	18
2.2.5	Resultados	19
3	ESTRATEGIA PROPUESTA PARA LA SEGMENTACIÓN DE LESIONES CEREBRALES POR TCEC	24
3.1	Conjunto de datos	25
3.1.1	Adquisición de datos y anotaciones de referencia	25
3.1.2	Características demográficas y clínicas	26
3.2	Preprocesamiento	29
3.2.1	Remoción del cráneo	29
3.2.2	Alineación de imágenes	30
3.2.3	Elección del atlas de referencia	31
3.2.4	Recorte basado en la región de interés	33
3.2.5	Normalización de intensidades	34
3.2.6	Remuestreo	37
3.3	Modelos de segmentación	37
3.3.1	No New U-Net (nnU-Net)	38

3.3.2	3D U-Net	40
3.4	Tecnologías usadas para la implementación del framework	41
4	EVALUACIÓN	42
4.1	Configuración de la evaluación	42
4.1.1	Coeficiente de similitud de Dice (DSC)	43
4.1.2	Índice de Jaccard (IoU)	43
4.1.3	Evaluación de los sujetos de control	43
4.2	Evaluación de la configuración del modelo y el atlas de registro	44
4.3	Evaluación de los métodos de normalización	46
4.4	Aprendizaje por Transferencia	47
4.4.1	Descripción del conjunto de datos para el aprendizaje por transferencia	48
4.4.2	Parámetros del preentrenamiento	48
4.4.3	Evaluación del aprendizaje por transferencia	49
4.5	Evaluación de la estrategia propuesta vs modelo de referencia 3D-UNET	49
4.5.1	Entrenamiento del modelo 3D U-Net	49
4.6	Análisis del desempeño del modelo final según variables demográficas	52
4.6.1	Análisis por rango de edad	52
4.6.2	Análisis por TSI	53
5	CONCLUSIONES Y TRABAJO FUTURO	56
5.1	Trabajos futuros	58
	REFERENCIAS	59

LISTA DE FIGURAS

Figura 1	Ejemplos de lesiones cerebrales en asociadas al TCEC . . .	2
Figura 2	Contusión hemorrágica aguda observada en imágenes cerebrales	9
Figura 3	Ejemplo de corte sagital, coronal y axial en IRM visualizada con FSLEyes, un visor incluido en la FMRIB Software Library (FSL).	10
Figura 4	Clasificación del Aprendizaje Automático	12
Figura 5	Representación gráfica de una red neuronal artificial tradicional y una red neuronal profunda	13
Figura 6	Comparación de los métodos de aprendizaje automático tradicional y por transferencia	14
Figura 7	Diagrama de flujo del proceso de selección	18
Figura 8	Recolección de datos con metodología PRISMA	23
Figura 9	Flujo de trabajo propuesto para la segmentación de lesiones cerebrales causadas por TCEC	24
Figura 10	Visualización de etiquetas de una IRM antes y después del mapeo	29
Figura 11	Pasos de ROBEX para segmentar todo el cerebro a partir de una imagen	30
Figura 12	Ejemplo de una IRM del conjunto AIMS-TBI antes y después de la remoción del cráneo realizada por ROBEX . . .	30
Figura 13	Ejemplo del registro de un sujeto con el atlas SRI24 utilizando ANTs	31
Figura 14	Propiedades de los atlas ICBM152 y SRI24	33
Figura 15	Ejemplo del recorte basado en el ROI de una IRM	34
Figura 16	Proceso de nnU-Net para la segmentación automática de imágenes médicas	38
Figura 17	Matrices de confusión de la evaluación de la configuración y el atlas	44
Figura 18	Matrices de confusión de la evaluación del método de normalización	46
Figura 19	Matriz de confusión de la evaluación del Aprendizaje por Transferencia	49
Figura 20	Matriz de confusión del modelo 3D U-Net	50
Figura 21	Segmentación de referencia vs Predicción de los modelos .	51
Figura 22	Número de sujetos por rango de edad y sexo	52

Figura 23	Promedio del DSC por rango de edad y sexo	53
Figura 24	Número de sujetos por TSI y sexo	54
Figura 25	Promedio del DSC por TSI y sexo	55

LISTA DE TABLAS

Tabla 1	Resultados obtenidos	6
Tabla 2	Características de atlas conocidos	10
Tabla 3	Características técnicas de las secuencias de IRM	26
Tabla 4	Características demográficas del conjunto de datos	27
Tabla 5	Distribución de los sujetos por sexo y rangos de edad	27
Tabla 6	División del conjunto de datos para entrenamiento y pruebas	28
Tabla 7	Etiquetas de referencia del conjunto de datos	28
Tabla 8	Características demográficas de los atlas cerebrales consi- derados para la alineación de imágenes	32
Tabla 9	Evaluación de configuración del modelo y el atlas de registro	46
Tabla 10	Evaluación de los métodos de estandarización	47
Tabla 11	Comparación de resultados de evaluación de los modelos	50

ACRÓNIMOS

TCE	Trauma Craneoencefálico
TCEC	Trauma Craneoencefálico Cerrado
IRM	Imagen por Resonancia Magnética
IA	Inteligencia Artificial
ML	Aprendizaje Automático
TL	Aprendizaje por Transferencia
CNN	Red Neuronal Convolutacional
DSC	Coeficiente de Similitud de Dice
IoU	Índice de Jaccard - Intersección sobre la Unión
TC	Tomografía Computarizada

INTRODUCCIÓN

1.1 PLANTEAMIENTO DEL PROBLEMA

El Traumatismo Craneoencefálico (TCE) es causado por una fuerza externa que impacta al paciente, generando una o múltiples lesiones que afectan las funciones cerebrales de forma transitoria o permanente. Las lesiones pueden ser clasificadas como abiertas o cerradas, dependiendo de si ocurre penetración en el cráneo. En este contexto, el presente trabajo de grado se enfocó en segmentar las lesiones causadas por el Trauma Craneoencefálico Cerrado (TCEC), que ocurre cuando el cráneo no es penetrado.

El TCEC causa lesiones como la conmoción cerebral, contusión, daño axonal difuso, hematomas o hemorragias intracraneales [1]. Según [2], estas pueden originar diversas secuelas neurológicas como la cefalea, alteraciones del lenguaje, déficits motores y sensitivos, deterioro cognitivo, demencia y epilepsia, además de trastornos psiquiátricos como la depresión, ansiedad y trastorno por estrés postraumático.

El TCE es un problema mundial de gran magnitud en el campo médico, siendo las causas principales las caídas y los accidentes de tránsito. Un estudio realizado en 2019 reportó 27.16 millones de casos nuevos de TCE, 48.99 millones de casos prevalentes, que incluyen tanto los casos nuevos como los existentes y 7.08 millones de años vividos con discapacidad resultado de las secuelas físicas, cognitivas y emocionales asociadas al TCE [3]. Además de ser un gran problema para la población afectada, también es un desafío importante para el sistema de salud, el cual debe atender a un alto número de personas que requieren hospitalizaciones, atención en urgencias o que, en casos graves pueden fallecer. Un estudio realizado en Estados Unidos en el año 2013 ejemplifica esta problemática, al registrar 2.5 millones de visitas a urgencias, alrededor de 282,000 hospitalizaciones y aproximadamente 56,000 muertes relacionadas con TCE [4].

En el contexto nacional, según una revisión de la literatura [5], una investigación realizada en 2014 registró cerca de 1,000 muertes anuales por TCE, de las cuales más del 50% fueron causadas por accidentes de tránsito, especialmente los ocasionados en motocicletas. Recientemente, en el año 2023 el Ministerio

de Salud de Colombia informó que aunque el 70% de los pacientes con esta condición logran recuperarse, el 9% fallece antes de recibir atención en las clínicas, el 6% fallece durante la hospitalización y el 15% presenta secuelas. Es necesario resaltar que en ese mismo año se registraron 200 casos de TCE por cada 100,000 habitantes en el país [6].

Dado que el TCE afecta a una gran parte de la población y puede tener consecuencias graves, es esencial que se realicen diagnósticos tempranos y precisos para tratar a los pacientes de manera rápida y efectiva. Para ello, se utilizan generalmente las Imágenes por Resonancia Magnética (IRM), un procedimiento que permite la visualización de las lesiones y estructuras internas del cerebro. Las IRM son analizadas por profesionales médicos, quienes segmentan las lesiones de forma manual, tarea que puede tomar varias horas dependiendo del número de imágenes y la complejidad de las lesiones.

Dichas lesiones varían en tamaño y número, pueden ser focales o difusas, y se extienden a través de múltiples tipos de tejido como la sustancia gris, la sustancia blanca y el líquido cefalorraquídeo. Además, pueden ocurrir en regiones homólogas de ambos hemisferios, como se puede ver en la Figura 1. Estas características dificultan el registro y la normalización de imágenes, e introducen errores tanto locales como globales en la segmentación [7].

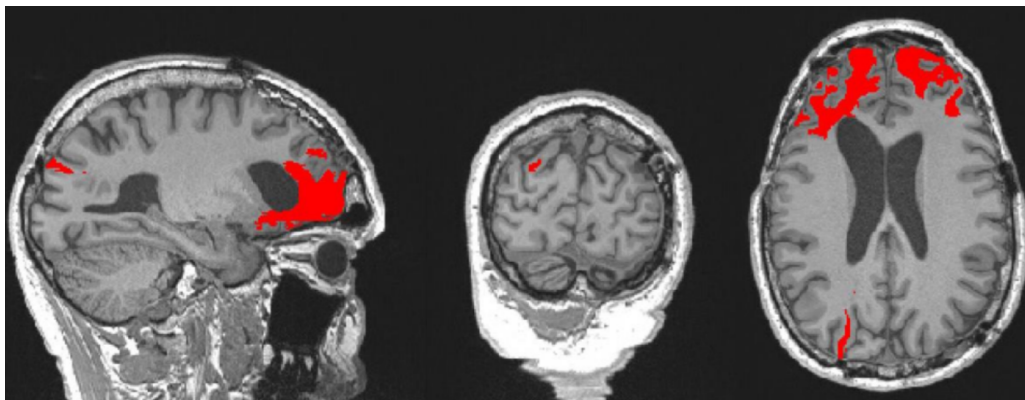


Figura 1: Ejemplos de lesiones cerebrales asociadas al TCEC. La imagen presenta lesiones que varían en tamaño y distribución, incluyendo afectación bilateral y difusa en la corteza frontal, así como lesiones focales profundas [7].

Asimismo, el resultado de la segmentación manual tiende a ser subjetivo y es propenso a errores humanos debido a que depende en gran medida de la experiencia del profesional, lo que dificulta obtener resultados consistentes y

reproducibles.

Para superar las limitaciones de la segmentación manual se han desarrollado distintos modelos de segmentación automática y semiautomática, cuyo objetivo es automatizar el proceso y reducir el tiempo al evaluar grandes conjuntos de datos. Esto se evidencia en [8], que compara métodos manuales, semiautomáticos y automáticos, mostrando que las alternativas automáticas pueden procesar los datos con mayor rapidez, aunque cada enfoque presenta ventajas y desventajas en cuanto a su precisión con diferentes tipos de lesiones.

Algunos métodos de segmentación semiautomática utilizados frecuentemente son técnicas de agrupamiento como K-means y Fuzzy C-means, que pueden procesar grandes volúmenes de datos rápidamente, pero requieren de una configuración inicial de parámetros, por lo que su funcionamiento aún depende de intervención humana. En cuanto a los modelos automáticos, se han desarrollado modelos basados en aprendizaje profundo, como las redes neuronales convolucionales (CNN) y los transformadores. Las herramientas que utilizan estos enfoques, como IcoBrain y FreeSurfer, pueden procesar grandes cantidades de datos y realizar la segmentación sin necesidad de intervención humana directa. No obstante, el entrenamiento de redes neuronales profundas tiene un alto costo en términos de tiempo y recursos computacionales, y su rendimiento depende de la cantidad y calidad del conjunto de datos [9] [10] [11].

A pesar de los avances del aprendizaje profundo, aún no existe un modelo de segmentación automática que se haya establecido como un estándar preciso para segmentar lesiones en el cerebro, específicamente aquellas relacionadas con el TCEC [12]. Como consecuencia, la segmentación imprecisa afecta negativamente a la detección temprana de las lesiones y ocasiona un diagnóstico tardío limitando las opciones de tratamiento oportuno para la población afectada por TCEC.

Por tanto, este trabajo de grado aborda la dificultad de obtener una segmentación precisa de lesiones presentadas en pacientes afectados por Trauma Craneoencefálico Cerrado mediante Imágenes por Resonancia Magnética.

De acuerdo con esto, se planteó la siguiente pregunta de investigación: **¿Cómo mejorar la segmentación de lesiones cerebrales causadas por TCEC en IRM?**

1.2 JUSTIFICACIÓN DEL PROBLEMA

El TCEC es una lesión a la que gran parte de la población puede estar expuesta, principalmente por caídas o accidentes de tránsito. El tiempo en que se tarda el personal médico en realizar un diagnóstico correcto es crucial para aumentar las posibilidades de un tratamiento adecuado y oportuno. A continuación, se indica por qué este problema es relevante en diferentes ámbitos.

Las personas con TCEC enfrentan distintos desafíos a nivel social, emocional y profesional, lo que afecta directamente su calidad de vida. Algunos síntomas asociados al trauma pueden presentarse de forma crónica y desencadenar enfermedades como la demencia, el deterioro cognitivo, la epilepsia, la esquizofrenia, entre otros. Un estudio realizado por [13], determinó que el deterioro cognitivo se relaciona con una disminución en la calidad de vida en los dominios de funcionamiento físico, social y emocional, así como en la salud mental. Estos problemas de salud tienen un impacto directo en el desempeño profesional, las interacciones sociales y las actividades diarias. Además, se ha documentado que la fatiga mental después de un TCE tiene un impacto negativo en el estado laboral y social de las personas [14].

Además de las consecuencias sociales, el TCEC también representa un reto para el campo de la salud y la economía. La atención médica, las hospitalizaciones y los tratamientos de los pacientes afectados implican un alto costo. En una revisión sistemática realizada en 2019, se analizaron los costos hospitalarios por paciente, que resultaron ser generalmente altos, variando entre 2,130 y 401,808 dólares [15]. En Colombia, enfermedades relacionadas como la demencia tienen un costo de tratamiento por paciente que puede llegar a los 99 millones de pesos, incluyendo los costos del cuidador. Asimismo, en 2013 se estimó que el costo directo mínimo anual estaba entre 116 a 140 mil millones de pesos [16].

Debido a esta problemática, en este proyecto se analizan los métodos de segmentación más utilizados en pacientes con TCEC presentes en la literatura. A partir de este análisis, se propone el método más adecuado en cuanto a precisión, sensibilidad y especificidad, y se implementa un framework que realiza la segmentación de las lesiones cerebrales empleando el enfoque seleccionado. El beneficio a nivel académico de este framework es que puede ser fácilmente reproducible y adaptable en otros proyectos, facilitando su integración en futuras investigaciones. En conjunto, estos aspectos resaltan la importancia e influencia del trabajo de grado tanto en el campo de la salud como en la formación profesional e investigativa.

1.3 OBJETIVOS

1.3.1 *Objetivo General*

Proponer un método de segmentación que permita delimitar lesiones del cerebro en IRM de personas afectadas por TCEC.

1.3.2 *Objetivos Especificos*

1. Revisar la literatura utilizando la metodología PRISMA para identificar estudios relevantes sobre segmentación de lesiones cerebrales en IRM de pacientes con TCEC.
2. Seleccionar conjuntos de datos públicos apropiados para el entrenamiento y evaluación de los algoritmos de segmentación de lesiones, asegurando su representatividad y calidad.
3. Seleccionar métodos de segmentación revisados en la literatura para la identificación de lesiones cerebrales en IRM.
4. Evaluar el rendimiento de los modelos de segmentación seleccionados en términos de precisión, sensibilidad y especificidad utilizando diversas métricas de validación.
5. Implementar un framework que integre los modelos seleccionados y permita segmentar IRM de personas con lesiones por TCEC.

1.4 ALCANCE DE LA PROPUESTA

La investigación contó con el criterio de un experto para su desarrollo. Sin embargo, no se incluyó una validación clínica que cuente con la participación de pacientes.

1.5 RESULTADOS OBTENIDOS

En la Tabla 1 se presentan los resultados obtenidos para cada objetivo específico planteado en la sección 1.3.

Tabla 1: Resultados obtenidos

OBJETIVO ESPECÍFICO	RESULTADO OBTENIDO
Revisar la literatura utilizando la metodología PRISMA para identificar estudios relevantes sobre segmentación de lesiones cerebrales en IRM de pacientes con TCEC	Se realizó una revisión sistemática de la literatura siguiendo la metodología PRISMA para identificar estudios relevantes sobre la segmentación de lesiones en IRM de pacientes con TCEC, sus enfoques más utilizados y limitaciones. El proceso de revisión y la síntesis de los resultados se encuentran en la sección 2.2
Seleccionar conjuntos de datos públicos apropiados para el entrenamiento y evaluación de los algoritmos de segmentación de lesiones, asegurando su representatividad y calidad	Se seleccionó el conjunto de datos perteneciente al desafío de Automated Identification of Mod-Sev TBI Lesions 2025 (identificación automatizada de lesiones de TCE grave a moderado, AIMS-TBI) ¹ . La procedencia, descripción y características del conjunto de datos se encuentran detallados en la sección 3.1
Seleccionar métodos de segmentación revisados en la literatura para la identificación de lesiones cerebrales en IRM	Se seleccionaron dos modelos de segmentación basados en la arquitectura U-Net. Su descripción, funcionamiento y demás detalles se presentan en la sección 3.3
Evaluar el rendimiento de los modelos de segmentación seleccionados en términos de precisión, sensibilidad y especificidad utilizando diversas métricas de validación	Se definieron las métricas y se evaluaron los modelos de segmentación seleccionados. La descripción de cada métrica y los resultados obtenidos se presentan en el capítulo 4

(Continúa en la siguiente página)

¹ <https://aims-tbi25.grand-challenge.org/>

(Continuación de la tabla anterior)

OBJETIVO ESPECÍFICO	RESULTADO OBTENIDO
Implementar un framework que integre los modelos seleccionados y permita segmentar IRM de personas con lesiones por TCEC	El framework que integra la propuesta para segmentar IRM de pacientes con TCEC, se encuentra disponible públicamente en: https://github.com/CastanoJenifer/TCEC-Segmentation .

MARCO REFERENCIAL

En este capítulo se presentan los conceptos fundamentales asociados con Inteligencia Artificial aplicada a IRM y su relación con el TCEC.

2.1 MARCO TEÓRICO

2.1.1 *Trauma Craneoencefálico Cerrado*

El TCEC es un daño cerebral causado por una fuerza externa que incide en el cráneo sin penetrarlo. Puede originarse por un golpe directo en la cabeza o por la aceleración-desaceleración rápida del cerebro dentro del cráneo. La magnitud del daño cerebral se puede clasificar en tres categorías: leve, moderada y grave, determinadas por la naturaleza, intensidad, orientación y duración del impacto. Los tipos de TCEC incluyen la conmoción cerebral, la contusión, la lesión axonal difusa y los hematomas intracraneales (hematoma epidural, hematoma subdural, hemorragia subaracnoidea y hemorragia intraparenquimatosa) [17].

Frecuentemente, los síntomas del TCEC no son específicos e incluyen náuseas, vómitos, cefalea, pérdida del conocimiento, sensación de confusión, mareo, tinnitus y alteraciones visuales. A largo plazo, muchos pacientes presentan síntomas persistentes como cambios emocionales, problemas de equilibrio, dificultades cognitivas y de memoria, ansiedad, depresión, trastornos del sueño y de la concentración, alucinaciones o delirios. El TCEC puede tener consecuencias a corto y largo plazo afectando la salud física, mental y funcional del individuo [17] [18].

En la Figura 2 se puede observar un paciente con un TCEC grave que muestra contusiones hemorrágicas agudas presentadas en imágenes cerebrales obtenidas mediante Tomografía Computarizada (TC) y Resonancia Magnética (RM).

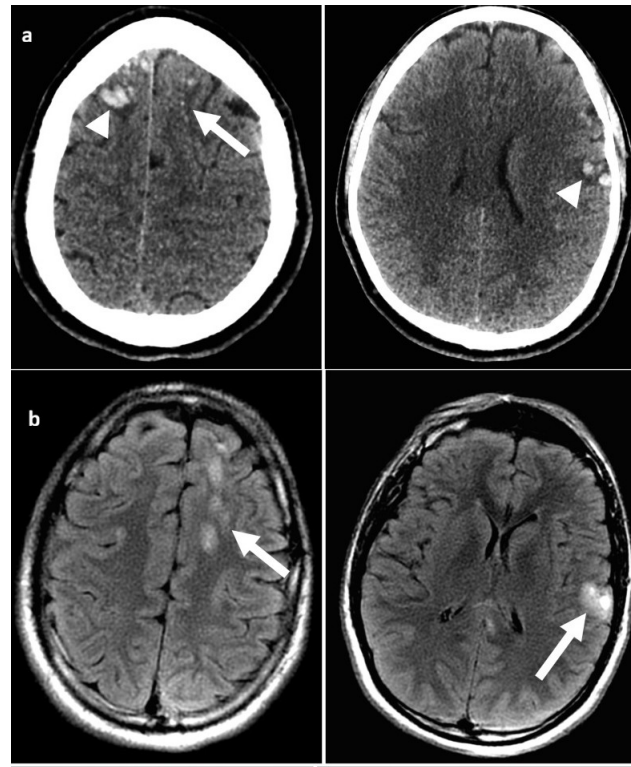


Figura 2: Contusión hemorrágica aguda observada en TC y RM. (a) TC sin contraste que muestra contusiones hemorrágicas en los lóbulos frontales bilaterales (puntas de flecha), y microhemorragias puntiformes en el giro frontal superior izquierdo, compatibles con lesiones axónicas hemorrágicas traumáticas por cizallamiento (flecha). (b) FLAIR T2 que muestra edema de sustancia blanca multifocal (imagen izquierda, flecha) asociado a microhemorragias y edema cortical/subcortical focal (imagen derecha, flecha) [19].

2.1.2 Imagen por Resonancia Magnética

La imagen obtenida por Resonancia Magnética Nuclear (RMN) es una técnica de diagnóstico avanzado que permite la obtención de imágenes detalladas de órganos y tejidos internos mediante el uso de imanes de campos magnéticos, ondas de radiofrecuencia y computadoras para la producción de imágenes. La RMN ha evolucionado hasta convertirse en IRM que es una técnica esencial para diagnosticar enfermedades gracias a su alto nivel de detalle y contraste entre los diferentes tejidos [20].

Entre las modalidades de IRM se tienen las secuencias ponderadas por densidad de protones, ponderadas en T1, en T2 y por difusión. En particular, la

secuencia T1-w muestra imágenes que se aproximan a la apariencia macroscópica de los tejidos, donde la sustancia gris se ve gris y la blanca más brillante. Por otro lado, la secuencia T2-w resalta el contenido de agua, mostrando el líquido cefalorraquídeo brillante. Finalmente, la secuencia FLAIR suprime el líquido cefalorraquídeo y presenta la sustancia blanca más oscura que la gris [21].

En la Figura 3 se puede observar un ejemplo de los cortes sagital, coronal y axial en una IRM.

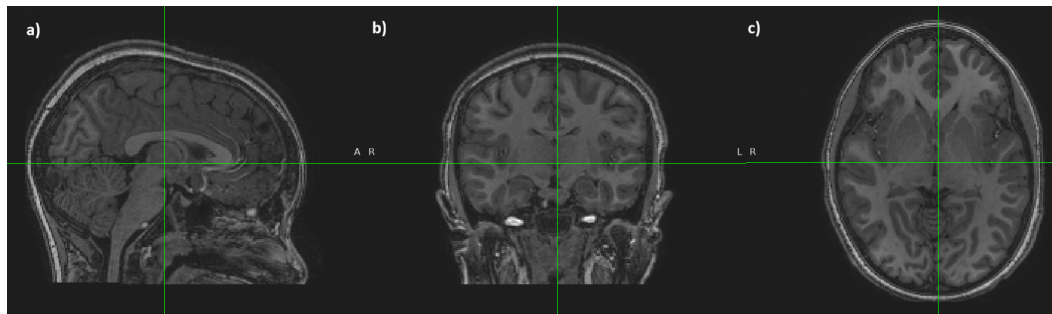


Figura 3: Ejemplo de corte sagital (a), coronal (b) y axial (c) de una IRM visualizada con FSLEyes, un visor incluido en la FMRIB Software Library (FSL) [Elaboración propia].

2.1.3 Atlas cerebral

El atlas cerebral es una representación del cerebro construida para facilitar la descripción de aspectos estructurales y funcionales, y proporcionar un sistema de coordenadas de referencia espacial estándar para la localización anatómica y funcional del cerebro humano. Aunque algunos atlas se originaron a partir de un solo sujeto como es el caso del atlas Talairach, muchos atlas modernos utilizan datos promediados de una población para ser más representativos [22]. En la Tabla 2 se presentan las características de algunos de los atlas más conocidos, tomadas de un estudio de 2017 [23].

Tabla 2: Características de atlas conocidos

ATLAS	EDAD	SEXO	SECUENCIAS	MÉTODO
Talairach & Tournoux	60	1 F	Coordenadas estereotáficas	Corte post mortem

(Continúa en la siguiente página)

(Continuación de la tabla anterior)

ATLAS	EDAD	SEXO	SECUENCIAS	MÉTODO
Chinese probabilistic atlas	18 – 70	1,000 N/A	T1, T2	Promedio por vóxel
MNI/ICBM ₁₅₂	18 – 44	86 M 66 F	T1, T2/PD	Promedio por vóxel
MNI-Colin ₂₇	N/A	1 M	T1, T2/PD	Promedio por vóxel
SRI ₂₄	19 – 84	12 M 12 F	T1, T2/PD	Promedio por vóxel

2.1.4 Aprendizaje Automático

El Aprendizaje Automático o Machine Learning (ML) es una rama de la IA que permite que los sistemas aprendan a partir de los datos sin ser programados explícitamente, utilizando algoritmos adaptados a las características del problema, los datos disponibles y el modelo más adecuado para su solución. El ML se clasifica en algoritmos supervisados, entre los cuales se encuentran los árboles de decisión y máquinas de vectores de soporte, algoritmos semisupervisados y algoritmos no supervisados como K-Means [24]. En la Figura 4 se muestra gráficamente la clasificación del ML.

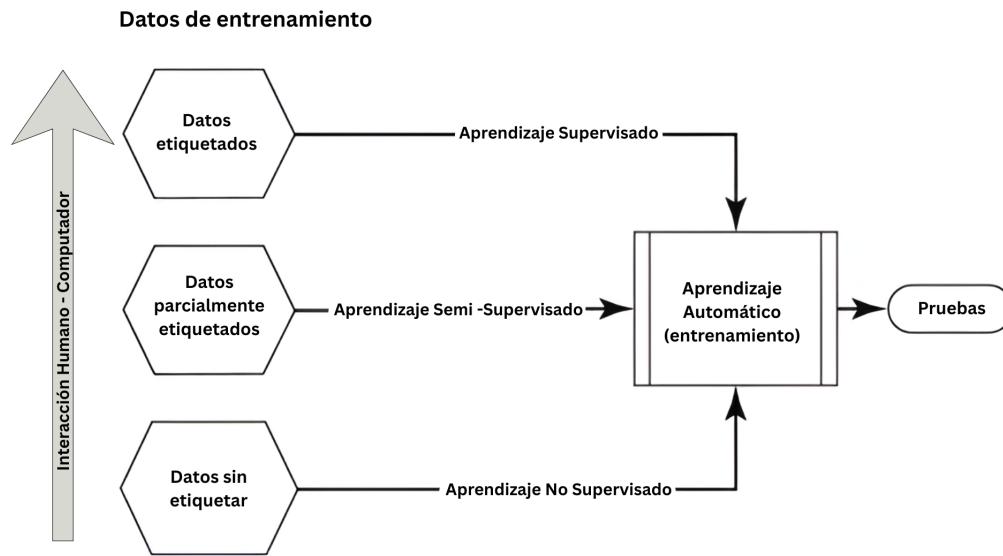


Figura 4: Clasificación del Aprendizaje Automático [25].

2.1.5 Aprendizaje Profundo

El aprendizaje profundo es un subconjunto del ML, que utiliza redes neuronales con múltiples capas ocultas. Sus elementos son nodos o neuronas, donde las señales de entrada se combinan mediante pesos y sesgos. Mientras en una red de una sola capa las entradas se multiplican por los pesos, se suman y procesan mediante una función de activación, en aprendizaje profundo las salidas de cada capa son entradas de la siguiente. Al contar con más pesos y conexiones adicionales, la red aprende y resuelve problemas más complejos. En la Figura 5 se muestra la diferencia entre el flujo de procesamiento de una red neuronal artificial tradicional y una red neuronal profunda.

Existen muchos tipos de redes de aprendizaje profundo, entre los cuales están el Perceptrón Multicapa (MLP), la Red Neuronal Convolucional (CNN), la Red Neuronal Recurrente (RNN) y la Red Generativa Antagónica (GAN) [26]. En particular, las CNN son arquitecturas diseñadas específicamente para procesar datos con múltiples arreglos, como imágenes y señales de audio, aprovechando propiedades como conexiones locales y pesos compartidos para detectar características jerárquicas desde bordes simples hasta formas complejas. Su excelente rendimiento en tareas de visión por computadora las ha convertido en una herramienta fundamental para la segmentación y clasificación de imágenes médicas [27].

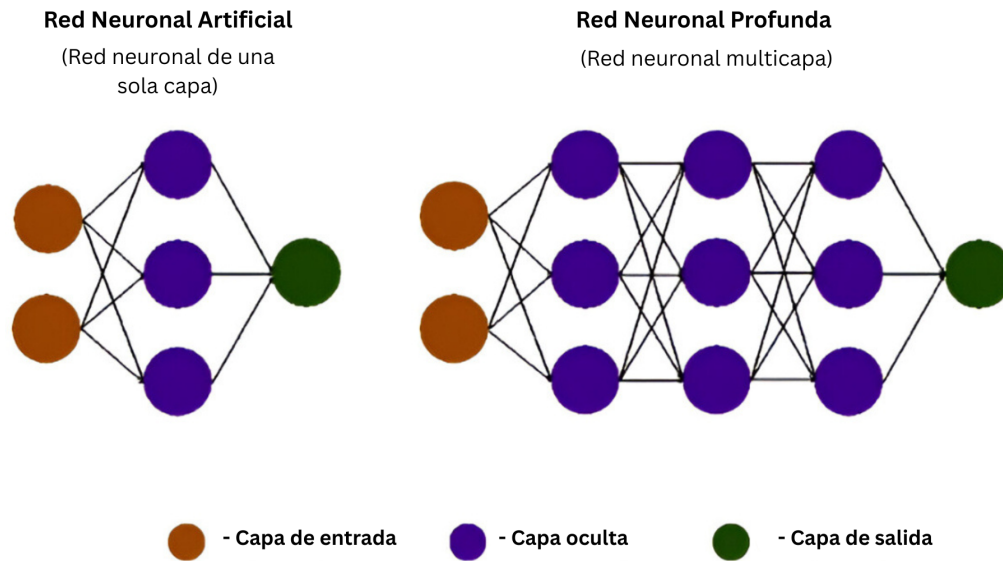


Figura 5: Representación gráfica de una red neuronal artificial tradicional y una red neuronal profunda [28].

2.1.6 Aprendizaje por Transferencia

El Aprendizaje por Transferencia (TL) es una técnica de aprendizaje automático que permite mejorar el rendimiento de un modelo en una tarea objetivo utilizando conocimientos adquiridos previamente en una tarea relacionada. Esta técnica es útil cuando se dispone de una cantidad limitada de datos etiquetados para el nuevo problema. Una de las primeras formas de TL fue la basada en características, que consistía en utilizar las características aprendidas en una tarea de origen para mejorar el rendimiento en la tarea objetivo. Otra forma de TL es el ajuste fino, que consiste en entrenar un modelo preentrenado con un conjunto de datos más pequeño [29].

Como se muestra en la Figura 6, el aprendizaje automático tradicional se entrena desde cero utilizando un conjunto de datos específico para la tarea objetivo, mientras que el TL tiene como objetivo transferir la información de las tareas base para aplicarlo a una nueva tarea con un conjunto de datos de entrenamiento limitado o de alta calidad [30].

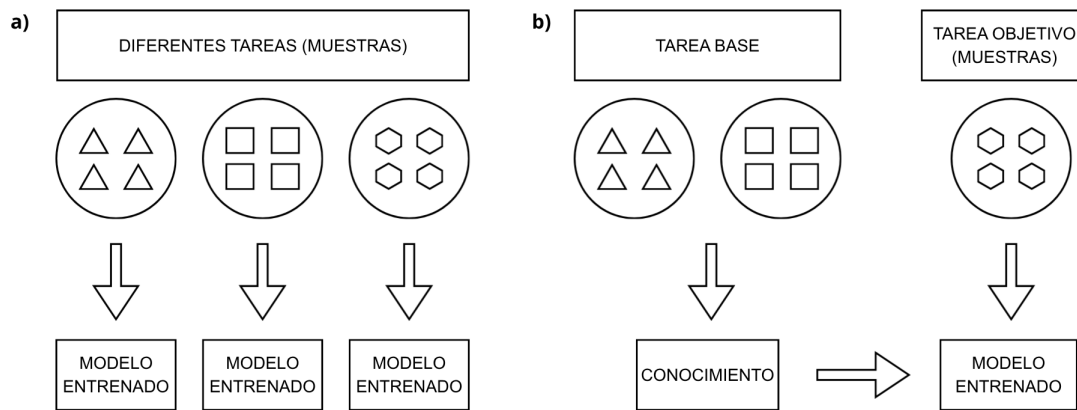


Figura 6: Comparación de los métodos de aprendizaje automático tradicional y por transferencia. (a) Método de aprendizaje tradicional. (b) Método de aprendizaje por transferencia [30].

2.2 ESTADO DEL ARTE

Se realizó una revisión sistemática utilizando la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [31], con el objetivo de analizar la literatura científica relacionada con la segmentación de IRM de pacientes con TCEC. Las siguientes preguntas se consideraron para orientar el proceso:

- ¿Qué métodos o técnicas se utilizan para la segmentación de lesiones cerebrales en IRM?
- ¿Cuáles son los métodos de preprocesamiento empleados en la preparación de IRM?
- ¿Cuál debe ser el tamaño mínimo de la muestra de un conjunto de datos?
- ¿Cuáles son los lenguajes de programación, frameworks, librerías o bibliotecas utilizadas para la segmentación de IRM?
- ¿Cuáles son las estructuras o regiones cerebrales más segmentadas en pacientes con TCEC?
- ¿Cuál es la proporción de los conjuntos de entrenamiento y de prueba?
- ¿Qué bases de datos de IRM de pacientes con cambios estructurales cerebrales o lesiones están disponibles públicamente para la investigación?

- ¿Existe una implementación disponible del modelo de segmentación utilizado?
- ¿Cuáles son las métricas o criterios de evaluación utilizados para medir el rendimiento de los métodos de segmentación en IRM?

Criterios de elegibilidad

Los criterios establecidos para la inclusión o exclusión de estudios en la revisión sistemática se presentan a continuación:

Inclusión

- **Estructura de interés:** Estudios que investiguen el área del cerebro.
- **Modalidad de la imagen:** Estudios que utilicen IRM incluyendo modalidades como T1-w, T2-w, FLAIR, IRM Estructural, entre otras relevantes.
- **Problemática:** Estudios que aborden temas como el daño cerebral, el TCEC, el deterioro cognitivo, la demencia y el Alzheimer.
- **Tipo de evaluación:** Estudios que utilicen técnicas de preprocesamiento, inteligencia artificial, aprendizaje profundo, algoritmos de aprendizaje, redes neuronales, modelos o técnicas de segmentación.

Exclusión

Se excluyeron del análisis estudios realizados en modelos animales, así como investigaciones preclínicas y aquellos publicados antes del año 2022.

2.2.1 Bases de datos

Las búsquedas se realizaron en tres (3) bases de datos científicas: PubMed, Web of science y Scopus, seleccionadas por su relevancia en áreas relacionadas con el tema de estudio. En total, se eligieron 33 artículos relevantes, distribuidos de la siguiente forma: 4 artículos de PubMed, 6 de Web of Science, 17 de Scopus y 6 adicionales a través de la revisión de citas relacionadas. Estas consultas fueron realizadas hasta marzo de 2025.

2.2.2 Estrategias de búsqueda

Las estrategias se establecieron según los criterios de inclusión, de esta manera, los términos de búsqueda están relacionados con la estructura de interés, la

problemática presentada, la modalidad de imagen y el tipo de evaluación. Estos términos se buscaron específicamente en el título y resumen de cada investigación. Asimismo, se establecieron filtros adicionales limitando la especie a humanos y el rango de publicación a los años 2022 – 2024.

Las ecuaciones de búsqueda específicas fueron:

- **Estructura de interés:**[AB=(“brain”)]
- **Modalidad de la imagen:** [AB=(“T1” OR “T1-w” OR “T1-weighted” OR “T2” OR “T2-w” OR “T2-weighted” OR “FLAIR” OR “structural magnetic resonance imaging” OR “structural MRI” OR “sMRI” OR “MRI” OR “magnetic resonance image”)]
- **Problemática:** [AB=(“brain lesions” OR “brain damage” OR “closed TBI” OR “closed brain injury” OR “closed TBI lesion” OR “traumatic brain injury” OR “closed head injury” OR “nonpenetrating head injury” OR “blunt head injury” OR “closed head trauma”)]
- **Algoritmos y técnicas:** [AB= (“artificial intelligence” OR “machine learning” OR “deep learning” OR “convolutional neural networks” OR “segmentation methods” OR “lesion segmentation” OR “image segmentation” OR “image segmentation methods” OR “automatic segmentation” OR “semi-automatic segmentation” OR “pre-processing” OR “segmentation algorithms” OR “segmentation models” OR “segmentation”)]

2.2.3 Proceso de selección

A partir de las consultas realizadas según la estrategia de búsqueda mencionada anteriormente, se identificaron inicialmente 522 artículos. El proceso de selección se desarrolló en varias etapas, en primer lugar, se eliminaron los artículos duplicados encontrados entre las bases de datos consultadas.

A continuación, se analizaron los títulos de los artículos restantes para identificar aquellos relacionados con demencia, deterioro cognitivo, identificación de lesiones por TCEC y otros temas afines. Se excluyeron aquellos estudios cuyo enfoque no estaba relacionado con el TCE, como los que trataban sobre esclerosis múltiple, accidentes cerebrovasculares u otras patologías no relacionadas. Sin embargo, algunos artículos enfocados en tumores fueron considerados relevantes debido a que estudiaban lesiones cerebrales.

Posteriormente, los artículos seleccionados tras el filtrado por título fueron evaluados a través de los resúmenes, lo que permitió profundizar en el contenido y determinar su relevancia para la revisión sistemática.

Adicionalmente, se realizó una búsqueda preliminar rápida para asegurar que los artículos cumplieran con los criterios de inclusión previamente establecidos. Durante esta etapa, se excluyeron artículos por las siguientes razones:

- No se aborda un proceso de segmentación.
- La segmentación realizada fue completamente manual.
- El estudio no se relaciona con TCE, ni sus lesiones son relevantes para la investigación.
- El objetivo del estudio se centra en áreas no relevantes para la investigación, como las variaciones anatómicas por sexo, sin enfocarse en lesiones cerebrales.

Finalmente, después de realizar todas las etapas, se obtuvo un conjunto definitivo de 33 artículos que cumplen con los criterios de inclusión establecidos. Estos estudios fueron utilizados en la síntesis y análisis final de la revisión sistemática. En la Figura 7 se ilustra el diagrama de flujo del proceso de selección de los 33 artículos.

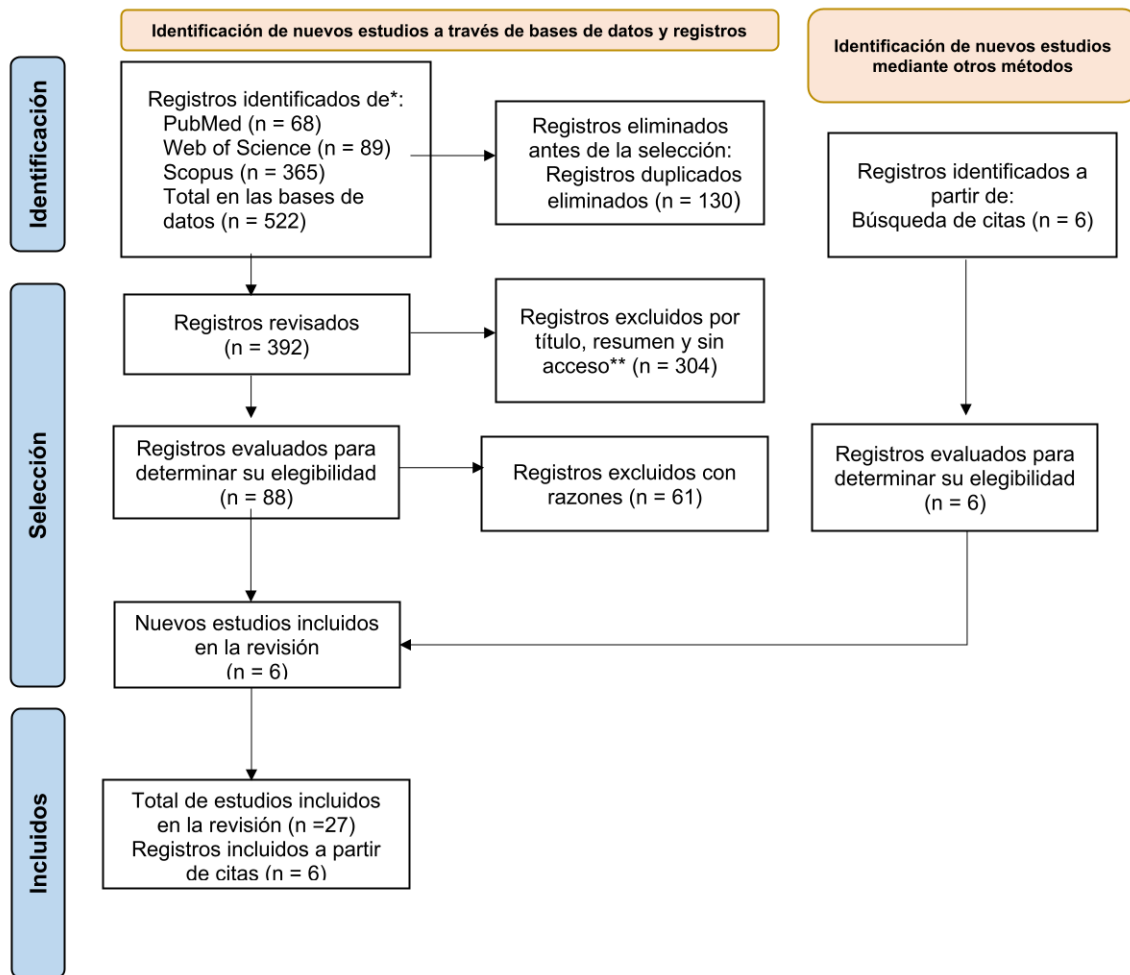


Figura 7: Diagrama de flujo del proceso de selección [Elaboración propia].

2.2.4 Proceso de recolección de datos

Los datos de los informes se recopilaron a través de una revisión exhaustiva de cada estudio incluido. Cada uno de los 33 estudios fue examinado de manera independiente, sin el uso de herramientas de automatización. El objetivo fue obtener información sobre diversos aspectos relevantes como el tamaño de los conjuntos de datos utilizados, técnicas de preprocesamiento, segmentación, regiones cerebrales de interés y las tecnologías aplicadas. Además, se realizó una revisión detallada de los criterios de inclusión para garantizar que los estudios seleccionados cumplieran con los requisitos establecidos.

2.2.5 Resultados

¿Qué métodos o técnicas se utilizan para la segmentación de lesiones cerebrales en IRM?

Las técnicas utilizadas para la segmentación de lesiones cerebrales relevantes en IRM pueden variar según el estudio. Se observaron las siguientes:

1. **Redes Neuronales:** Se destacan modelos basados en U-Net y en redes neuronales convolucionales:
 - HUT: Combinación de U-Net con un transformador de resolución cruzada. Este modelo es utilizado para mejorar el rendimiento de la segmentación de lesiones en una sola modalidad de imagen [32].
 - U-Net: La arquitectura U-Net combina mapas de características de alto y bajo nivel, y puede recuperar estructuras más finas en la imagen. Esta arquitectura se utiliza tanto para segmentación 2D como 3D de lesiones en la materia blanca y estructuras relacionadas con la consciencia. Además, se menciona una variación denominada VB-Net con canales de salida como el fondo, la sustancia blanca y otras lesiones intracraneales. También se emplea una combinación para segmentar los ganglios basales, que incluye la arquitectura U-Net junto con un modelo probabilístico CRF, el modelo basado en transformadores UNTER y MSGS-Net [33] [34] [35] [36] [37].
 - SwinUNETR 2D: Es una red neuronal profunda con un enfoque binario de segmentación de hiperintensidades [38].
 - Red neuronal convolucional: Se utilizó para la segmentación de los espacios perivasculares, el hipocampo y para medir el área o volumen de los hematomas subdurales [10] [39] [40].
2. **Técnicas de agrupamiento:** Se mencionan diversas técnicas de agrupamiento como K-Means y Fuzzy C-Means, utilizadas para segmentar el tejido cerebral.[9] [41] [42]. K-Means asigna a cada vóxel a un solo clúster, mientras que Fuzzy C-Means asigna grados de pertenencia a todos los clústeres, generando como resultado un mapa de pertenencias difusas.
3. **Técnicas de optimización:** Se emplean distintas técnicas de optimización como el enjambre de partículas (PSO) que utiliza el análisis de varianza fijo de dos vías (ANOVA) como función de aptitud. Otro ejemplo es la optimización CUF-GW que combina características de dos algoritmos de optimización: el comportamiento de búsqueda del lobo gris (Canis lupus) y las características de búsqueda del calamar (cuttlefish).

4. **Modelo probabilístico:** Se utilizó un enfoque bayesiano de campo aleatorio de Markov utilizado para segmentar el cerebro y un procedimiento automatizado de cuantificación basado en atlas que divide el cerebro en regiones definidas utilizando un enfoque Potholes and Molehills [43].
5. **Herramientas:** Entre las principales herramientas de segmentación utilizadas se encuentran Icobrain, FreeSurfer, SPM y FSL.
 - Icobrain: Software de aprendizaje automático desarrollado por Icometrix para la segmentación de lesiones axonales difusas [10].
 - FreeSurfer: Usado para la segmentación de estructuras subcorticales y la generación de segmentaciones volumétricas mediante el programa recon-all. Este permitió la detección de daños a lo largo de los tractos de sustancia blanca, además de proporcionar representaciones del grosor cortical y de los volúmenes subcorticales [11] [44] [45].
 - Statistical Parametric Mapping (SPM2) y FSL (FAST): Herramientas de código abierto usadas para generar máscaras de segmentación del líquido cefalorraquídeo, materia blanca y materia gris [46] [47].

¿Cuáles son los métodos de preprocesamiento empleados en la preparación de IRM?

En el procesamiento de IRM se aplican distintos métodos para mejorar la calidad de los datos y minimizar artefactos que podrían interferir con el análisis, lo que se denomina preprocesamiento. Entre estos se incluyen la eliminación del cráneo y los tejidos no cerebrales mediante herramientas como FSL (BET) y algoritmos basados en umbralización de histograma, así como la corrección de distorsiones y del campo de sesgo para corregir inhomogeneidades. También, se realizan procesos de normalización de intensidades, eliminación de artefactos de movimiento y suavizado utilizando kernels gaussianos. Asimismo, se usaron técnicas de aumento de datos aplicando transformaciones como escalado, rotación, inversión de ejes y desplazamiento de la intensidad, además del correregistro a un atlas estándar y el registro multimodal para alinear diferentes imágenes. Finalmente, se mencionan estrategias de eliminación de ruido y mejora de brillo y contraste, como el filtro de ecualización de histograma adaptativo de contraste (AHE), la transformada wavelet de Bayes (WT) y herramientas específicas como dwidenoise de MRtrix [48] [49] [38] [50]. La aplicación de estas técnicas permite obtener resultados precisos y confiables.

¿Cuál debe ser el tamaño mínimo de la muestra del conjunto de datos?

El tamaño del conjunto de datos utilizado fue determinado por factores como la disponibilidad de datos y los criterios de inclusión establecidos en cada estudio,

tanto en relación con el estado del paciente como con la calidad de las imágenes. Es común que algunos estudios descartaran imágenes que no cumplieran con la calidad necesaria necesitada. De esta manera, los tamaños de los conjuntos de datos variaron ampliamente, desde estudios con muestras de 18 pacientes hasta aquellos con 1045 pacientes, adaptándose a las necesidades específicas de cada investigación [35] [11] [44].

¿Cuáles son los lenguajes de programación, frameworks, librerías o bibliotecas utilizadas para la segmentación de IRM?

Las herramientas utilizadas abarcan diversos recursos, se puede destacar el uso del entorno Python junto con bibliotecas como Numpy, Pandas, Matplotlib, Seaborn, SciPy y ML Scikit-learn. Además, se usaron frameworks como PyTorch DL y PyTorch Geometric, aplicados para la segmentación y modelado geométrico. También se incluyeron herramientas de análisis y visualización como FreeSurfer y OpenCV. Para el análisis estadístico, se utilizó R en su versión 4.1.1, mientras que Matlab se implementó en tareas de preprocesamiento y en algunas etapas del análisis estadístico.

¿Cuáles son las estructuras o regiones cerebrales más segmentadas en pacientes con TCEC?

- **Materia blanca:** En particular, se identifican lesiones axonales difusas en tractos como el cuerpo calloso, los tractos corticoespinales, la corona radiata, los fascículos longitudinales inferiores y el pedúnculo cerebeloso medio.
- **Hipocampo:** Región fundamental para la memoria y frecuentemente afectada en diversas patologías neurológicas como el alzheimer. Está relacionado con el deterioro emocional y cognitivo.
- **Materia gris:** Se destacan áreas como el giro frontal inferior, el giro cingulado anterior y posterior, el parahipocampo, la ínsula, la amígdala, el caudado, el putamen y el tálamo, que tienen un rol crucial en la emoción, la memoria y la función cognitiva.
- **Lóbulo frontal, temporal y ganglios basales:** Relacionados con el procesamiento de la información emocional, la memoria a largo plazo, las funciones ejecutivas, la memoria y la percepción.

¿Cuál es la proporción de los conjuntos de entrenamiento y de prueba?

La proporción de los conjuntos de entrenamiento y de prueba cambian en los diferentes estudios analizados. Por ejemplo, en el conjunto de ATLAS se utilizaron

212 sujetos para el entrenamiento y 27 para las pruebas, lo que representa aproximadamente el 88% del conjunto de datos para el entrenamiento y el 12% para las pruebas. En otros estudios, como el del conjunto de datos de CIKA, la división fue establecida con 120 sujetos para el entrenamiento (54%) y 103 sujetos para las pruebas (46%). También se encuentran estudios que utilizan proporciones como 60% del conjunto de datos para el entrenamiento y 40% para las pruebas o 70% para el entrenamiento y 30% para pruebas. Además, en algunos casos se menciona el uso de la herramienta de división de datos de la biblioteca Scikit learn para asegurar una división coherente y balanceada [32] [40] [41] [44] [51].

¿Qué bases de datos de IRM de pacientes con cambios estructurales cerebrales o lesiones están disponibles públicamente para la investigación?

Entre las bases de datos utilizadas en los estudios se destacan las siguientes:

- Alzheimer's Disease Neuroimaging Initiative (ADNI): Incluye diversas colecciones de IRM clasificadas por etapa de estudio, tales como ADNI1, ADNI2, ADNI-GO y ADNI3.
- Federal Interagency Traumatic Brain Injury Research repository¹.
- Conjuntos de datos de entrenamiento encontrados en IXI que contiene imágenes de T1, T2 y DWI².

Además de bases de datos como el National Alzheimer's Coordinating Center (NACC) y Open Access Series of Imaging Studies (OASIS) que se puede encontrar online ³.

¿Existe una implementación disponible del modelo de segmentación utilizado?

Se encontraron dos implementaciones disponibles de los modelos de segmentación utilizados. Para la segmentación de hiperintensidades en la materia blanca se utilizó el modelo 2D VB-Net ⁴. Además de otra implementación de un modelo de segmentación de lesiones utilizando una arquitectura 3D ⁵ [35] [36].

¹ <https://fitbir.nih.gov>

² <http://brain-development.org/ixi-dataset/>

³ <https://github.com/haohaohuang-tj/WMH>

⁴ <https://github.com/simonsf/wmh-segmentation>

⁵ https://github.com/marianocabezas/rethinking_dsc

¿Cuáles son las métricas o criterios de evaluación utilizados para medir el rendimiento de los métodos de segmentación en IRM?

- **Recuperación (Recall):** También conocida como sensibilidad, mide la porción de vóxeles de casos positivos que se comprueban con exactitud como positivos.
- **Precisión (Precision):** Mide la porción de vóxeles segmentados que efectivamente pertenecen a las regiones de verdad fundamental.
- **Especificidad:** Mide la capacidad del modelo para identificar correctamente los pixeles que no pertenecen a la clase de interés.
- **Coeficiente de Dice (Dice Coefficient):** Evalúa la superposición entre las regiones de segmentación y las regiones de verdad fundamental o referencia.
- **Área bajo la curva (AUC):** Evalúa el rendimiento del modelo en términos de la tasa de verdaderos positivos frente a la tasa de falsos positivos.
- **Distancia de Hausdorff (Hausdorff Distance):** Mide las distancias de los dos conjuntos de vóxeles de las regiones segmentadas y de las regiones de verdad fundamental con un valor de percentil del 95%.

En la Figura 8 se evidencia la forma en que se tabularon las características de la revisión de cada estudio.

Nombre del artículo	Estructuras/regiones cerebrales más afectadas en pacientes con Trauma Craneoencefálico	Tamaño del conjunto de datos usado	Base de datos de IRM públicas	Proporción de los conjuntos de entrenamiento y de prueba	Técnicas utilizadas para la segmentación de lesiones cerebrales relevantes en IRM	Técnicas de preprocesamiento utilizadas para la preparación de IRM
HUT: Hybrid UNet transformer for brain lesion and tumour segmentation	N/A	239 escaneos de IRM ponderados por T1-w	Anatomical Tracings of Lesion After Stroke (ATLAS) y el conjunto de datos BraTS 2020	212 sujetos para entrenamiento y 27 para prueba	Arquitectura de red neuronal profunda Hybrid UNet transformer (HUT), es una combinación de UNet y un transformador de resolución cruzada. El transformador de resolución cruzada genera dos resoluciones diferentes combinadas con las conexiones de salto de UNet.	N/A
Particle Swarm Optimization and Two-Way Fixed-Effects Analysis of Variance for Efficient Brain Tumor Segmentation	N/A	En CIKA: 223 personas contribuyeron En BraTS: 54 imágenes de gliomas de bajo grado y 220 imágenes de gliomas de alto grado	CIKA y BraTS 2015	En CIKA: 120 para entrenamiento y 103 para prueba En BraTS: 274 imágenes de entrenamiento y 110 fueron para el conjunto de prueba	K-means: Modelo de aprendizaje no supervisado que aplica particionamiento de datos no jerárquico. Además, se utiliza una técnica de optimización por enjambre de partículas (PSO) que utiliza el análisis de varianza fijo de dos vías (ANOVA) como función de aptitud. Se aplicó la técnica metaheurística PSO para identificar la región de interés (ROI), la función de aptitud que determinó el mejor bloque que podía considerarse como ROI se basó en ANOVA. Finalmente, se utilizó el método K-means para segmentar la ROI.	Skull stripping: El primer paso de la técnica propuesta es un procedimiento de eliminación del cráneo, en el cual la imagen en escala de grises se convierte en una imagen binaria utilizando un umbral fijo. Luego, se aplican dos operaciones morfológicas: llenado y erosión, a la imagen binaria. Finalmente, la imagen original se enmascara con la imagen binaria obtenida.

Figura 8: Recolección de datos con metodología PRISMA [Elaboración propia].

La recolección de datos se puede visualizar completamente en una hoja de calculo compartida por Google Drive ⁶.

⁶ <https://docs.google.com/spreadsheets/d/1Pfxf5fpw0v19NPb8NcRu-WCkVj9UTwRQNV0K4ZA0d8/edit?usp=sharing>

ESTRATEGIA PROPUESTA PARA LA SEGMENTACIÓN DE LESIONES CEREBRALES POR TCEC

En la Figura 9 se presenta el diagrama que muestra la propuesta para la segmentación de lesiones cerebrales asociadas al TCEC a partir de IRM utilizando técnicas de procesamiento de imágenes y aprendizaje profundo.

Para ello, se procede a preprocesar las IRM de los sujetos de control y de lesión del conjunto de datos descrito en la sección 3.1. Posteriormente, se entrena el modelo nnU-Net, tomado del repositorio mencionado en el pie de página¹. Finalmente, se obtiene como resultado la predicción de las lesiones del conjunto de imágenes.

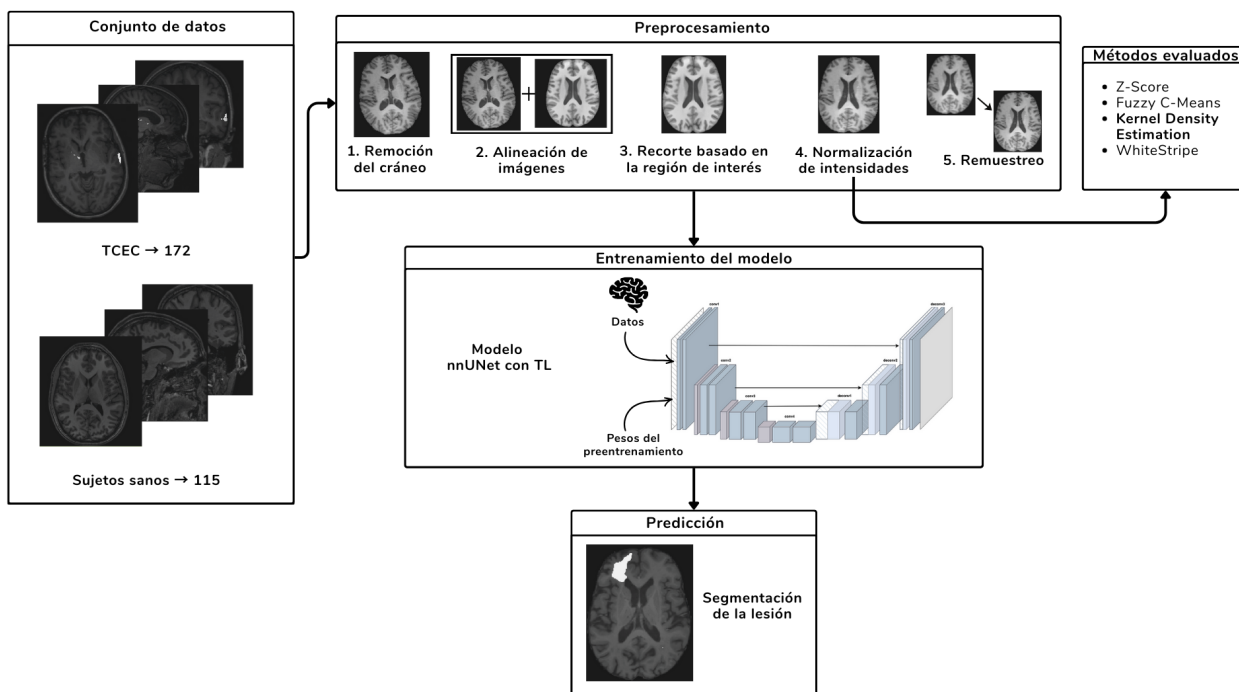


Figura 9: Flujo de trabajo propuesto para la segmentación de lesiones cerebrales asociadas al TCEC [Elaboración propia].

¹ <https://github.com/MIC-DKFZ/nnUNet>

3.1 CONJUNTO DE DATOS

El conjunto de datos seleccionado para llevar a cabo el entrenamiento y evaluación de los modelos de segmentación de lesiones cerebrales es el perteneciente a la competencia Automated Identification of Moderate-Severe Traumatic Brain Injury (Identificación Automática de Lesiones por Traumatismo Craneoencefálico Moderado-Severo) AIMS-TBI 2025 ², que forma parte de los desafíos Medical Image Computing and Computer Assisted Intervention Society (Sociedad de Computación de Imágenes Médicas e Intervención Asistida por Computadora) MICCAI 2025 ³.

El objetivo de la competencia AIMS-TBI es generar algoritmos que puedan detectar y segmentar con precisión lesiones asociadas al TCE que incluyen contusiones, hemorragias, hematomas, encefalomalacia, gliosis, lesiones de la sustancia blanca y tractos de drenaje quirúrgico. Esta competencia se centra en la identificación de lesiones únicamente en datos de resonancia magnética ponderada en T1 (T1-w), aprovechando los datos compartidos dentro del grupo de trabajo sobre lesiones cerebrales traumáticas del consorcio ENIGMA (Enhancing Neuro-Imaging Genetics through Meta-Analysis) específicamente de los subgrupos de trauma leve pediátrico y del adulto [52].

3.1.1 *Adquisición de datos y anotaciones de referencia*

La adquisición de los datos para la competencia AIMS-TBI 2025 se realizó en múltiples centros utilizando diferentes escáneres de resonancia magnética. Entre los centros están: la Escuela de Medicina de Baylor (Baylor College of Medicine), la Universidad de California en Los Ángeles (University of California Los Angeles), la Universidad Estatal de Pensilvania (Pennsylvania State University), el instituto Kennedy Krieger (Kennedy Krieger Institute), la Universidad de Loma Linsa (Loma Linda University), el Hospital Infantil Nationwide (Nationwide Children's Hospital), el Instituto de Investigación Infantil Murdoch (Murdoch Children's Research Institute), la Universidad Deakin (Deakin University), la Universidad de Texas en Houston (University of Texas Houston), la Fundación Kessler (Kessler Foundation), el Hospital de Asuntos de Veteranos de Palo Alto (VA Palo Alto), y la Universidad de Oslo (University of Oslo).

Los datos incluyen IRM ponderadas en T1, adquiridas con escáneres de 1.5T y 3T. La obtención de las imágenes fue realizada por profesionales sanitarios

² <https://aims-tbi25.grand-challenge.org/>

³ <https://conferences.miccai.org/2025/en/challenges.asp>

o radiólogos cualificados en cada una de las instituciones mencionadas. La identificación de lesiones para las anotaciones de referencia se realizó por medio de segmentación automatizada y revisión manual por parte de expertos. Primero, una red U-Net básica generó una segmentación inicial usando ANTsPy/ANTsPyNet. Luego, el personal de la Universidad de Utah revisó y corrigió manualmente las máscaras siguiendo un protocolo estandarizado en donde cada imagen fue revisada por dos anotadores formados, y posteriormente revisada y aprobada por un anotador experto. Todas las anotaciones fueron hechas con la herramienta ITK-SNAP, y las imágenes fueron desfiguradas con pydeface ⁴ para proteger la privacidad de los sujetos, sin aplicar ningún otro preprocesamiento.

Como se muestra en la tabla 3, las IRM fueron adquiridas con equipos de tres fabricantes diferentes y todas tienen las mismas dimensiones y tamaño del vóxel.

Tabla 3: Características técnicas de las secuencias de IRM

SCAN	SUJETOS	DIMENSIONES	TAMAÑO DEL VÓXEL mm^3
GE	109	$256 \times 256 \times 256$	$1 \times 1 \times 1$
PHILIPS	9	$256 \times 256 \times 256$	$1 \times 1 \times 1$
SIEMENS	169	$256 \times 256 \times 256$	$1 \times 1 \times 1$

3.1.2 Características demográficas y clínicas

El conjunto de datos contiene imágenes de sujetos menores de edad, sin embargo, se decidió excluirlas en el proyecto debido a que las estructuras cerebrales presentan variaciones según la edad, por lo que el modelo se enfoca únicamente en sujetos adultos. Se seleccionaron en total 287 sujetos, de los cuales 121 son mujeres y 166 son hombres. Su rango de edad varía desde los 18 hasta los 85 años, con una edad promedio de 43.4 años ($DE \pm 16.4$).

En la Tabla 4 se presentan las características demográficas y clínicas de los sujetos que conforman el conjunto de datos. Asimismo, en la Tabla 5 se detalla la distribución del número de sujetos divididos por rangos de edad y sexo.

⁴ <https://pypi.org/project/pydeface/>

Tabla 4: Características demográficas del conjunto de datos

SEXO	NÚMERO DE SUJETOS	RANGO DE EDAD	EDAD PROMEDIO	TSI PROMEDIO (DÍAS)
Femenino	121	18 – 78	41.9 SD $\pm$ 16.5	407.6, SD $\pm$ 684.2
Masculino	166	18 – 85	44.5 SD $\pm$ 16.3	414.4, SD $\pm$ 610.6
Todos	287	18 – 85	43.4 SD $\pm$ 16.4	411.5, SD $\pm$ 641.6

Tabla 5: Distribución de los sujetos por sexo y rangos de edad

RANGO	NÚMERO DE SUJETOS	MUJERES	HOMBRES
18 – 31	85	36	49
32 – 47	77	35	42
48 – 63	95	41	54
64 – 78	29	9	20
85	1	0	1
Total	287	121	166

En la Tabla 6 se muestra la división del conjunto de datos para el entrenamiento y las pruebas del modelo. La partición fue del 80% para entrenamiento y 20% para pruebas. Del total de sujetos que presentan lesiones (172), 140 se asignaron al conjunto de entrenamiento (82 hombres y 58 mujeres) y 32 al de pruebas (20 hombres y 12 mujeres). Por otra parte, de los sujetos sanos o de control (115), 90 se asignaron al conjunto de entrenamiento (50 hombres y 40 mujeres) y 25 al de pruebas (14 hombres y 11 mujeres).

Tabla 6: División del conjunto de datos para entrenamiento y pruebas

TIPO	TOTAL DE SUJETOS	CONTROL	LESIÓN
Sujetos para entrenamiento (80%)	230	90	140
Sujetos para pruebas (20%)	57	25	32
Total	287	115	172

En la Tabla 7 se encuentran las etiquetas de referencia originales del conjunto de datos. La etiqueta 0 corresponde al fondo, es decir, a regiones sin lesión. Asimismo, con base en el conocimiento del Dr. Alejandro Herrera, se identificó que las etiquetas del 1 al 5 corresponden a diferentes tipos de lesiones, como las mencionadas anteriormente. Es así, que algunos sujetos presentan múltiples tipos de lesiones, por lo que tienen asignadas varias etiquetas.

Debido a que el objetivo del modelo es detectar y segmentar cualquier lesión cerebral relacionada al TCEC independientemente de su tipo, se decidió mapear todas las etiquetas de lesión (2 a 5) en una sola etiqueta (1). De esta manera, el modelo solo maneja dos clases: fondo (0) y lesión (1). En la Figura 10 se muestra un ejemplo de un sujeto con múltiples lesiones antes y después de mapear las etiquetas.

Tabla 7: Etiquetas de referencia del conjunto de datos

ETIQUETA	CANTIDAD
0	115
0, 1	102
0, 3	1
0, 5	11
0, 1, 2	1
0, 1, 5	54
0, 1, 2, 5	1
0, 1, 3, 5	1
0, 1, 4, 5	1

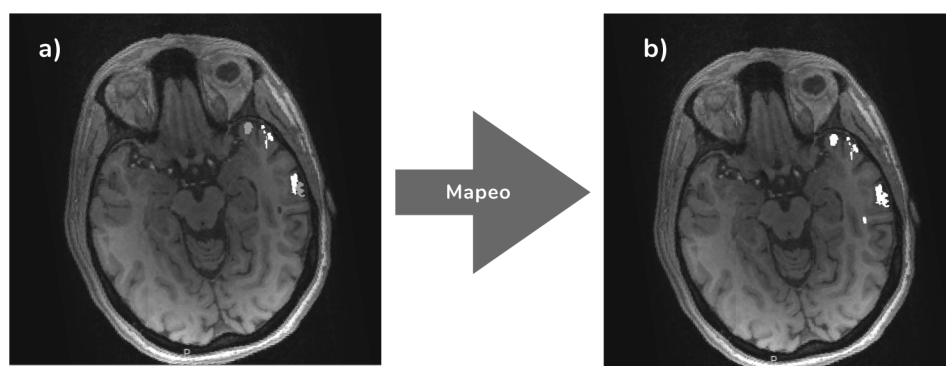


Figura 10: Visualización de etiquetas de una IRM antes y después del mapeo [Elaboración propia].

3.2 PREPROCESAMIENTO

Las IRM de los conjuntos de entrenamiento y de prueba fueron procesadas siguiendo los procedimientos que se describen a continuación, con el objetivo de estandarizar los datos, mejorar su calidad y asegurar que todas las imágenes se encuentren en condiciones comparables.

3.2.1 Remoción del cráneo

Las IRM sin procesar normalmente incluyen tejidos que no pertenecen al cerebro, como la piel, los músculos, el cuello y los globos oculares, lo que puede ser un obstáculo para los modelos de segmentación automática. Por esta razón, la remoción del cráneo es un paso fundamental del preprocesamiento, para el que se utilizó un sistema robusto de extracción cerebral basado en aprendizaje (ROBEX), en las imágenes del conjunto de datos AIMS-TBI [53]. ROBEX es un enfoque híbrido para la remoción del cráneo de IRM ponderada en T₁, está diseñado para ser robusto frente a las variaciones de intensidad y para funcionar adecuadamente en todos los conjuntos de datos. El método fue entrenado con un dataset de 92 IRM de sujetos sanos, delineadas manualmente por un fisiólogo experto, y evaluado utilizando los siguientes conjuntos de datos: el Repositorio de Segmentación Cerebral en Internet (IBSR), el conjunto LPBA40 y el conjunto transversal de resonancia magnética del proyecto OASIS.

En la Figura 11 se muestran los pasos que realiza ROBEX al procesar una imagen, y en la Figura 12 se presenta un ejemplo de una IRM del conjunto AIMS-TBI antes y después del preprocesamiento realizado por el sistema.

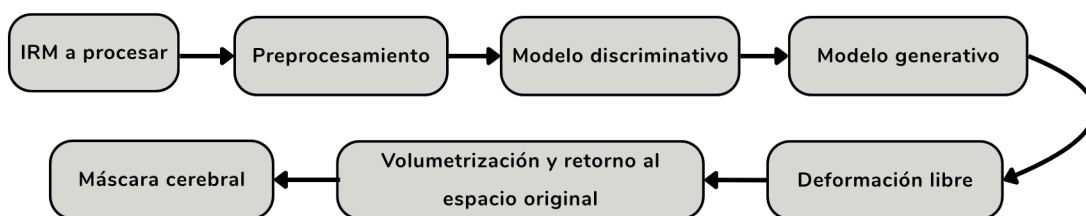


Figura 11: Pasos de ROBEX para segmentar todo el cerebro a partir de una imagen [Elaboración propia].

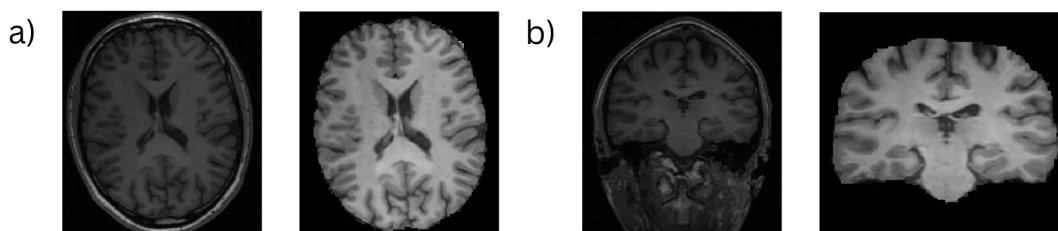


Figura 12: Ejemplo de una IRM del conjunto AIMS-TBI antes y después de la remoción del cráneo realizada por ROBEX [Elaboración propia].

3.2.2 Alineación de imágenes

La alineación de imágenes o registro es un proceso cuyo objetivo es superponer múltiples IRM a una plantilla o atlas de referencia, para que todas se ajusten al mismo sistema de coordenadas y compartan una orientación espacial común. Esto permite ubicar los sujetos dentro del mismo espacio y comparar sus datos, facilitando y mejorando el desempeño de la segmentación. Para registrar las imágenes del conjunto de datos AIMS-TBI se utilizó el método de normalización de imágenes simétricas (SyN) implementado en el paquete Advanced Normalization Tools (ANTs) ⁵. Este algoritmo es reconocido en la literatura por su alta precisión en aplicaciones de imágenes médicas.

SyN es un método de registro que deforma tanto la imagen fija (normalmente el atlas o plantilla) como la imagen movable (las IRM) en direcciones opuestas, hasta encontrarse en un punto medio. Gracias a esto, el resultado del registro es simétrico y no depende de cuál imagen se elija como fija o movable, a diferencia de métodos como Demons que solo deforman la imagen movable hacia la fija, por lo que su resultado sí se ve afectado por esta elección [54].

⁵ <https://github.com/ANTsX/ANTs>

En evaluaciones comparativas realizadas en [55], se analizaron diferentes algoritmos de registro considerando factores como la variabilidad anatómica inter-sujeto, variaciones en las intensidades de los vóxeles, el ruido, la presencia de estructuras no cerebrales, las diferencias de protocolo en bases de datos multi-sitio y la ausencia de correspondencias debido a patologías. En estas situaciones, el método SyN mostró un desempeño competitivo destacando en escenarios patológicos, donde fue el algoritmo más preciso, alcanzando el mayor nivel de coincidencia con expertos, por ese motivo, fue seleccionado para este trabajo.

En la Figura 13 se presenta un ejemplo del registro de un sujeto con el atlas SRI24 utilizando ANTs.

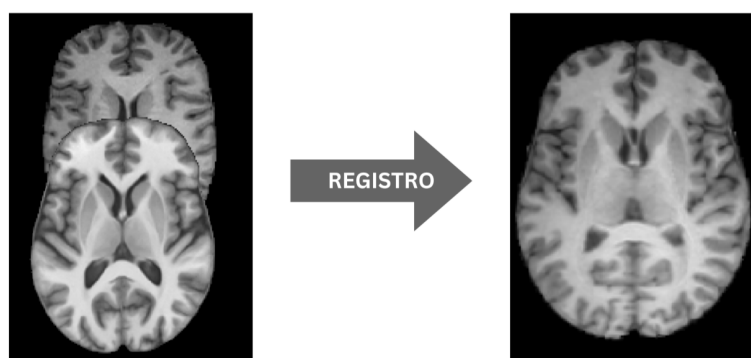


Figura 13: Ejemplo del registro de un sujeto con el atlas SRI24 utilizando ANTs [Elaboración propia].

3.2.3 Elección del atlas de referencia

Para la plantilla a utilizar para el registro de las imágenes se tuvieron en cuenta diferentes atlas cerebrales, los cuales fueron analizados previamente en el trabajo de grado titulado *Creation and Characterisation of a Colombian Brain Atlas Based on Magnetic Resonance Imaging: A Study of Southwestern Colombian Population* [56]. En la Tabla 8 se presentan las características principales de cada uno de los atlas, incluyendo el número de sujetos que los componen, su sexo, edad, origen y disponibilidad pública.

Tabla 8: Características demográficas de los atlas cerebrales considerados para la alineación de imágenes

ATLAS	SUJETOS	SEXO	EDAD	ORIGEN	PÚBLICO
Talairach and Tournoux	1	1F	60	Francia	No
ICBM152	152	86 M 66 F	18 - 44	129 Canadá 15 Asia 1 Mixto	Sí
Chinese56	56	56M	20 – 30	China	Sí
SRI24	24	6M 6F 6M 6F	19 – 33 67 – 84	América	Sí
MNI305	305	239M 66F	19 – 28	Canadá	Sí
Colin27	1	M	28	Canadá	Sí
IBA100	100	50M 50F	21 – 30	India	Sí
Chinese2020	1000	N/A	20 – 75	China	No
Korean96	96	48M 48F	> 60	Corea	No

Para las experimentaciones del proyecto se tuvieron en cuenta particularmente los atlas Chinese2020, ICBM152 y SRI24, debido a que todos fueron construidos a partir de sujetos adultos con rangos de edad similares a los del conjunto de datos AIMS-TBI. El atlas Chinese2020 incluye sujetos entre los 20 y 75 años, cubriendo casi todo el rango del conjunto de datos, sin embargo, no está disponible públicamente. Por otro lado, el atlas ICBM152 ⁶ abarca edades entre los 18 y 44 años, que es una parte importante del rango requerido, y ofrece un equilibrio en la distribución entre hombres y mujeres. Finalmente, el atlas SRI24 ⁷ incluye sujetos entre los 19 y 33 años y entre los 67 y 84 años, cubriendo los extremos del rango y presentando también una proporción equilibrada entre hombres y

⁶ <https://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152Nlin2009>

⁷ <https://www.nitrc.org/projects/sri24>

mujeres. Los detalles y propiedades de los dos atlas seleccionados en este trabajo para evaluación se observan en la Figura 14.

a) --- Análisis del archivo: ICBM152_atlas.nii.gz

Dimensiones: (193, 229, 193)
Tamaño de voxel (mm): (1.0, 1.0, 1.0)

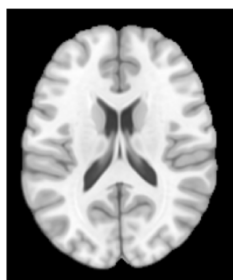
Matriz afín (direcciones y origen):

```
[[ 1.  0. -0. -96.]
 [ 0.  1. -0. -132.]
 [ 0.  0.  1. -78.]
 [ 0.  0.  0.  1.]]
```

Origen (mm): [-96. -132. -78.]

Orientación: ('R', 'A', 'S')

Tipo de datos: float32



b) --- Análisis del archivo: SRI24_atlas.nii.gz

Dimensiones: (240, 240, 155, 1)
Tamaño de voxel (mm): (1.0, 1.0, 1.0)

Matriz afín (direcciones y origen):

```
[[ 1.  0.  0. -120.]
 [ 0.  1.  0. -129.]
 [ 0.  0.  1. -68.]
 [ 0.  0.  0.  1.]]
```

Origen (mm): [-120. -129. -68.]

Orientación: ('R', 'A', 'S')

Tipo de datos: int16

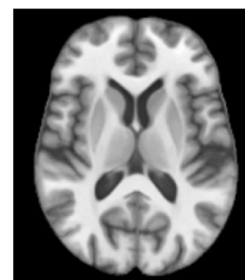


Figura 14: Propiedades de los atlas (a) CBM152 y (b) SRI24 [Elaboración propia].

3.2.4 Recorte basado en la región de interés

El recorte basado en la región de interés (ROI) tiene como objetivo eliminar o ignorar los vóxeles que no pertenecen a la región de interés, que en este caso es el cerebro. Normalmente, los vóxeles que son irrelevantes pertenecen al fondo. Al realizar el recorte, se reduce la cantidad de información no útil que el modelo debe procesar, mejorando el tiempo de procesamiento, disminuyendo el espacio usado en memoria y permitiendo que el modelo se enfoque en las características más importantes para la segmentación.

El proceso empieza inicializando una máscara binaria del mismo tamaño que la imagen de entrada, donde los vóxeles dentro del ROI se establecen en 1 y los que están fuera se establecen en 0. Luego, utilizando la máscara, se genera una caja mínima rectangular que contiene toda la región etiquetada con 1. Después, la imagen original se multiplica por la máscara binaria, produciendo una nueva imagen donde solo los vóxeles dentro del ROI con-

servan sus valores originales, mientras que los vóxeles fuera del ROI se eliminan.

En el caso de la segmentación de referencia, esta se recorta utilizando la misma máscara binaria. Sin embargo, los vóxeles que estaban fuera de la región ROI se marcan por defecto como -1 , que corresponde a una etiqueta para ignorar, la etiqueta 0 para el fondo real y la etiqueta 1 para las lesiones. En la Figura 15 se presenta un ejemplo del recorte realizado en este trabajo basado en el ROI de una IRM del conjunto de datos.

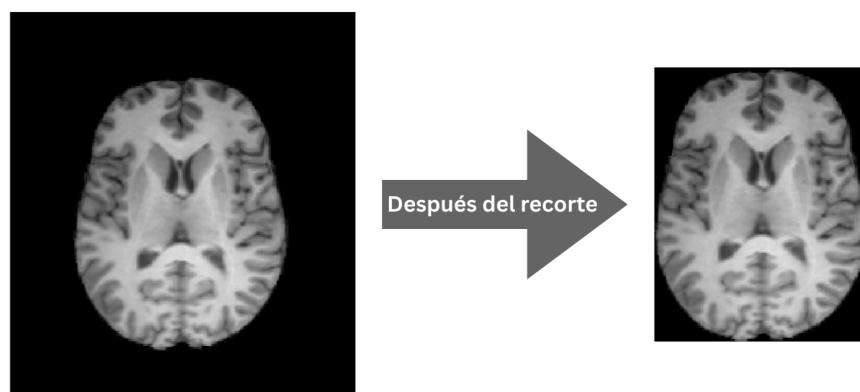


Figura 15: Ejemplo del recorte basado en el ROI de una IRM [Elaboración propia].

3.2.5 Normalización de intensidades

Los valores de intensidad de los vóxeles en las IRM no se encuentran en una escala estandarizada, sino que son arbitrarios y pueden variar según el fabricante del escáner, la intensidad del campo y los parámetros de adquisición de secuencia de la imagen. Esto ocasiona que las intensidades no solo varíen entre pacientes, sino también entre diferentes exámenes del mismo paciente, lo que puede dificultar la detección de información biológica y patológica.

Es por esto, que en el preprocesamiento del conjunto de datos es importante incluir la normalización de intensidades, utilizada para transformar los datos a una distribución normal estándar. Este proceso garantiza que todas las intensidades se encuentren en la misma escala y evita que ciertas características predominen sobre otras. De esta manera, el modelo no se ve afectado por cambios o variaciones presentadas en las diferentes IRM. Asimismo, la normalización tiene un papel significativo en el rendimiento de los métodos de procesamiento automatizado de imágenes, porque se aplica como un paso de

preparación del cual dependen los resultados posteriores [57].

Los métodos de normalización utilizados fueron seleccionados al analizar el artículo [58], en el cual se evalúa el efecto de estrategias de estandarización de intensidad en la cuantificación de la atrofia longitudinal, demostrando que la normalización puede conducir a una mejor cuantificación de la atrofia. En este proyecto se decidió evaluar estos mismos métodos en el conjunto de datos AIMS-TBI, con el objetivo de observar su impacto en la segmentación de lesiones cerebrales traumáticas. Es importante mencionar que todas las técnicas fueron implementadas utilizando la librería Intensity Normalization⁸ presentada en [59]. Las técnicas de normalización utilizadas en este trabajo fueron:

3.2.5.1 *Z-Score*

Z-score es un método de normalización muy utilizado que transforma las intensidades para que tengan una media cero y una desviación estándar unitaria. El proceso consiste en que, dada una imagen de entrada I , se resta el valor medio de intensidad de cada vóxel y el resultado se divide por la desviación estándar. La fórmula se presenta en la Ecuación 1, donde μ y σ son la media y la desviación estándar de las intensidades dentro del volumen intracraneal.

$$I_{\text{norm}} = \frac{I - \mu}{\sigma} \quad (1)$$

3.2.5.2 *Fuzzy c-means FCM*

La normalización basada en FCM utiliza la agrupación difusa de la materia blanca. El proceso comienza cuando se lleva a cabo el agrupamiento difuso de c-medias. Dado un conjunto de vóxeles $x_i \in I$, FCM encuentra un conjunto de K clústeres, donde $C = \{c_1, \dots, c_K\}$, minimizando la función de pérdida presentada en la Ecuación 2, donde C_k es el centroide del clúster K y $W_{ik} \in [0, 1]$ es el grado de pertenencia de x_i del clúster k .

$$\arg \min_C \sum_{x_i \in I} \sum_{c_k \in C} w_{ik} \|x_i - c_k\|_2^2 \quad (2)$$

⁸ <https://github.com/jcreinhold/intensity-normalization>

El grado de pertenencia de un vóxel a un clúster se calcula utilizando la Ecuación 3.

$$w_{ik} = \frac{1}{\sum_{j=1}^K \frac{\|x_i - c_k\|_2^2}{\|x_i - c_j\|_2^2}} \quad (3)$$

El proceso se realiza iterativamente hasta que los valores no tengan un cambio significativo. Luego, se divide todas las intensidades de los vóxeles por la intensidad media detectada de la sustancia blanca. La selección de k puede depender de las regiones de interés que se vayan a segmentar o de la presencia de tejido anormal. En este trabajo se estableció el $K = 3$ (sustancia blanca, sustancia gris y líquido cefalorraquídeo), como en la implementación original [59].

3.2.5.3 Kernel Density Estimation KDE

Este método estima una función de densidad de probabilidad a partir de puntos de datos (en este caso, las intensidades). La función de densidad de probabilidad $p(x)$ se calcula utilizando la Ecuación 4, donde K es el núcleo de suavizado, h es el parámetro de suavizado y $\text{card}()$ es la cardinalidad del conjunto de intensidades. La contribución de cada punto depende de sus vecinos: entre más cerca estén los puntos de x , mayor será el valor de $p(x)$. El parámetro h controla el grado de suavidad de $p(x)$: la función de densidad será más suave con un valor de h más grande. Luego de que el método calcula $p(x)$, identifica el modo (o pico más alto) que se encuentra hacia la derecha de la distribución, el cual representa el modo de la materia blanca en imágenes T1-w, este valor se utiliza como base para la normalización. En este trabajo se utilizó un núcleo gaussiano y un $h = \max(I)/80$, como se establece en la implementación original [59].

$$p(x) = \frac{1}{\text{card}(I)} \sum_{x_i \in I} \mathcal{K}\left(\frac{x - x_i}{h}\right) \quad (4)$$

3.2.5.4 WhiteStripe WS

El método WhiteStripe estima las estadísticas de las intensidades de la sustancia blanca para normalizar todas las demás regiones de interés. El proceso consta de

tres pasos. Primero, se suaviza el histograma mediante splines penalizados para identificar el modo correspondiente a la sustancia blanca, evitando que el ruido de la imagen o el modo correspondiente al fondo afecten la estimación. Para esto, se ignora el primer modo, que normalmente pertenece al fondo, y se selecciona el último, que en imágenes T1-w representa la sustancia blanca, μW_s . Segundo, se estima la desviación estándar σW_s , a partir de una ventana que cubre el 10% de los valores de intensidad alrededor del modo detectado. Por último, todas las intensidades se normalizan mediante la operación de Z-score presentada en la Ecuación 1, utilizando las dos estadísticas obtenidas.

3.2.6 Remuestreo

El remuestreo se utiliza para cambiar el tamaño de los vóxeles de una imagen sin modificar sus límites espaciales en el sistema de coordenadas. Este paso es importante dentro del preprocesamiento debido a que el tamaño del vóxel se ha identificado como una de las fuentes de variabilidad con mayor impacto en la reproducibilidad del análisis de IRM [60]. En conjuntos de datos de diferentes centros, las imágenes pueden tener diferentes tamaños de vóxel por el uso de diversos escáneres y resoluciones. Por esta razón, es importante que el modelo sea robusto a estas variaciones y estandarice las imágenes para que sean comparables. De acuerdo con lo anterior, se incluyó el remuestreo de las imágenes a un tamaño de vóxel estándar de $1 \times 1 \times 1 \text{ mm}^3$. En este proceso se utiliza una interpolación suave en las imágenes de los sujetos y una interpolación por vecinos más cercanos en sus respectivas segmentaciones. En los casos en que las imágenes son anisotrópicas, es decir, cuando su espaciado no es igual en los tres ejes y supera un umbral predefinido, la interpolación se aplica por cada corte en dimensión 2D para evitar artefactos o distorsiones.

Es importante señalar que, en el conjunto de datos AIMS-TBI, todas las imágenes tienen un tamaño de vóxel de $1 \times 1 \times 1 \text{ mm}^3$, por lo que no fue necesario aplicar remuestreo en este caso particular. Sin embargo, el remuestreo se decidió incluir en el flujo de preprocesamiento para que el modelo pueda adaptarse a estos casos y no se vea afectado por dicha variabilidad.

3.3 MODELOS DE SEGMENTACIÓN

En este trabajo se emplearon los modelos nnU-Net y 3D U-Net. Ambos siguen la misma estructura básica de U-Net, ampliamente utilizada en el campo de la segmentación médica, y permiten trabajar directamente con imágenes en tres dimensiones. Es importante señalar que no se seleccionaron los modelos encon-

trados en la literatura porque estos han sido diseñados para detectar tipos de lesiones muy específicas, mientras que el conjunto de datos utilizado incluye diferentes tipos de lesión. Por esta razón, dichos métodos no resultarían adecuados o podrían no ofrecer un buen desempeño. A continuación, se describen con mayor detalle los modelos seleccionados.

3.3.1 No New U-Net (nnU-Net)

No New U-Net (nnU-Net) es un framework de segmentación de imágenes médicas basado en la arquitectura estándar U-Net [61]. El framework es de código abierto y está disponible en el enlace presentado en el pie de página ⁹.

nnU-Net está diseñado para tratar con diferentes modalidades como MRI, CT, microscopía electrónica y fluorescencia, y generar máscaras de segmentación de las estructuras o regiones de interés como los órganos, tumores, lesiones y células. Su propósito es facilitar la tarea de segmentación en escenarios de investigación clínica, sin tener que diseñar y ajustar manualmente cada componente del proceso.

En la Figura 16 se presenta su flujo de procesamiento, formado a partir de parámetros fijos, basados en reglas y empíricos, los cuales se describen a continuación.

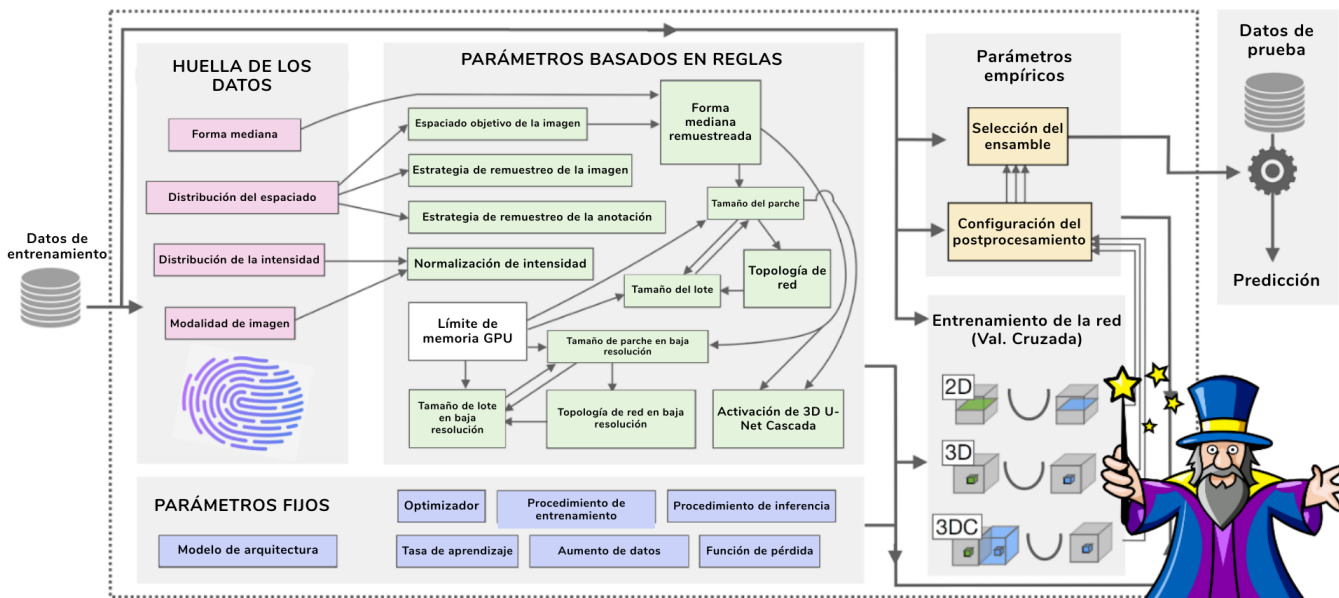


Figura 16: Proceso de nnU-Net para la segmentación automática de imágenes médicas [61].

⁹ <https://github.com/MIC-DKFZ/nnUNet>

3.3.1.1 *Huella digital de los datos*

La primera etapa en el flujo del modelo es extraer la huella digital, que consiste en identificar características del conjunto de datos y, con base en ello, determinar los pasos que deben ejecutarse posteriormente. Esta huella se contruye a partir de una muestra máxima de $1e^7$ vóxeles seleccionados aleatoriamente, aplicando una semilla fija para garantizar la reproducibilidad. Para cada modalidad de imagen, el modelo calcula el tamaño espacial promedio de las imágenes después del recorte (forma mediana), el promedio del tamaño de los vóxeles (distribución del espaciado) y la distribución de intensidades obteniendo estadísticas como la media, la desviación estándar y percentiles del 5% y el 9.5%. De esta manera, el modelo determina si el conjunto de datos es anisotrópico o isotrópico, y como debe manejar las intensidades y la relación entre tamaños de imagen.

3.3.1.2 *Parámetros fijos*

Los parámetros fijos no dependen de la información del dataset, estos ya están establecidos por los desarrolladores del framework. Entre estos se encuentra el modelo de arquitectura, el optimizador, la tasa de aprendizaje, la función de pérdida, el aumento de los datos, y el procedimiento de entrenamiento e inferencia.

3.3.1.3 *Parámetros basados en reglas*

El modelo establece los parámetros basados en reglas con base en la huella de los datos. A continuación, se presentan dos ejemplos de parámetros que condicionan a otros:

- La distribución del espaciado: Se utiliza para determinar la estrategia de remuestreo de las imágenes y sus etiquetas.
- El límite de memoria GPU y la forma mediana remuestreada: Condicionan el tamaño del parche, el tamaño del lote y la topología de la red como su profundidad y tamaño del núcleo.

Las decisiones se fundamentan en un conjunto de reglas heurísticas. Por ejemplo, en el remuestreo, si el conjunto de datos es anisotrópico, se aplica una interpolación continua en el plano y remuestreo con vecino más cercano fuera del plano para las etiquetas. Asimismo, para la selección del tamaño del parche, el tamaño del lote y la topología de la red el modelo sigue un procedimiento iterativo, donde el parche se inicializa con la forma mediana de las imágenes y se va reduciendo hasta ajustarse al límite de memoria disponible, adaptando la topología en cada iteración. Estas reglas permiten que nnU-Net genere configuraciones coherentes sin la búsqueda manual.

3.3.1.4 *Parámetros empíricos*

De acuerdo con la huella de los datos y los parámetros basados en reglas, nnU-Net genera un plan para cada configuración de entrenamiento. El entrenamiento 2D consiste en procesar la imagen por cortes, mientras que el entrenamiento 3D utiliza el volumen completo. La opción 3D Cascada se activa solamente cuando las imágenes son muy grandes, y el procedimiento se realiza entrenando primero una red en baja resolución y luego otra en la resolución original para refinar los detalles.

Por otra parte, el framework permite entrenar estas configuraciones con o sin validación cruzada, con el fin de seleccionar empíricamente la mejor variante o el mejor ensamble según las métricas de validación. Finalmente, decide si aplicar postprocesado basándose en si esto mejora los resultados de validación.

3.3.2 *3D U-Net*

El modelo 3D U-Net fue desarrollado por Adam Wolny y otros colaboradores como parte de la investigación en eLife en 2020 [62]. Su propósito fue ofrecer una implementación flexible, eficaz y tridimensional de la arquitectura U-Net, diseñada especialmente para segmentar estructuras biológicas. Esta iniciativa fue creada para permitir el uso de redes 3D en estudios que requieren modelos reproducibles, configurables y aptos para manejar grandes cantidades de datos biomédicos sin tener que programar una arquitectura desde el inicio.

El modelo implementa una versión completa de 3D U-Net¹⁰, una arquitectura que se basa en un codificador-decodificador con conexiones de salto, lo que permite la combinación de información global y detalles locales de la imagen. Consiste en convoluciones tridimensionales, funciones de activación no lineales y procesos de downsampling y upsampling que permiten a la red aprender representaciones semánticas profundas sin perder la estructura espacial del volumen. Su diseño modular permite modificar algunos parámetros como profundidad, cantidad de canales, tamaño de los parches o función de pérdida a través de archivos de configuración.

3D U-Net sigue la lógica tradicional de U-Net: durante la etapa de codificación, la red disminuye gradualmente la resolución del volumen para capturar características de alto nivel, mientras que en la etapa de decodificación recupera la resolución original al combinar estas representaciones profundas con caracterís-

¹⁰ <https://github.com/wolny/pytorch-3dunet>

ticas espaciales del encoder mediante conexiones de salto. Este enfoque facilita la conservación de bordes y detalles, generando la segmentación del volumen.

3.4 TECNOLOGÍAS USADAS PARA LA IMPLEMENTACIÓN DEL FRAMEWORK

El framework fue implementado mayormente con Python, que se utilizó para automatizar el flujo de trabajo, la ejecución secuencial de los procesos y la integración de sus diferentes fases. Asimismo, se utilizó Docker como contenedor del sistema para garantizar un ambiente de procesamiento uniforme en diferentes plataformas y así asegurar la reproducibilidad de los resultados, evitando inconsistencias derivadas de diferencias en el entorno computacional.

Además del entorno tecnológico base, el framework incorpora herramientas especializadas para cada etapa del procesamiento. En particular, se empleó ROBEX para la extracción del cráneo, ANTs para el alineamiento de las imágenes con un atlas de referencia, la librería Intensity Normalization para la normalización de intensidades, y el modelo nnU-Net para la segmentación de lesiones mediante aprendizaje profundo.

EVALUACIÓN

4.1 CONFIGURACIÓN DE LA EVALUACIÓN

Para el entrenamiento y evaluación del modelo de segmentación basado en nn-Unet optimizado para lesiones asociadas al TCEC, se evaluó el efecto en el desempeño de la segmentación de (i) atlas empleado para el registro ICBM152 ó SRI24; y modalidad de nn-Unet, 2D ó 3D (ii) el método de normalización Z-score, FCM, WS, ó KDE, (iii) efecto de aprendizaje por transferencia empleando el conjunto de datos BRATS2021.

Para cada caso de prueba, los modelos fueron entrenados empleando la arquitectura estándar de U-Net definida por nnU-Net durante 1,000 épocas, con 125 iteraciones por época, un tamaño de parche de [128, 160, 112] y un tamaño de lote de 2. La optimización se ejecutó mediante el Descenso de Gradiente Estocástico (SGD) con momentum de 0.99 y Nesterov, usando un esquema polinomial para la tasa de aprendizaje con un valor inicial de $1e^{-2}$ y un decaimiento de peso de $3e^{-5}$. Asimismo, la función de pérdida fue la combinación estándar de Dice + Cross-Entropy, con un suavizado de $1e^{-5}$. Debido a que el conjunto de datos es isotrópico, el aumento de datos consistió en aplicar rotaciones 3D en el rango permitido por nnU-Net y espejamiento en los tres ejes.

Adicionalmente, como métricas de evaluación en cada caso se empleó el Coeficiente de Similitud de Dice e Índice de Jaccard. Para esto, la segmentación de referencia se comparó con la predicción del modelo, calculando los siguientes valores a partir de los vóxeles, que se utilizaron tanto en las métricas de evaluación como en el análisis de resultados.

- Verdaderos Positivos (VP): Vóxeles correctamente etiquetados como lesión.
- Verdaderos Negativos (VN): Vóxeles correctamente etiquetados como no lesión.
- Falsos Positivos (FP): Vóxeles incorrectamente etiquetados como lesión.
- Falsos Negativos (FN): Vóxeles incorrectamente etiquetados como no lesión.

4.1.1 *Coeficiente de similitud de Dice (DSC)*

El Coeficiente de Dice (DSC) es una métrica utilizada para comparar la similitud de la segmentación de referencia y la predicha. Su rango varia desde 0, que indica que no hay superposición, hasta 1, que representa una superposición perfecta. El DSC es frecuentemente utilizado en la segmentación de imágenes médicas para evaluar la capacidad del modelo para identificar las regiones segmentadas. Su fórmula se presenta en la Ecuación 5, donde los VP se multiplican por 2 para darle mayor peso a los aciertos del modelo.

$$DSC = \frac{2 VP}{2 VP + FP + FN} \quad (5)$$

4.1.2 *Índice de Jaccard (IoU)*

El Índice de Jaccard, también conocido como Intersección sobre la Unión (IoU) mide la superposición entre la segmentación predicha y la referencia, calculando la proporción de aciertos del modelo sobre todos los vóxeles que debió segmentar. Su fórmula se presenta en la Ecuación 6.

$$IoU = \frac{VP}{VP + FP + FN} \quad (6)$$

4.1.3 *Evaluación de los sujetos de control*

Se denominan sujetos de control a aquellos pacientes que no presentan ninguna lesión relacionada con el TCEC y cuya imagen cerebral corresponde a un cerebro sano. Estos casos son relevantes y fueron incluidos en el estudio porque permiten que el modelo, además de segmentar lesiones, aprenda a identificar las imágenes que no contienen ninguna alteración.

Por otra parte, las métricas de DSC e IoU no consideran como un acierto cuando el modelo identifica correctamente un sujeto sin lesión. En estos casos, al no haber vóxeles positivos en la segmentación de referencia ni en la predicción, la división resulta en 0.0/0.0, por lo que se reporta como NaN. Aún así, la evaluación se realizó de la misma forma que lo hace la competencia AIMS-TBI, donde asignan un valor de 1.0 a las métricas cuando esto sucede, debido a que el modelo identificó correctamente la ausencia total de la lesión. Por el contrario, si

el modelo etiquetó al menos un vóxel como lesión en un sujeto de control, las métricas arrojan el peor valor posible, es decir, 0.0.

4.2 EVALUACIÓN DE LA CONFIGURACIÓN DEL MODELO Y EL ATLAS DE REGISTRO

La primera evaluación consistió en determinar cuál atlas era el más adecuado para el registro, considerando las dos opciones presentadas en la sección de preprocesamiento 3.2: SRI24 e ICBM152. Al mismo tiempo, también se evaluó con cuál configuración de entrenamiento obtenía un mejor desempeño nnU-Net, comparando las opciones disponibles 2D y 3D. En esta etapa se establecieron las bases para las pruebas posteriores y se aseguró que el flujo de procesamiento partiera desde una configuración óptima. Los resultados del primer experimento se muestran en la Figura 17.

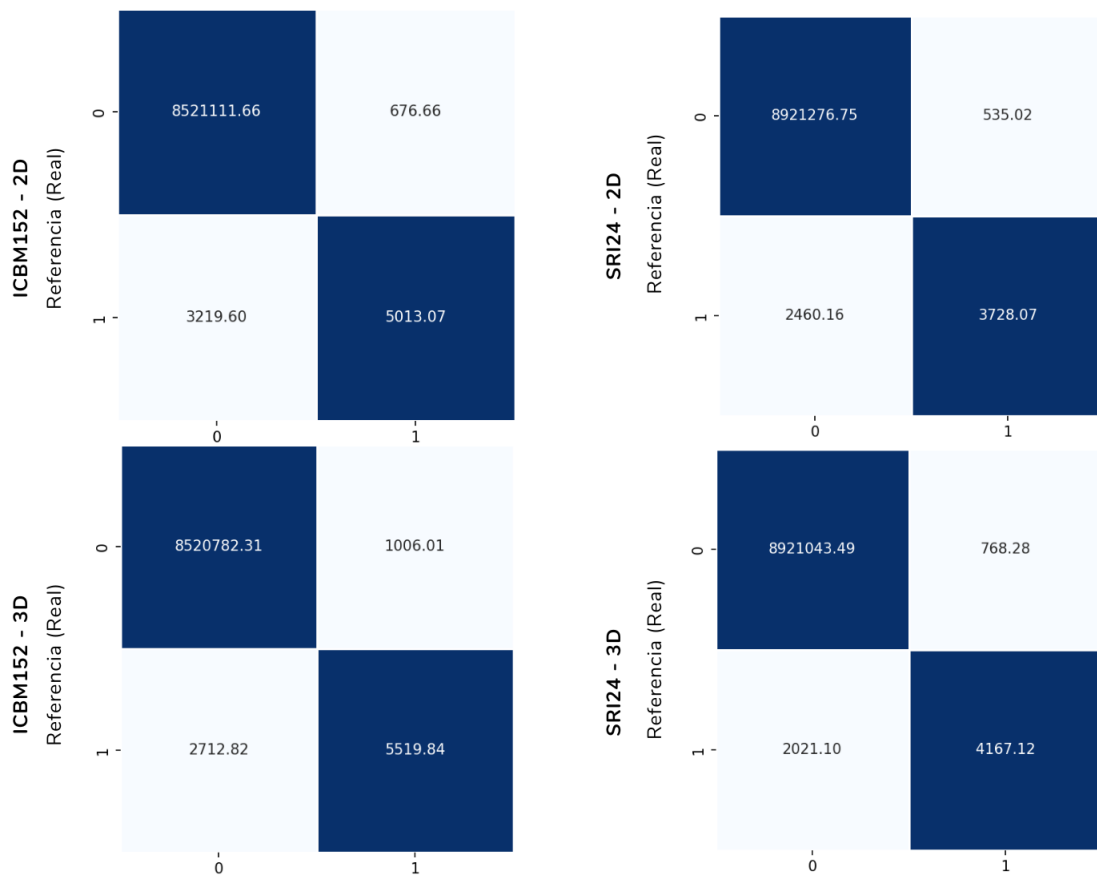


Figura 17: Matrices de confusión de la evaluación de configuración y atlas [Elaboración propia].

En el primer caso, donde las imágenes fueron alineadas con el atlas ICBM152 y el modelo se entrenó usando la configuración 2D, este reconoció adecuadamente 5,013 vóxeles positivos de un total de 8,232. No obstante, también identificó 676 FP y mostró la cifra más alta de FN, con 3,219 vóxeles con lesión que no logró detectar.

En el segundo caso, donde las imágenes fueron alineadas con el atlas ICBM152 y el modelo fue entrenado utilizando la configuración 3D, el modelo consiguió detectar correctamente el mayor número de vóxeles positivos, con 5,519 identificados de un total de 8,232. Sin embargo, también fue el segundo con el peor número de FN con 2,712 vóxeles.

En el tercer caso, donde las imágenes se registraron utilizando el atlas SRI24 y el modelo se entrenó con la configuración 2D, este identificó correctamente 3,728 vóxeles positivos de un total de 6,188, lo que representa aproximadamente el 60% de la lesión. Sin embargo, identificó erróneamente 535 vóxeles sanos como si fueran lesiones y no detectó 2,460 vóxeles con lesiones, lo que señala una pérdida significativa de áreas lesionadas.

Finalmente, en el caso en que las imágenes fueron registradas utilizando el atlas SRI24 y el modelo se entrenó con la configuración 3D, dicho modelo fue capaz de identificar de manera correcta 4,167 vóxeles positivos, lo que significa 439 más que en la configuración 2D, representando aproximadamente el 67.3% del total. Asimismo, evidenció una mejora significativa en los FN, al reducirlos a 2,021 vóxeles, lo que representa una disminución cercana al 17.8% en comparación con el modelo 2D. No obstante, produjo 768 FP, que son 233 más que en el experimento previo.

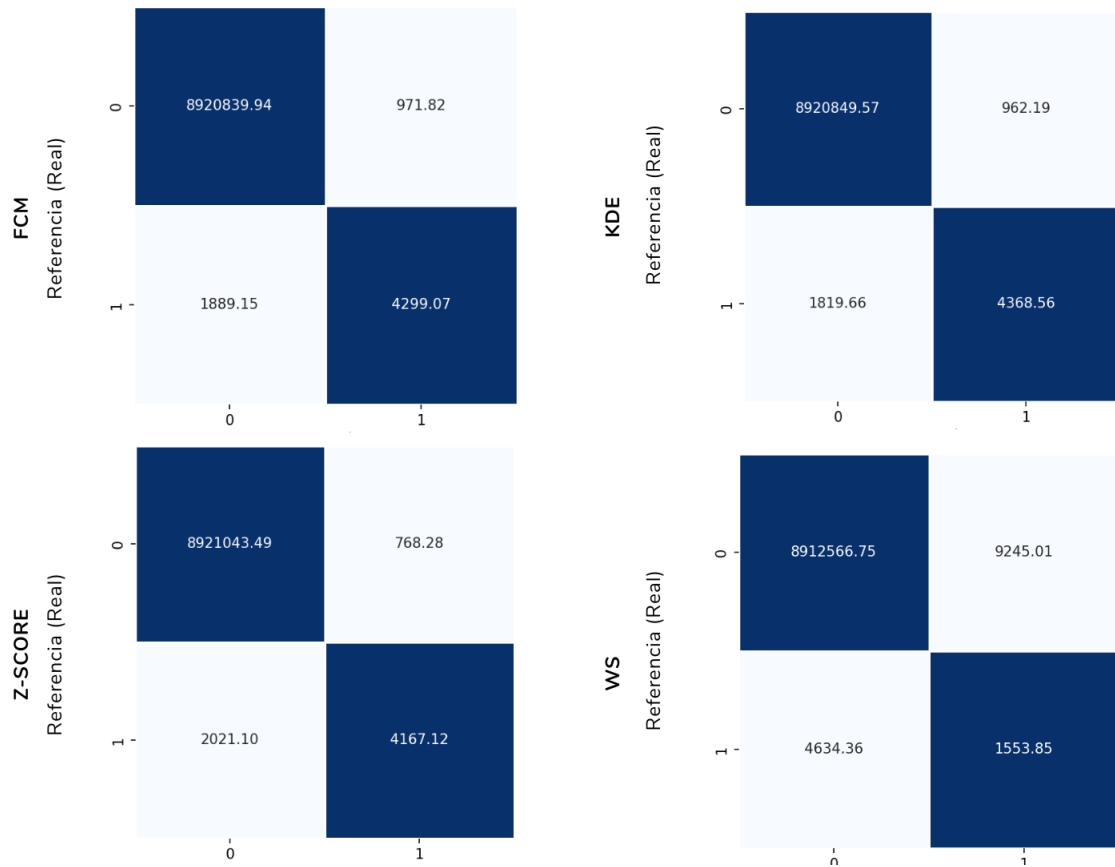
Asimismo, en la Tabla 9 se muestran los resultados de las métricas para cada situación. En términos generales, el mejor desempeño se logró con el registro que empleó el atlas SRI24 y el entrenamiento con la configuración 3D. En los modelos 2D, el atlas ICBM152 supera a SRI24 en las métricas DSC con 0.4999 e IoU con 0.4484. Sin embargo, en los modelos 3D, SRI24 muestra una ventaja clara sobre ICBM152, lo que demuestra que el atlas de registro y la configuración del modelo tienen un impacto importante en la precisión de la segmentación.

Tabla 9: Evaluación de configuración del modelo y el atlas de registro

MÉTRICA	2D - SRI	3D - SRI	2D - ICBM	3D - ICBM
DSC	0.4547	0.5262	0.4999	0.4619
IoU	0.4058	0.4712	0.4484	0.4045

4.3 EVALUACIÓN DE LOS MÉTODOS DE NORMALIZACIÓN

Con el atlas y la configuración seleccionados, se realizaron las pruebas con los métodos de normalización Z-Score, FCM, KDE y WS. En la Figura 18 se presentan los resultados de la evaluación.

**Figura 18:** Matrices de confusión de la evaluación del método de normalización [Elaboración propia].

El método con Z-score corresponde al modelo anteriormente evaluado y seleccionado en el experimento de configuración de entrenamiento y atlas de

registro: 3D y SRI24. En cuanto a FCM, el método presentó un ligero aumento en los VP en comparación con el método Z-Score. Mientras Z-Score logró 4,167 VP de un total de 6,188, FCM obtuvo 4,299 presentando una mejora del 2.13%.

Asimismo, KDE también evidenció un aumento en los VP con respecto a Z-Score. Mientras que Z-Score logró 4,167 VP, KDE obtuvo 4,368, lo que indica una mejora del 3.25%, destacando con la mejor cifra de VP. Igualmente, se detectó un leve aumento en los FP.

Por último, WS fue el que obtuvo el peor desempeño entre los demás métodos evaluados, teniendo 1,553 VP y 4,634 FN, además de 9,245 FP. Esto indica que el modelo no solo dejó de identificar gran parte de las lesiones reales, sino que también clasificó como lesión numerosas regiones donde no había ninguna.

En la Tabla 10 se observa que el método KDE alcanzó el mejor desempeño, con un DSC de 0.5322 y un IoU de 0.4750, lo que lo convirtió en la opción más adecuada para este conjunto de datos. Esto sugiere que KDE realizó la normalización más consistente de las intensidades. El siguiente que presentó un buen desempeño fue Z-Score, con un DSC de 0.5262 y un IoU de 0.4712.

Por el contrario, WS obtuvo el peor desempeño. Esto pudo ocurrir debido a que el método asume la existencia de materia blanca normal estable, lo que no resulta adecuado para cerebros con TCEC, donde la estructura de la materia blanca suele estar alterada.

Tabla 10: Evaluación de los métodos de estandarización

MÉTRICA	Z-SCORE	FCM	KDE	WS
DSC	0.5262	0.4436	0.5322	0.0602
IoU	0.4712	0.3866	0.4750	0.0355

4.4 APRENDIZAJE POR TRANSFERENCIA

Con el atlas de registro, la configuración de entrenamiento y el método de estandarización óptimos ya establecidos, se evaluó si el aprendizaje por transferencia podría mejorar el desempeño del modelo en la segmentación. Se eligió el conjunto de datos BRATs 2021, cuyo propósito es segmentar tumores cerebrales en tres subregiones. Aunque no se trata exactamente del mismo tipo de lesión, es una tarea relacionada con la segmentación de lesiones en IRM, por lo que ofre-

ce un volumen significativo de datos con características similares. El objetivo de este preentrenamiento es que el modelo mejore su capacidad de segmentación utilizando patrones aprendidos en un problema muy similar.

4.4.1 *Descripción del conjunto de datos para el aprendizaje por transferencia*

El conjunto de datos para el preentrenamiento proviene de los datos de entrenamiento de la competencia BRATs 2021, cuyo objetivo es evaluar métodos automáticos para la segmentación de gliomas cerebrales en IRM. Este conjunto incluye escaneos 3D adquiridos clínicamente de pacientes con distintos tipos de gliomas, sus anotaciones fueron generadas manualmente y validadas por expertos. Para cada paciente, se proporcionan las siguientes modalidades de imagen: T1, T1 con contraste, T2 y T2-FLAIR, sin embargo solo se utilizó la IRM ponderada en T1.

Aunque las segmentaciones originales incluyen múltiples etiquetas que representan diferentes subregiones del tumor, estas fueron mapeadas a tumor (1) y no tumor (0) para centrar la tarea en la segmentación del tumor en general, de manera consistente con el enfoque utilizado para el conjunto de datos AIMS-TBI. Las imágenes fueron descargadas del repositorio mencionado en el pie de página ¹.

4.4.2 *Parámetros del preentrenamiento*

Para llevar a cabo el preentrenamiento, nnU-Net generó un plan específico adaptado al conjunto de datos BRATs 2021, pero asegurando la compatibilidad con el plan previamente creado para AIMS-TBI, de modo que ambos conjuntos pudieran integrarse sin inconsistencias durante el entrenamiento. El preentrenamiento se realizó durante 1,000 épocas, con 250 iteraciones por época, utilizando exactamente los mismos parámetros empleados en el entrenamiento inicial con AIMS-TBI: tamaño de parche, tamaño de lote, función de pérdida, optimizador, tasa de aprendizaje y peso de decaimiento.

Una vez finalizado el proceso, los mejores pesos obtenidos en este preentrenamiento se utilizaron como punto de partida para entrenar nuevamente el modelo con el conjunto AIMS-TBI, repitiendo todo el proceso de entrenamiento completo bajo los mismos hiperparámetros descritos en el entrenamiento de nnU-Net.

¹ <https://www.kaggle.com/dschettler8845/brats-2021-task1>

4.4.3 Evaluación del aprendizaje por transferencia

En la matriz de confusión se muestra que, aunque los VP disminuyeron ligeramente a 4,359 en comparación con el modelo sin preentrenamiento, los FN mejoraron significativamente, reduciéndose a 907 cuando antes tenía 962. Esto indica que el modelo detectó mejor entre los sujetos de control y los sujetos con lesión. Este comportamiento se ve reflejado en los resultados de las métricas, donde obtuvo un DSC = 0.5642 e IoU = 0.5097, convirtiéndose en el mejor desempeño entre todas las variantes evaluadas.

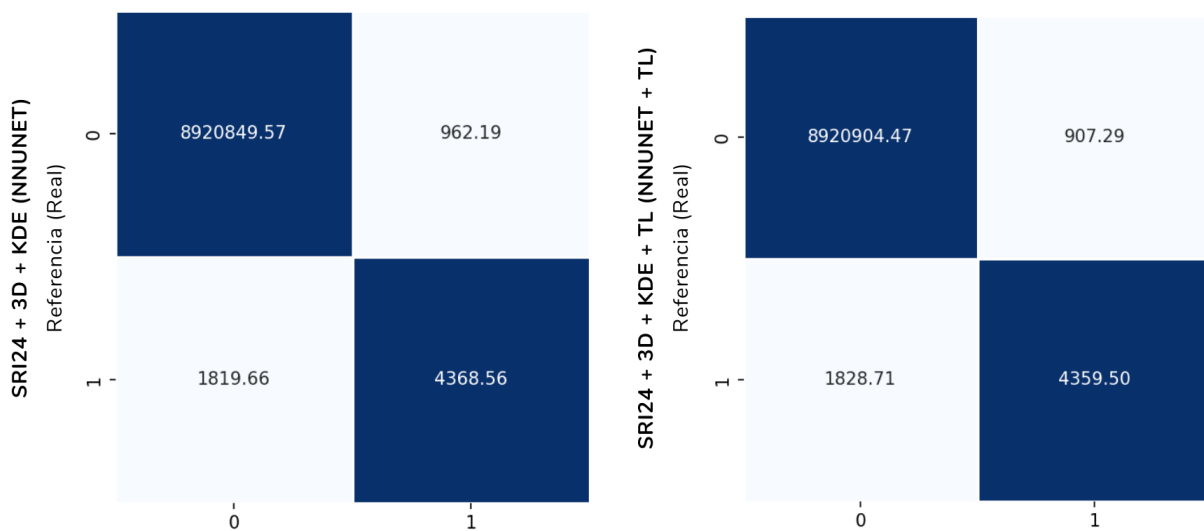


Figura 19: Matriz de confusión de la evaluación del Aprendizaje por Transferencia [Elaboración propia].

4.5 EVALUACIÓN DE LA ESTRATEGIA PROPUESTA VS MODELO DE REFERENCIA 3D-UNET

4.5.1 Entrenamiento del modelo 3D U-Net

3D U-Net fue utilizado como modelo para comparar el desempeño de nnU-Net debido a que ambos comparten una arquitectura estándar y permiten el procesamiento de datos en tres dimensiones. En este modelo, se usa la arquitectura UNet3D con 32 mapas de características iniciales. La optimización se llevó a cabo con un esquema similar al utilizado por nnU-Net, manteniendo el entrenamiento del modelo por el mismo lapso que se demoró el entrenamiento de nnU-Net. Se programó para un máximo de 125,000 iteraciones totales, con una validación

y registro cada 2,500 iteraciones. Se utilizó el optimizador Adam con una tasa de aprendizaje inicial de $1e^{-2}$ y un decaimiento de peso de $3e^{-5}$. La función de pérdida fue la combinación de Binary Cross-Entropy y DiceLoss. El tamaño del lote para entrenamiento se fijó en 2 y el tamaño del parche para la extracción de sub-volúmenes fue de [128,160,112]. La métrica de evaluación utilizada fue el Coeficiente Dice con un umbral de 0.5.

Observando los resultados de la matriz de confusión en la Figura 20 y el promedio de la métricas de evaluación, el modelo no mostró un buen desempeño en la segmentación e identificación de las lesiones. Solo obtuvo 662 VP, mientras que los valores de FN con 5,525 y FP 4,795 fueron considerablemente altos. Como consecuencia, su promedio de DSC fue de 0.0207 e IoU de 0.0119.

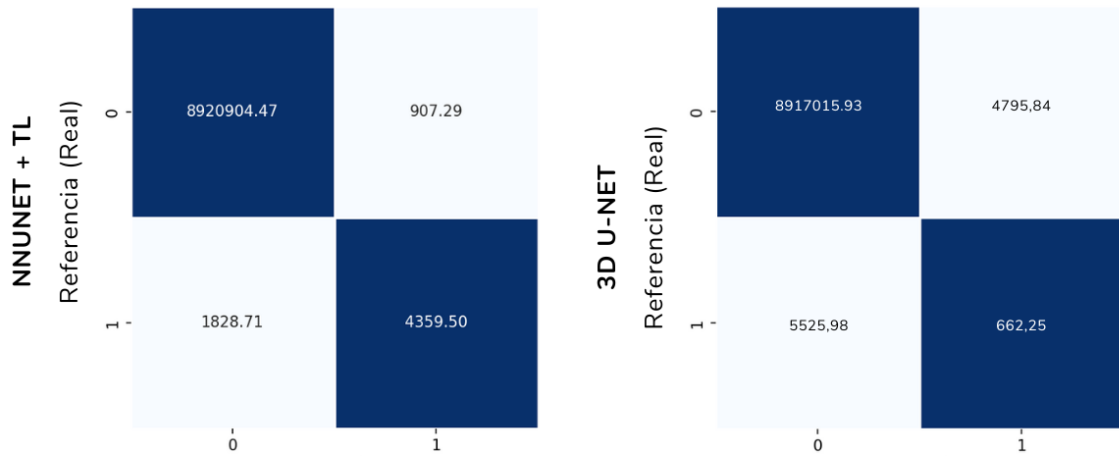


Figura 20: Matriz de confusión del modelo 3D U-Net [Elaboración propia].

En la Tabla 1 se compara el desempeño de los modelos: nnU-Net sin TL, nnU-Net con TL y 3D U-Net. Entre ellos, el modelo nn-UNet con TL obtuvo los mejores resultados, por lo que se considera el modelo final del trabajo.

Tabla 11: Comparación de resultados de evaluación de los modelos

MÉTRICA	NNU-NET SIN TL	NNU-NET + TL	3D U-NET
DSC	0.5322	0.5642	0.0207
IoU	0.4750	0.5097	0.0119

Finalmente, en la Figura 21 se muestran cuatro casos de sujetos con lesión, donde se compara la segmentación de referencia y las predicciones generadas por

cada uno de los modelos. En los dos primeros sujetos, el modelo nnU-Net generó una segmentación altamente alineada con la referencia. En el tercer y último sujeto, si bien logró identificar la región general de la lesión, no segmentó varios vóxeles que pertenecen a esta, en el tercer paciente la detección fue parcial pero aceptable, mientras que en el cuarto solo logró identificar unos pocos vóxeles. Por otro lado, el modelo 3D U-Net no mostró un buen desempeño en ninguno de los casos; además de no segmentar correctamente las lesiones, confundió estructuras anatómicas como los ventrículos, segmentandolas erróneamente como lesión.

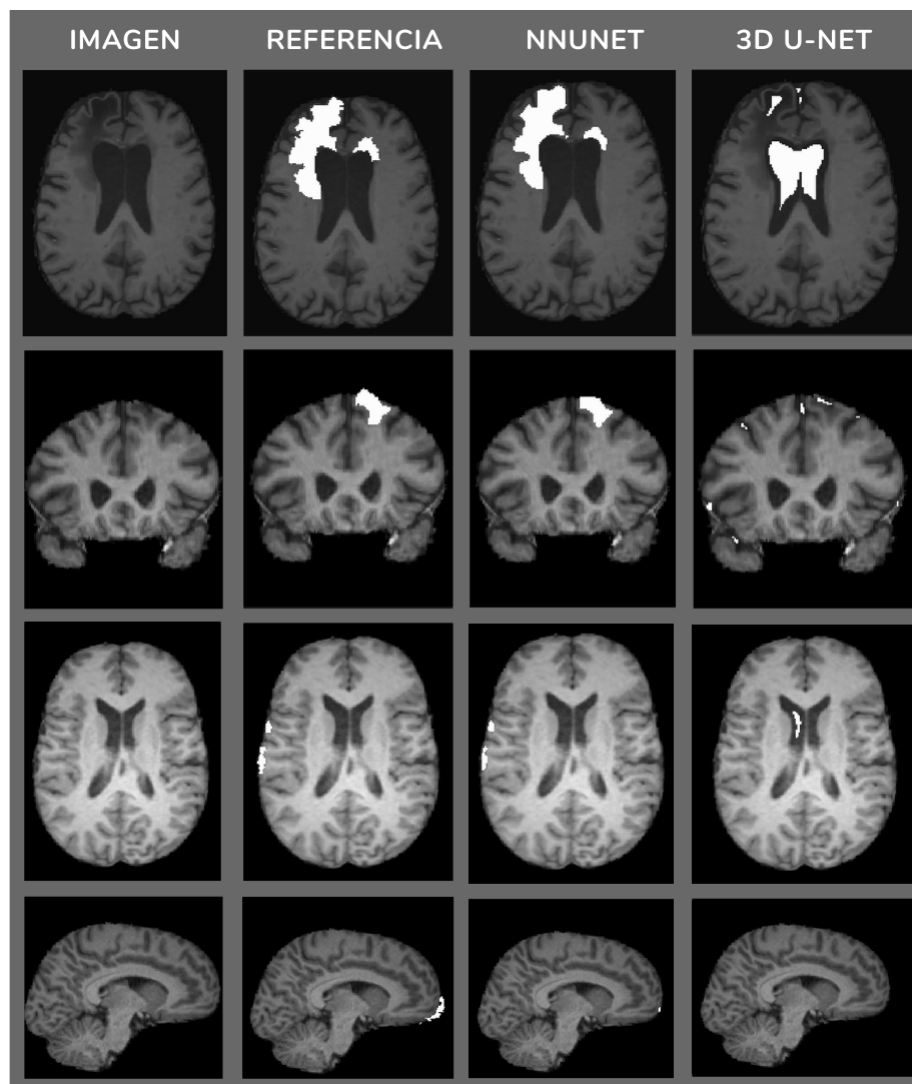


Figura 21: Segmentación de referencia vs predicción de los modelos [Elaboración propia].

4.6 ANÁLISIS DEL DESEMPEÑO DEL MODELO FINAL SEGÚN VARIABLES DEMOGRÁFICAS

Para el análisis del desempeño del modelo final se consideraron las variables edad y TSI. Es importante señalar que en este análisis se excluyeron los sujetos de control, debido a que se va a observar cómo varía la segmentación del modelo en función de estas dos variables clínicas.

Es importante mencionar que el TSI representa el tiempo transcurrido desde que ocurrió la lesión: valores bajos indican un trauma reciente, mientras que valores más altos pueden corresponder a lesiones en etapa subaguda o crónica. De este modo, se busca determinar si la edad del paciente y el estado temporal de la lesión influyen en la capacidad del modelo para segmentar adecuadamente.

4.6.1 *Análisis por rango de edad*

En la Figura 22 se presenta el número de sujetos del conjunto de pruebas agrupados en los siguientes cuatro rangos de edad: (18 – 31), (32 – 47), (48 – 63) y (64 – 78), y divididos por sexo. A partir de esta gráfica, se observa que el mayor número de sujetos corresponde al rango de (48 – 63) años en ambos sexos, mientras que el menor número se encuentra en el rango de (64 – 78) años.

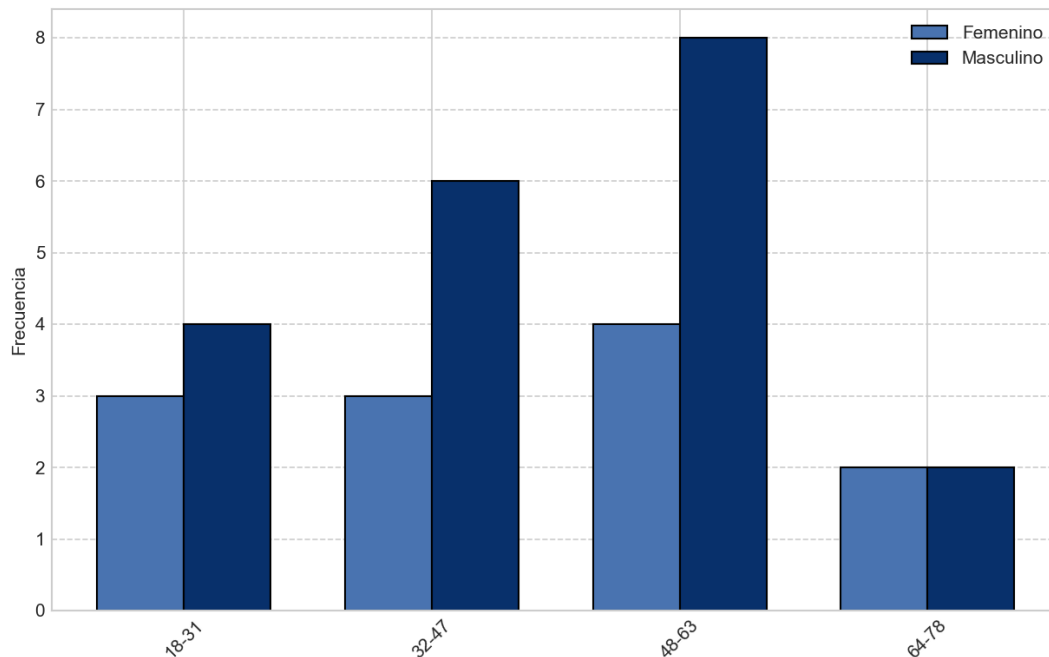


Figura 22: Número de sujetos por rango de edad y sexo [Elaboración propia].

La Figura 23 presenta la distribución media del DSC agrupada en los mismos cuatro rangos de edad, y dividida por sexo. Basandose en esto, se puede observar que el modelo alcanza su mejor desempeño para los hombres en el grupo de (48 – 63) años, y para las mujeres en el rango de (64 – 78) años. Por el contrario, el desempeño más bajo se observa en el rango de (32 – 47) años para los hombres, y de (18 – 31) para las mujeres. En general, lo que más se destaca es la marcada variabilidad de los resultados y la gran extensión de las barras de error en todos los grupos, lo cual sugiere una dispersión considerable en los valores de la métrica DSC. Esto indica que, aunque el modelo muestra un buen desempeño promedio, también existen casos individuales donde el modelo tiene un rendimiento muy bajo.

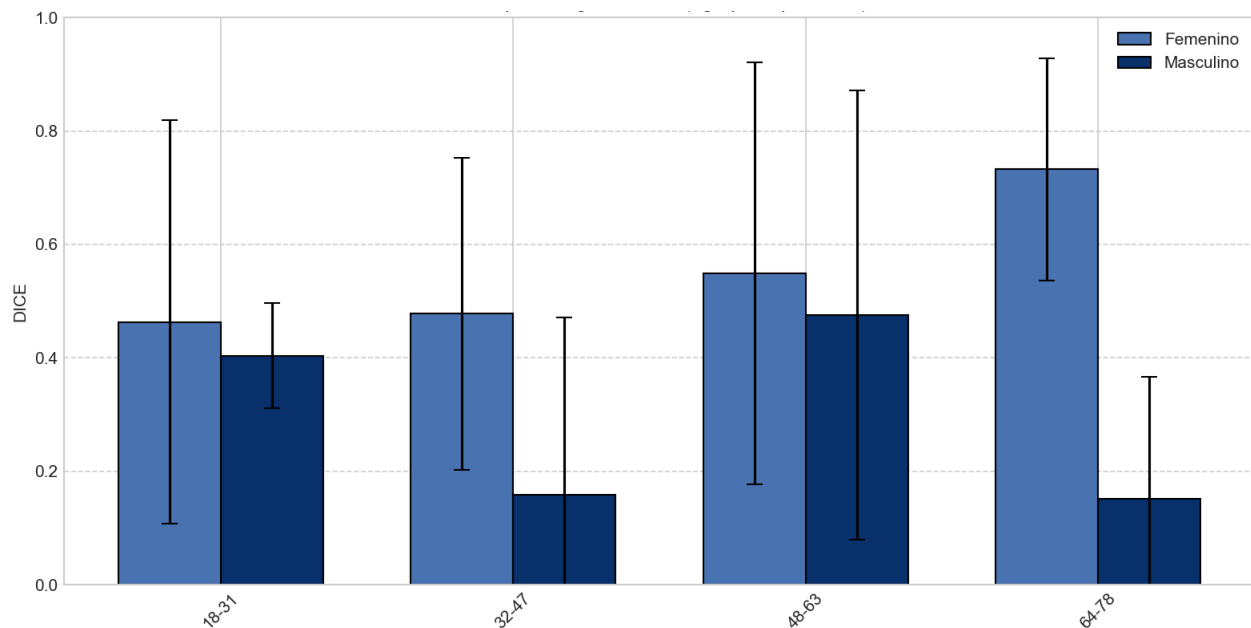


Figura 23: Promedio del DSC por rango de edad y sexo [Elaboración propia].

4.6.2 Análisis por TSI

En la Figura 24 se presenta el número de sujetos del conjunto de pruebas agrupados en los rangos de TSI calculado en semanas, comenzando con $TSI < 1$ semana y terminando con $TSI > 128$ semanas (alrededor de 2.5 años) y divididos por sexo. Por medio de esta gráfica, se observa que el mayor número de sujetos está entre las semanas (8 – 16) y (64 – 128). Además, no se tienen registros con TSI de (1 – 4) semanas.

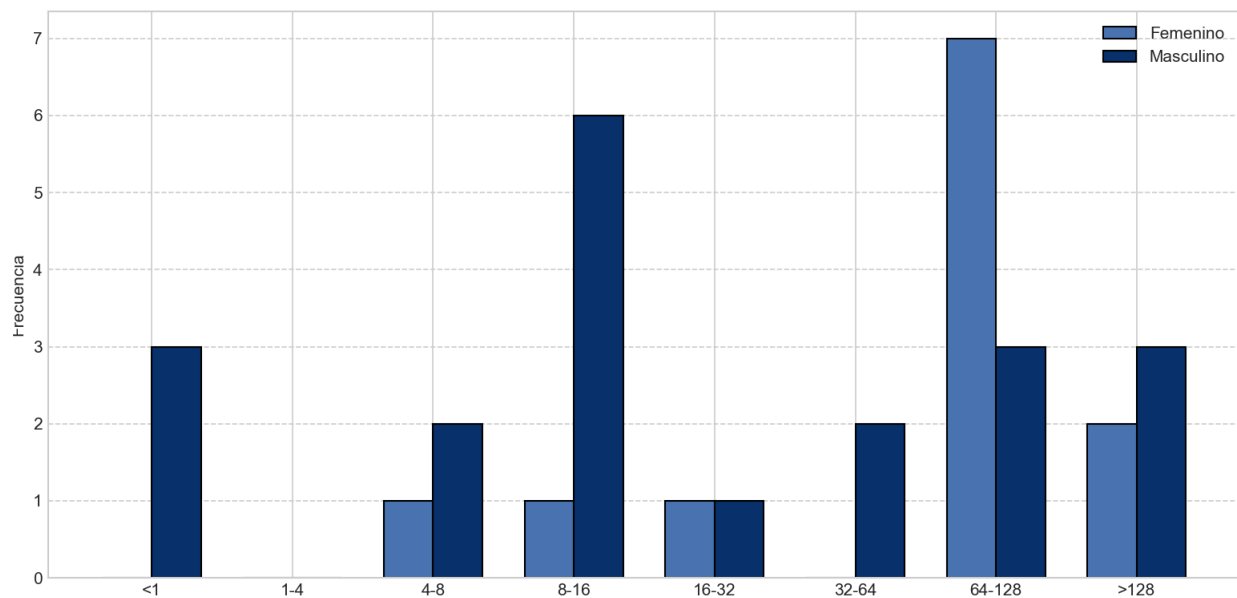


Figura 24: Número de sujetos por TSI y sexo [Elaboración propia].

En la Figura 25 se muestran los rangos de TSI calculado en semanas, y agrupados en los mismos rangos de la gráfica anterior. En esta, se puede observar que el modelo obtuvo el mejor desempeño de manera equitativa para ambos sexos en el rango de TSI (64 – 128) semanas, con un DSC de aproximadamente 0.62. Además, el desempeño promedio más bajo de TSI se encuentra en el rango (8 – 16) semanas, con un DSC cercano a 0.1. En general, la variabilidad es alta en la mayoría de los grupos. Sin embargo, la variabilidad se mantiene muy baja en el rango (8 – 16) semanas, lo que indica que el modelo produce resultados bajos de forma consistente en esta etapa. Esto podría estar asociado a la menor disponibilidad de datos en el caso de las mujeres para este rango de TSI. Por otro lado, aunque los hombres presentan un número relativamente mayor de casos, también se observa un rendimiento bajo en la mayoría de ellos, lo que sugiere que el modelo podría tener dificultades específicas para este intervalo temporal.

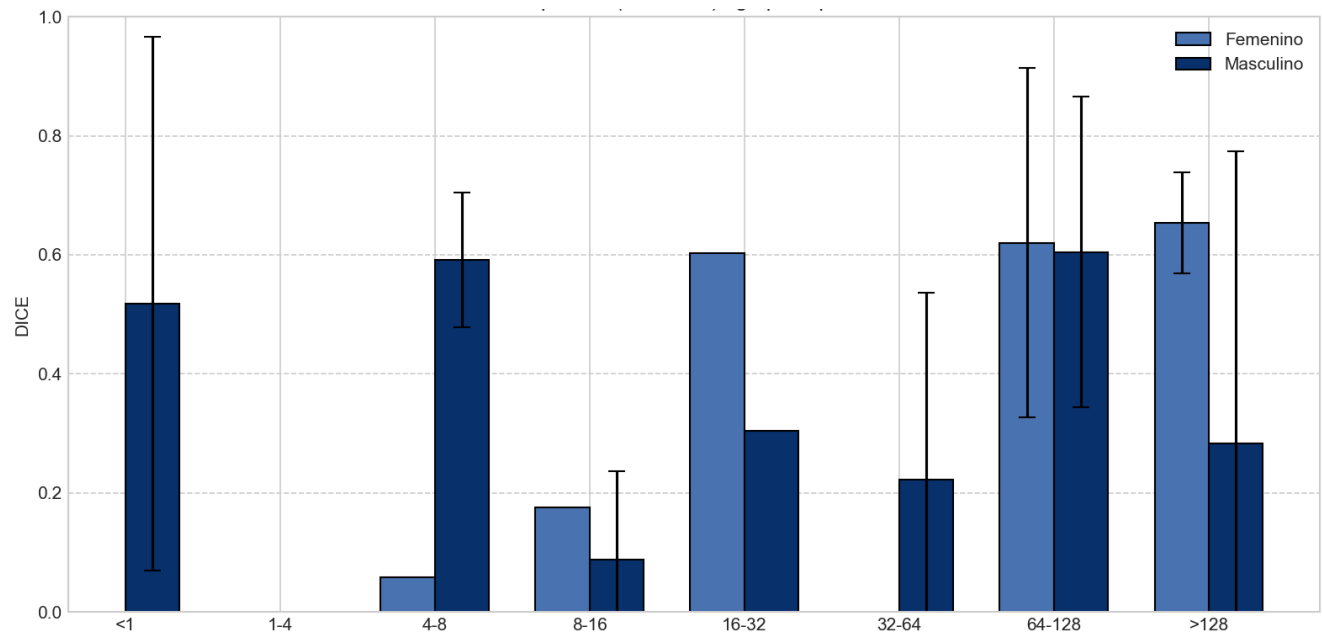


Figura 25: Promedio del DSC por TSI y sexo [Elaboración propia].

CONCLUSIONES Y TRABAJO FUTURO

En el presente trabajo de grado se ha explorado el uso de U-Net como un modelo para la segmentación e identificación de lesiones asociadas al TCEC a través de IRM ponderadas en T1, dando como resultado una estrategia para la segmentación automática de lesiones asociadas al TCEC.

En particular, se utilizaron modelos basados en la arquitectura U-Net, nnUNet y 3D U-Net. La estrategia propuesta fue el resultado de explorar el efecto de diferentes métodos de normalización durante el preprocesamiento de las imágenes y diferentes atlas para el registro. Se evidenció que el atlas SRI24 y el método KDE fue el que mejor manejó la normalización del conjunto de datos de TCEC. Aunque Z-score sigue siendo una opción útil, KDE logró mejorar las métricas de evaluación. Esto puede ocurrir debido que KDE modela la distribución de intensidades de forma no paramétrica, capturando mejor las variaciones propias de las IRM de pacientes con TCEC, las cuales suelen presentar distribuciones heterogéneas debido a la presencia de lesiones de distintos tipos, intensidades atípicas y artefactos. Por otro lado, Z-score asume que las intensidades tienen una distribución normal, por lo que puede perder precisión cuando los datos tienen variaciones.

Además, se evaluó el efecto del aprendizaje por transferencia (TL), demostrando que el desempeño del modelo mejora significativamente cuando parte de una base de conocimiento más amplia, como la obtenida con el conjunto de datos de tumores, BRATs. Esto permitió reducir la limitación del modelo debido al tamaño moderado del conjunto de TCEC.

En términos de desempeño, nnUNet se estableció como la mejor opción para la segmentación de lesiones en pacientes con un trauma tan complejo como el TCEC. De igual manera, el TL ayudó al modelo a detectar entre sujetos de control y sujetos con lesiones, además de refinar la segmentación.

Cabe resaltar, que todo el proceso de entrenamiento y evaluación fue realizado empleando el conjunto de datos AIMS-TBI. Sin embargo, sería útil adaptar y evaluar los modelos en más conjuntos de datos que incluyan mayor variabilidad anatómica y de adquisición. También, es importante mencionar que, aunque la diferencia no fue muy marcada, el conjunto de datos contenía más casos de cerebros con lesiones de hombres que de mujeres, lo que podría generar un sesgo en el proceso de aprendizaje.

Finalmente, el análisis de desempeño demostró gran variedad entre los grupos de edad, obteniendo un mejor desempeño del modelo en pacientes entre 48 y 63

años, y en aquellos con un TSI entre 64 y 128 semanas. Esto indica que tanto las características fisiológicas como la evolución en tiempo de las lesiones influyen en la calidad de la segmentación.

5.1 TRABAJOS FUTUROS

Como trabajo futuro se plantean tres líneas de acción:

- **Diversidad de conjuntos de datos:** Evaluar y adaptar el modelo a más conjuntos de datos que contengan mayor variedad de características, como imágenes obtenidas en diferentes plataformas, con distintos protocolos de adquisición y tipos de lesiones.
- **Incluir sujetos menores de edad:** Incluir datos de pacientes menores de edad para analizar cómo se comporta el modelo en cerebros en desarrollo, donde las estructuras pueden presentar diferencias significativas respecto a las de adultos.
- **Aplicación de postprocesamiento y validación cruzada:** Realizar la validación cruzada empleando las herramientas que ofrece el modelo para obtener una evaluación más sólida, además de investigar técnicas de postprocesamiento orientadas a mejorar específicamente la segmentación de lesiones asociadas al TCEC.

REFERENCIAS

- [1] J. Ginsburg y T. Smith, "Traumatic Brain Injury", *Encyclopedia of Child Behavior and Development*, págs. 374-374, feb. de 2025. DOI: [10.1007/978-0-387-79061-9_6270](https://doi.org/10.1007/978-0-387-79061-9_6270) dirección: <https://www.ncbi.nlm.nih.gov/books/NBK557861/>
- [2] O. Bernal-Pacheco, M. Vega-Rincón y J. F. Hernández-Preciado, *Consecuencias neuropsiquiátricas del trauma craneoencefálico Neuropsychiatric consequences of traumatic brain injury*, 2009.
- [3] B. Guan, D. B. Anderson, L. Chen, S. Feng y H. Zhou, "Global, regional and national burden of traumatic brain injury and spinal cord injury, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019", *BMJ Open*, vol. 13, 10 oct. de 2023, ISSN: 20446055. DOI: [10.1136/bmjopen-2023-075049](https://doi.org/10.1136/bmjopen-2023-075049)
- [4] A. Schuchat et al., *Morbidity and Mortality Weekly Report Traumatic Brain Injury-Related Emergency Department Visits, Hospitalizations, and Deaths-United States, 2007 and 2013 Surveillance Summaries Centers for Disease Control and Prevention MMWR Editorial and Production Staff (Serials) MMWR Editorial Board*, 2017.
- [5] J. D. Charry, J. F. Cáceres, A. C. Salazar, L. P. López y J. P. Solano, *Trauma craneoencefálico. Revisión de la literatura*, 2017.
- [6] infobae, "Cerca de 16.000 personas mueren al año por accidentes cerebro vasculares: estas son las ciudades con los mayores índices", 2023.
- [7] E. Dennis, N. Tustison, E. Deutscher, E. Wilde y K. Caeyenberghs, "AIMS TBI - Automated Identification of Moderate-Severe Traumatic Brain Injury Lesions. 27th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2024).", págs. 1-15, abr. de 2024. DOI: [10.5281/ZENODO.10990745](https://doi.org/10.5281/ZENODO.10990745) dirección: <https://zenodo.org/records/10990745>
- [8] M. Wilke, B. de Haan, H. Juenger y H. O. Karnath, "Manual, semi-automated, and automated delineation of chronic brain lesions: A comparison of methods", *NeuroImage*, vol. 56, págs. 1-9, 4 jun. de 2011, ISSN: 1053-8119. DOI: [10.1016/J.NEUROIMAGE.2011.04.014](https://doi.org/10.1016/J.NEUROIMAGE.2011.04.014) dirección: <https://www.sciencedirect.com/science/article/pii/S1053811911003922>

- [9] E. H. Ali, S. Sadek y Z. F. Makki, "Hybrid improved fuzzy C-means and watershed segmentation to classify Alzheimer's using deep learning", *IAES International Journal of Artificial Intelligence*, vol. 13, págs. 4080-4094, 4 dic. de 2024, ISSN: 22528938. DOI: [10.11591/ijai.v13.i4.pp4080-4094](https://doi.org/10.11591/ijai.v13.i4.pp4080-4094)
- [10] C. Bohyn, T. V. Vyvere, F. D. Keyzer, D. M. Sima y P. Demaerel, "Morphometric evaluation of traumatic axonal injury and the correlation with post-traumatic cerebral atrophy and functional outcome", *Neuroradiology Journal*, vol. 35, 4 2022, ISSN: 23851996. DOI: [10.1177/19714009211049714](https://doi.org/10.1177/19714009211049714)
- [11] C. Maffei et al., "Automated detection of axonal damage along white matter tracts in acute severe traumatic brain injury", *NeuroImage: Clinical*, vol. 37, 2023, ISSN: 22131582. DOI: [10.1016/j.nicl.2022.103294](https://doi.org/10.1016/j.nicl.2022.103294)
- [12] A. Bennett et al., "Manual lesion segmentations for traumatic brain injury characterization", *Frontiers in Neuroimaging*, vol. 2, 2023, ISSN: 28131193. DOI: [10.3389/FNIMG.2023.1068591/PDF](https://doi.org/10.3389/FNIMG.2023.1068591/PDF) dirección: <https://pubmed.ncbi.nlm.nih.gov/37554636/>
- [13] N. Gorgoraptis et al., "Cognitive impairment and health-related quality of life following traumatic brain injury", *NeuroRehabilitation*, vol. 44, 3 2019, ISSN: 18786448. DOI: [10.3233/NRE-182618](https://doi.org/10.3233/NRE-182618)
- [14] S. Palm, L. Rönnbäck y B. Johansson, "Long-term mental fatigue after traumatic brain injury and impact on employment status", *Journal of Rehabilitation Medicine*, vol. 49, págs. 228-233, 3 mar. de 2017, ISSN: 16501977. DOI: [10.2340/16501977-2190](https://doi.org/10.2340/16501977-2190)
- [15] J. T. van Dijck, M. D. Dijkman, R. H. Ophuis, G. C. de Ruiter, W. C. Peul y S. Polinder, *In-hospital costs after severe traumatic brain injury: A systematic review and quality assessment*, mayo de 2019. DOI: [10.1371/journal.pone.0216743](https://doi.org/10.1371/journal.pone.0216743)
- [16] S. I. Prada, Y. Takeuchi e Y. Ariza, *Costo monetario del tratamiento de la enfermedad de Alzheimer en Colombia Monetary cost of treatment for Alzheimer's disease in Colombia Artículo original*, 2014.
- [17] J. Ginsburg y T. Smith, "Traumatic Brain Injury", *Encyclopedia of Child Behavior and Development*, págs. 374-374, feb. de 2025. DOI: [10.1007/978-0-387-79061-9_6270](https://doi.org/10.1007/978-0-387-79061-9_6270) dirección: <https://www.ncbi.nlm.nih.gov/books/NBK557861/>
- [18] A. I. Maas, N. Stocchetti y R. Bullock, "Moderate and severe traumatic brain injury in adults", *The Lancet Neurology*, vol. 7, págs. 728-741, 8 ago. de 2008, ISSN: 1474-4422. DOI: [10.1016/S1474-4422\(08\)70164-9](https://doi.org/10.1016/S1474-4422(08)70164-9) dirección: <https://www.thelancet.com/action/showFullText?pii=S1474442208701649>

- [19] A. J. Tsiouris e Y. W. Lui, "Neuroimaging Update on Traumatic Brain Injury", págs. 87-112, feb. de 2024. DOI: [10.1007/978-3-031-50675-8_7](https://doi.org/10.1007/978-3-031-50675-8_7) dirección: <https://www.ncbi.nlm.nih.gov/books/NBK608606/>
- [20] I. R. R. Díaz et al., "Imágenes diagnósticas: conceptos y generalidades", *Revista de la Facultad de Ciencias Médicas*, págs. 35-42, 2014.
- [21] F. Gaillard, Y. Baba y D. Bell, "MRI sequences (overview)", *Radiopaedia.org*, jun. de 2015. DOI: [10.53347/RID-37346](https://doi.org/10.53347/RID-37346)
- [22] A. W. Toga y P. M. Thompson, "Image Registration and the Construction of Multidimensional Brain Atlases", *Handbook of Medical Image Processing and Analysis*, págs. 707-724, ene. de 2009. DOI: [10.1016/B978-012373904-9.50053-2](https://doi.org/10.1016/B978-012373904-9.50053-2) dirección: <https://www.sciencedirect.com/science/article/pii/B9780123739049500532>
- [23] D. A. Dickie et al., "Whole brain magnetic resonance image atlases: A systematic review of existing atlases and caveats for use in population imaging", *Frontiers in Neuroinformatics*, vol. 11, pág. 226 915, ene. de 2017, ISSN: 16625196. DOI: [10.3389/FNINF.2017.00001/FULL](https://doi.org/10.3389/FNINF.2017.00001/FULL) dirección: <http://www.brainmap.org/>;
- [24] B. Mahesh, "Machine learning algorithms-a review", *International Journal of Science and Research (IJSR)*. [Internet], vol. 9, n° 1, págs. 381-386, 2020.
- [25] I. El, N. R. Li y M. J. Murphy, *Theory and Applications Machine Learning in Radiation Oncology*, 2015.
- [26] M. Paluszczek y S. Thomas, "What Is Deep Learning?", *Practical MATLAB Deep Learning*, págs. 1-24, 2020. DOI: [10.1007/978-1-4842-5124-9_1](https://doi.org/10.1007/978-1-4842-5124-9_1) dirección: https://link.springer.com/chapter/10.1007/978-1-4842-5124-9_1
- [27] Y. LeCun, Y. Bengio y G. Hinton, "Deep learning", *nature*, vol. 521, n° 7553, págs. 436-444, 2015.
- [28] J. Kufel et al., "What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine", *Diagnostics 2023, Vol. 13, Page 2582*, vol. 13, pág. 2582, 15 ago. de 2023, ISSN: 2075-4418. DOI: [10.3390/DIAGNOSTICS13152582](https://doi.org/10.3390/DIAGNOSTICS13152582) dirección: <https://www.mdpi.com/2075-4418/13/15/2582>
- [29] A. H. Ali, M. G. Yaseen, M. Aljanabi y S. A. Abed, "Transfer Learning: A New Promising Techniques", *Mesopotamian Journal of Big Data*, vol. 2023, págs. 29-30, dic. de 2023, ISSN: 29586453. DOI: [10.58496/MJBD/2023/004](https://doi.org/10.58496/MJBD/2023/004)

- [30] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung y M. A. Azim, "Transfer learning: a friendly introduction", *Journal of Big Data*, vol. 9, págs. 1-19, 1 dic. de 2022, ISSN: 21961115. DOI: [10.1186/s40537-022-00652-w](https://doi.org/10.1186/s40537-022-00652-w) dirección: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00652-w>
- [31] M. J. Page et al., "Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas", *Revista española de cardiología*, vol. 74, nº 9, págs. 790-799, 2021.
- [32] W. K. Soh, H. Y. Yuen y J. C. Rajapakse, "HUT: Hybrid UNet transformer for brain lesion and tumour segmentation", *Heliyon*, vol. 9, 12 2023, ISSN: 24058440. DOI: [10.1016/j.heliyon.2023.e22412](https://doi.org/10.1016/j.heliyon.2023.e22412)
- [33] R. A. Hazarika, A. K. Maji, R. Syiem, S. N. Sur y D. Kandar, "Hippocampus Segmentation Using U-Net Convolutional Network from Brain Magnetic Resonance Imaging (MRI)", *Journal of Digital Imaging*, vol. 35, 4 2022, ISSN: 1618727X. DOI: [10.1007/s10278-022-00613-y](https://doi.org/10.1007/s10278-022-00613-y)
- [34] X. Zhao et al., "sTBI-GAN: An adversarial learning approach for data synthesis on traumatic brain segmentation", *Computerized Medical Imaging and Graphics*, vol. 112, 2024, ISSN: 18790771. DOI: [10.1016/j.compmedimag.2024.102325](https://doi.org/10.1016/j.compmedimag.2024.102325)
- [35] W. Zhu et al., "Automatic segmentation of white matter hyperintensities in routine clinical brain MRI by 2D VB-Net: A large-scale study", *Frontiers in Aging Neuroscience*, vol. 14, 2022, ISSN: 16634365. DOI: [10.3389/fnagi.2022.915009](https://doi.org/10.3389/fnagi.2022.915009)
- [36] M. Cabezas e Y. Diez, "An Analysis of Loss Functions for Heavily Imbalanced Lesion Segmentation", *Sensors*, vol. 24, nº 6, pág. 1981, 2024.
- [37] P. Coupeau, J.-B. Fasquel, L. Hertz-Pannier y M. Dinomais, "GNN-based structural information to improve DNN-based basal ganglia segmentation in children following early brain lesion", *Computerized Medical Imaging and Graphics*, vol. 115, pág. 102396, 2024.
- [38] M. F. Rachmadi, M. Byra y H. Skibbe, "A new family of instance-level loss functions for improving instance-level segmentation and detection of white matter hyperintensities in routine clinical brain MRI", *Computers in Biology and Medicine*, vol. 174, pág. 108414, mayo de 2024, ISSN: 0010-4825. DOI: [10.1016/J.COMPBIOMED.2024.108414](https://doi.org/10.1016/J.COMPBIOMED.2024.108414)
- [39] K. Pierre et al., "Machine Learning in Neuroimaging of Traumatic Brain Injury: Current Landscape, Research Gaps, and Future Directions", *Trauma Care*, vol. 4, págs. 31-43, 1 ene. de 2024. DOI: [10.3390/traumacare4010004](https://doi.org/10.3390/traumacare4010004)

- [40] S. Venkatasubramanian, J. N. Dwivedi, S. Raja, N. Rajeswari, J. Logeshwaran y A. P. Kumar, "Prediction of Alzheimer's Disease Using DHO-Based Pretrained CNN Model", *Mathematical Problems in Engineering*, vol. 2023, 2023, ISSN: 15635147. DOI: [10.1155/2023/1110500](https://doi.org/10.1155/2023/1110500)
- [41] S. Prabha, K. S. Sankaran y D. Chitradevi, "Efficient optimization based thresholding technique for analysis of alzheimer MRIs", *International Journal of Neuroscience*, vol. 133, 2 2023, ISSN: 15635279. DOI: [10.1080/00207454.2021.1901696](https://doi.org/10.1080/00207454.2021.1901696)
- [42] B. F. Stanley y S. W. Franklin, "Effective feature extraction for Cerebral Microbleed detection using Edge Emphasized Weber Maximum Directional Co-occurrence Matrix", *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, págs. 13683-13696, 10 oct. de 2023, ISSN: 18685145. DOI: [10.1007/s12652-022-04023-4](https://doi.org/10.1007/s12652-022-04023-4)
- [43] T. Mistral et al., "Automated Quantification of Brain Lesion Volume From Post-trauma MR Diffusion-Weighted Images", *Frontiers in Neurology*, vol. 12, 2022, ISSN: 16642295. DOI: [10.3389/fneur.2021.740603](https://doi.org/10.3389/fneur.2021.740603)
- [44] S. Holmes et al., "Integrated Features for Optimizing Machine Learning Classifiers of Pediatric and Young Adults With a Post-Traumatic Headache From Healthy Controls", *Frontiers in Pain Research*, vol. 3, 2022, ISSN: 2673561X. DOI: [10.3389/fpain.2022.859881](https://doi.org/10.3389/fpain.2022.859881)
- [45] G. M. D'Souza et al., "Interaction of Alzheimer Disease and Traumatic Brain Injury on Cortical Thickness", *Alzheimer Disease and Associated Disorders*, vol. 38, 1 2024, ISSN: 08930341. DOI: [10.1097/WAD.0000000000000607](https://doi.org/10.1097/WAD.0000000000000607)
- [46] X. Zhang et al., "Brain age prediction using interpretable multi-feature-based convolutional neural network in mild traumatic brain injury", *NeuroImage*, vol. 297, ago. de 2024, ISSN: 10959572. DOI: [10.1016/j.neuroimage.2024.120751](https://doi.org/10.1016/j.neuroimage.2024.120751)
- [47] L. Berger, B. Holshouser, J. G. Nichols, J. Pivonka-Jones, S. Ashwal y B. Bartnik-Olson, "White Matter Metabolite Ratios Predict Cognitive Outcome in Pediatric Traumatic Brain Injury", *Metabolites*, vol. 13, 7 2023, ISSN: 22181989. DOI: [10.3390/metabo13070778](https://doi.org/10.3390/metabo13070778)
- [48] A. M. Azor, D. J. Sharp, A. E. Jolly, N. J. Bourke y P. J. Hellyer, "Automation and standardization of subject-specific region-of-interest segmentation for investigation of diffusion imaging in clinical populations", *PLoS ONE*, vol. 17, 12 December 2022, ISSN: 19326203. DOI: [10.1371/journal.pone.0268233](https://doi.org/10.1371/journal.pone.0268233)

- [49] J. Heholt et al., "Simultaneous Arterial Spin Labelling Functional MRI and Fluorodeoxyglucose PET in Mild Chronic Traumatic Brain Injury", *Journal of Neuroradiology*, pág. 101 211, 2024.
- [50] J. J. Muller et al., "Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging", *Frontiers in Neuroscience*, vol. 17, pág. 1 182 509, ago. de 2023, ISSN: 1662453X. DOI: [10.3389/FNINS.2023.1182509](https://doi.org/10.3389/FNINS.2023.1182509) / BIBTEX dirección: https://www.nitrc.org/projects/noddi_
- [51] N. Atia et al., "Particle Swarm Optimization and Two-Way Fixed-Effects Analysis of Variance for Efficient Brain Tumor Segmentation", *Cancers*, vol. 14, 18 2022, ISSN: 20726694. DOI: [10.3390/cancers14184399](https://doi.org/10.3390/cancers14184399)
- [52] P. M. Thompson et al., "The Enhancing NeuroImaging Genetics through Meta-Analysis Consortium: 10 Years of Global Collaborations in Human Brain Mapping", *Human Brain Mapping*, vol. 43, pág. 15, 1 ene. de 2021, ISSN: 10970193. DOI: [10.1002/HBM.25672](https://doi.org/10.1002/HBM.25672) dirección: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8675422/>
- [53] J. E. Iglesias, C. Y. Liu, P. M. Thompson y Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods", *IEEE Transactions on Medical Imaging*, vol. 30, págs. 1617-1634, 9 sep. de 2011, ISSN: 02780062. DOI: [10.1109/TMI.2011.2138152](https://doi.org/10.1109/TMI.2011.2138152) dirección: <https://ieeexplore-ieee-org.bd.univalle.edu.co/document/5742706>
- [54] B. B. Avants, C. L. Epstein, M. Grossman y J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain", *Medical Image Analysis*, vol. 12, págs. 26-41, 1 feb. de 2008, ISSN: 1361-8415. DOI: [10.1016/J.MEDIA.2007.06.004](https://doi.org/10.1016/J.MEDIA.2007.06.004) dirección: <https://www.sciencedirect.com/science/article/pii/S1361841507000606>
- [55] Y. Ou, H. Akbari, M. Bilello, X. Da y C. Davatzikos, "Comparative Evaluation of Registration Algorithms in Different Brain Databases With Varying Difficulty: Results and Insights", *IEEE transactions on medical imaging*, vol. 33, pág. 2039, 10 oct. de 2014, ISSN: 1558254X. DOI: [10.1109/TMI.2014.2330355](https://doi.org/10.1109/TMI.2014.2330355) dirección: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4371548/>
- [56] M. T. J. B. A. H. D. Toro, "Creation and characterisation of a colombian brain atlas based on magnetic resonance imaging: A study of southwestern colombian population", Tesis doct., 2023, pág. 29.

- [57] Y. Li, S. Ammari, C. Balleyguier, N. Lassau y E. Chouzenoux, "Impact of Preprocessing and Harmonization Methods on the Removal of Scanner Effects in Brain MRI Radiomic Features", *Cancers*, vol. 13, pág. 3000, 12 jun. de 2021, ISSN: 20726694. DOI: [10.3390/CANCERS13123000](https://doi.org/10.3390/CANCERS13123000) dirección: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8232807/>
- [58] E. E. Carvajal-Camelo, J. Bernal, A. Oliver, X. Lladó y M. Trujillo, "Evaluating the Effect of Intensity Standardisation on Longitudinal Whole Brain Atrophy Quantification in Brain Magnetic Resonance Imaging", *Applied Sciences* 2021, Vol. 11, Page 1773, vol. 11, pág. 1773, 4 feb. de 2021, ISSN: 2076-3417. DOI: [10.3390/APP11041773](https://doi.org/10.3390/APP11041773) dirección: [https://www.mdpi.com/2076-3417/11/4/1773](https://www.mdpi.com/2076-3417/11/4/1773/htm%20https://www.mdpi.com/2076-3417/11/4/1773)
- [59] J. C. Reinhold, B. E. Dewey, A. Carass y J. L. Prince, "Evaluating the Impact of Intensity Normalization on MR Image Synthesis", *Proceedings of SPIE—the International Society for Optical Engineering*, vol. 10949, 109493H, mar. de 2019, ISSN: 16057422. DOI: [10.1117/12.2513089](https://doi.org/10.1117/12.2513089) dirección: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6758567/>
- [60] M. Ligeró et al., "Minimizing acquisition-related radiomics variability by image resampling and batch effect correction to allow for large-scale data analysis", *European Radiology* 2020 31:3, vol. 31, págs. 1460-1470, 3 sep. de 2020, ISSN: 1432-1084. DOI: [10.1007/S00330-020-07174-0](https://doi.org/10.1007/S00330-020-07174-0) dirección: <https://link.springer.com/article/10.1007/s00330-020-07174-0>
- [61] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen y K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation", *Nature Methods* 2020 18:2, vol. 18, págs. 203-211, 2 dic. de 2020, ISSN: 1548-7105. DOI: [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z) dirección: <https://www.nature.com/articles/s41592-020-01008-z>
- [62] A. Wolny et al., "Accurate and versatile 3D segmentation of plant tissues at cellular resolution", *eLife*, vol. 9, págs. 1-34, jul. de 2020, ISSN: 2050084X. DOI: [10.7554/ELIFE.57613](https://doi.org/10.7554/ELIFE.57613)