



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

INSTITUTO DE CIENCIAS APLICADAS Y TECNOLOGÍA
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS

SEGMENTACIÓN DE IMÁGENES MÉDICAS PARA EL RECONOCIMIENTO Y RECONSTRUCCIÓN DE
GLIOBLASTOMAS IMPLEMENTANDO REDES NEURONALES PROFUNDAS

TESIS
QUE PARA OPTAR POR EL GRADO DE
MAESTRO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

PRESENTA:
MIGUEL ÁNGEL VELOZ LUCAS

TUTOR
DR. MIGUEL ÁNGEL PADILLA CASTAÑEDA
ICAT UNAM

CIUDAD UNIVERSITARIA, CD. MX., NOVIEMBRE 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**PROPUESTA UNIVERSITARIA DE INTEGRIDAD Y
HONESTIDAD ACADÉMICA Y PROFESIONAL
(Graduación con trabajo escrito)**

De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción 1, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la Institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado SEGMENTACIÓN DE IMÁGENES MÉDICAS PARA EL RECONOCIMIENTO Y RECONSTRUCCIÓN DE GLIOBLASTOMAS IMPLEMENTANDO REDES NEURONALES PROFUNDAS que presenté para obtener el grado de Maestría es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi programa de posgrado, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de graduación.

Atentamente



Miguel Ángel Veloz Lucas, 304183371
(Nombre, firma y Número de cuenta de la persona alumna)

A la motivación y la curiosidad.

Agradecimientos

Institucionales

Agradecemos el financiamiento recibido por el grupo de Bioinstrumentación del ICAT, UNAM para el Laboratorio de Cirugía, por la Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, México, al proyecto DGAPA-PAPIIT-TA101422 “Esquema de neuronavegación quirúrgica por imágenes multiespectrales y simulación biomecánica para resección de tumores cerebrales guiada por computadora”; Secretaría de Educación, Ciencia, Tecnología e Innovación de la Ciudad de México, al proyecto SECTEI 219/2019 “Laboratorio de Investigación y Desarrollo de Simuladores Clínicos y Cirugía Asistida por Computadora”; Consejo Nacional de Humanidades, Ciencia y Tecnología, México, para el proyecto CONAHCyT PCC 319585 “Metodología de realidad aumentada multiespectral para la identificación y localización de tejido patológico en intervenciones quirúrgicas guiadas por computadora”.

Personales

Al Posgrado en Ciencia e Ingeniería de la Computación y al personal administrativo, por el seguimiento y atenciones brindadas durante todo el proceso de elaboración de este proyecto.

A mi director de tesis el Dr. Miguel Ángel Padilla Castañeda, por su enseñanza, consejo y dedicación que sirvieron como guía en la realización de este proyecto.

Al equipo de trabajo del Laboratorio de Bioinstrumentación del ICAT por su compañerismo y solidaridad.

A mis amigos Daniel, Roberto y Dalia, por su pasión de saber más y su invaluable amistad.

A mi madre Irma Lucas, mi padre Jaime Veloz y mi hermano Christian J. Veloz Lucas, por su amor y su apoyo en todos estos años y los que vendrán.

Resumen

El presente trabajo tiene como objetivo la exploración y aplicación de modelos de redes neuronales profundas para realizar tareas de segmentación de imágenes y hacer frente a los desafíos de la interpretación de imágenes médicas. Se propone la implementación de un sistema automatizado para realizar la tarea de segmentación de glioblastomas en imágenes de resonancia magnética de contraste, implementando modelos de aprendizaje profundo que se sitúan en el estado del arte del campo de las redes neuronales y la inteligencia artificial. A partir de la segmentación automática realizada por el modelo, se genera la reconstrucción en tres dimensiones del tumor para referenciar las posibles áreas de alta funcionalidad con afectación por la presencia de tejido tumoral, todo esto con la finalidad de mostrar y proponer el desarrollo de nuevas herramientas especializadas con la implementación de nuevas tecnologías de aprendizaje profundo, para detectar, diagnosticar y planear estrategias en el tratamiento de este tipo de neoplasia.

Índice general

Agradecimientos	III
Resumen	v
Lista de figuras	x
Lista de tablas	xvi
1 Introducción	1
1.1 Aspectos generales	1
1.2 Estructura del trabajo	5
2 Caso de estudio: Glioblastoma	7
2.1 Astrocitomas y Glioblastoma multiforme	7
2.1.1 Incidencia	9
2.2 Diagnóstico	12
2.2.1 Evaluación clínica del paciente	12
2.2.2 Diagnóstico con imagenología	14
2.3 Tratamiento	16
2.3.1 Manejo preoperatorio	17
2.3.2 Cirugía	17

2.3.3	Técnicas intraoperatorias, mayor seguridad en el tratamiento	18
2.3.4	Biopsia estereotáxica	19
2.3.5	Radioterapia	20
2.3.6	Quimioterapia y Terapia dirigida	22
2.4	Seguimiento del tratamiento	22
3	Imagen de resonancia magnética	24
3.1	Aspectos generales	24
3.2	Propiedades microscópicas de la IRM	25
3.3	Tipos de secuencia RM	31
3.3.1	Secuencia eco de espín SE	32
3.3.2	Secuencia T1	33
3.3.3	Secuencia T2	33
3.3.4	Secuencia FLAIR	34
3.3.5	Medios de contraste	35
4	Estado del arte de la segmentación de imágenes médicas	36
4.1	Segmentación y análisis de imágenes médicas	36
4.2	Aprendizaje profundo en análisis de imágenes médicas	37
4.3	Algoritmos de aprendizaje profundo para segmentación de imágenes médicas	41
4.3.1	Herramientas de software para segmentación automática de imágenes médicas	42

<i>ÍNDICE GENERAL</i>	VIII
5 Base de datos	
y arquitectura Swin UNETR	47
5.1 Repositorios de imágenes médicas	
cerebrales de libre acceso	47
5.1.1 Conjunto de datos BraTS 2023	49
5.1.2 Descripción de datos de entrenamiento	50
5.1.3 Segmentación de tumor	51
5.1.4 Formato de imágenes	52
5.2 Arquitectura del modelo	
Swin UNEt TRransformers (Swin UNETR)	54
5.2.1 Transformador de Visión ViT	55
5.2.2 Swin UNETR	59
6 Metodología	61
6.1 Herramientas de software	61
6.1.1 Bibliotecas de Python	61
6.1.2 Visualization Toolkit (VTK)	62
6.1.3 Insight Toolkit (ITK)	63
6.2 Preprocesamiento de imagen	63
6.3 Implementación y entrenamiento del modelo Swin UNETR . .	65
6.4 Validación de predicciones	68
6.4.1 Coeficiente DICE	68
6.4.2 Distancia de Hausdorff	70
7 Resultados, reconstrucción y análisis	72
7.1 Resultados	72

<i>ÍNDICE GENERAL</i>	IX
7.1.1 Introducción	72
7.1.2 Resultados de entrenamiento	73
7.1.3 Resultados de prueba	74
7.2 Reconstrucción	81
8 Discusión y Conclusiones	86
8.1 Discusiones	86
8.2 Conclusiones y trabajo a futuro	88

Índice de figuras

1.1	La imagen muestra diferentes tipos de células gliales las cuales dan soporte a las neuronas y conforman las vainas de mielina de los axones [3] (Imagen: “Illustration Of Neurons And Neuroglial Cells - Neuroglia@seekpng.com”).	2
2.1	A) IRM de encéfalo con contraste preoperatorio, donde se evidencia el tumor sólido con bordes irregulares y el componente quístico frontal izquierdo que capta el medio de contraste . Desviación de la línea media cerebral y compresión del ventrículo lateral izquierdo con efecto de masa: A) imagen axial. B) Imagen coronal. C) RM encéfalo T1 con contraste posoperatorio luego de un mes que demuestra los cambios luego de la cirugía, no evidencia realce de contraste. D) RM encéfalo T1 con contraste posoperatorio, control a los 3 años de la cirugía, no evidencia realce de contraste [16] (Imagen: “Revista Peruana de Medicina Experimental y Salud Publica [16]”).	10
2.2	Anatomía del interior del encéfalo. La región supratentorial contiene el cerebro, los ventrículos, el plexo coroideo, el hipotálamo, la glándula pineal, la hipófisis y el nervio óptico. La región infratentorial contiene el cerebelo y el tronco encefálico [21] (Imagen: “Tratamiento del ependimoma infantil PDQ®–Versión para profesionales de salud [21]”).	13

2.3	Anotación manual a través de evaluadores expertos. Se muestran parches de imagen con las estructuras tumorales que se anotan en las diferentes modalidades (arriba a la izquierda) y las etiquetas finales para todo el conjunto de datos (derecha). Los parches de imagen muestran de izquierda a derecha: todo el tumor visible en FLAIR, (A) el núcleo del tumor visible en T2 (B), las estructuras tumorales realizadas visibles en t1c (azul), que rodean los componentes quísticos/necróticos del núcleo (verde) (C). Las segmentaciones se combinan para generar las etiquetas finales de las estructuras tumorales (D): edema (amarillo), núcleo sólido sin realce (rojo), núcleo necrótico/quístico (verde), núcleo realzado (azul) [27] (Imagen: “The Multimodal Brain Tumor Image Segmentation Benchmark [27]”).	16
2.4	Médula cervical con luz violeta-azul. Se aprecia la fluorescencia rosada por transiluminación [32] (Imagen: “Resección de tumor intramedular guiada por fluorescencia con ácido aminolevulínico [32]”).	19
2.5	Realización de biopsia de alta precisión con posicionamiento de marco estereotáctico en lesiones del tallo cerebral (A, B) utilizando software de alta precisión para la planeación de trayectorias estereotácticas con fusión de imágenes (C, D) [34](Imagen: “Revista Gastrohup [34]”).	21
3.1	Propiedades microscópicas responsables de la MRI. a) Tejidos en un voxel b) Células en los tejidos c) Moléculas de agua en los tejidos [40] (Imagen: “MR pulse sequences: what every radiologist wants to know but is afraid to ask”, 2006 [40]).	26
3.2	a) Precesión de un protón. La imagen corresponde al movimiento del eje de rotación sobre las curvas discontinuas. b) La alineación del espín con el campo magnético nunca llega a realizarse del todo, en virtud del origen cuántico del espín; lo más que puede lograrse es una alineación en el sentido de la dirección del campo magnético que llamamos “paralela”, o una alineación en sentido contrario que llamamos “antiparalela” [38] (Imagen: “MRI, basic principles and applications [38]”).	27

3.3 En la figura se muestra el efecto de los “pulso de 90 grados”, en *a)* se puede observar el efecto de un pulso de radiofrecuencia de 90 grados. En *b)* los protones entran en sincronismo y comienzan a precesar en fase. En *c)* sus vectores se suman en dirección transversal al campo magnético y se establece una magnetización transversal o un vector suma que representa el momento magnético total de un tejido [43] (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación, Universitas Médica” [43]). 30

3.4 En la figura *a)* observamos que los protones regresan a su nivel de menor energía, disminuye la magnetización transversal T2 y se recupera la magnetización longitudinal T1. En *b)* se representa el proceso que genera la emisión de la señal de resonancia o eco que va a ser captado por una antena [43] (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación, Universitas Médica” [43]). 30

3.5 La figura muestra la secuencia SE de forma esquemática (Imagen: “Introducción biofísica a la resonancia magnética”. Centre Diagnòstic Pedralbes, 1993 [48]). 32

3.6 La figura muestra una tabla donde se puede observar las diferencias en las intensidades de niveles de gris entre las secuencias T1 y T2 (Imagen: “Introducción biofísica a la resonancia magnética”. Centre Diagnòstic Pedralbes, 1993 [48]). 34

3.7 La figura compara la imagen del mismo caso de RM entre cada una de las secuencias *a)* T1, *b)* T2 y *c)* FLAIR (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación”. Universitas Médica, 2011 [43]). 35

4.1 Descripción general y detalles de una arquitectura CNN para el reconocimiento de imágenes. *a)* Expande una entrada RGB de tres canales, convolución + función de activación. *b)* A partir de una imagen de entrada, se crean los mapas de características por convolución y se redimensionan mediante operaciones de agrupación máxima que reducen la resolución pero los mapas de características son más profundos hasta que llegan a un clasificador. *c)* Clasificación realizada con los pesos entrenados de la capa convolucional así como de la capa completamente conectada [53] (Imagen: “Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends”[53]). 40

5.1 Estudio de un paciente con diagnóstico confirmado de GBM, donde se presentan las diferentes secuencias capturadas de RM y la segmentación del área donde se encuentra el tumor. Cada fila representa un plano anatómico distinto (axial, sagital y coronal, de arriba hacia abajo respectivamente). Las columnas muestran cada una de las secuencias capturadas (T1, T1ce, T2, Flair) adicionando en la última columna la segmentación de las áreas donde está presente el tejido tumoral. Conjunto de datos procedentes de BraTS 2023 [68]. 52

5.2 Segmentación manual de una imagen de RM realizada por los expertos en anotación. Los paneles de imágenes *A – C* indican las regiones consideradas de la segmentación: *A*) el tumor realzado (ET - amarillo), *B*) el núcleo del tumor (TC - magenta) y *C*) todo el completo (WT - cian). *D*) muestra las segmentaciones combinadas que generan las etiquetas finales de las subregiones tumorales, núcleo mejorado (amarillo), núcleo necrótico/quístico (rojo) y edema/invasión (verde) (Imagen: The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification [68]). . . . 53

5.3 El Transformador sigue esta arquitectura general utilizando autoatención apilada y capas puntuales y completamente conectadas tanto para el codificador como para el decodificador, como se muestra en las mitades izquierda y derecha de la figura, respectivamente [71] (Imagen: Advances in neural information processing systems [71]). 55

5.4 A) Swin Transformer es un tipo de ViT que crea mapas de características jerárquicas fusionando parches de imágenes en capas más profundas y tiene una complejidad de cálculo lineal debido al cálculo de la autoatención solo dentro de cada ventana local. B) Los modelos ViT producen mapas de características de una única resolución baja y tienen una complejidad de cálculo cuadrático (Imagen: Swin transformer: Hierarchical vision transformer using shifted windows [73]). 58

5.5 El modelo UNETR consiste de un transformador codificador que utiliza directamente parches 3D y está conectado a un decodificador basado en CNN mediante una conexión de salto (Imagen: Unetr: Transformers for 3d medical image segmentation [74]). 59

5.6	Descripción general de la arquitectura Swin UNETR. La entrada del modelo son IRM multimodal en 3D con 4 canales. El Swin UNETR crea parches sin traslape de los datos de entrada y utiliza una capa de partición de parches para crear ventanas con el tamaño deseado para calcular la autoatención. Las representaciones de características codificadas en Swin Transformer se alimentan a un decodificador CNN a través de una conexión de salto en múltiples resoluciones. La salida de segmentación final consta de 3 canales de salida correspondientes a las subregiones tumor completo (WT), tumor realzado (ET) y núcleo tumoral (TC) [70](Imagen:Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images [70]). . . .	60
6.1	La imagen muestra un corte del estudio de RM de la secuencia FLAIR y la segmentación de las subestructuras que conforman el tumor. <i>A)</i> Pertene a un corte de la secuencia FLAIR de un caso confirmado de GBM, <i>B)</i> muestra el área completa del tumor que contiene las tres subestructuras que lo componen, <i>C)</i> es el área segmentada que presenta edema peritumoral, <i>D)</i> área segmentada donde se detecta el tumor con realce y <i>E)</i> es la segmentación del núcleo del tumor. Conjunto de datos procedentes de BraTS'23 [68].	66
6.2	El índice de Sørensen-Dice mide la similitud y diversidad de conjuntos de muestras, porcentaje de superposición entre dos conjuntos, que es un valor entre 0 y 1 (Imagen: “9 Distance Measures in Data Science: The advantages and pitfalls of common distance measures” [79]).	69
7.1	Gráficas que muestran el descenso de la pérdida y el incremento de la exactitud a través de las épocas de entrenamiento del modelo.	74
7.2	Gráficos que muestran el incremento de la exactitud promedio para cada una de las clases de salida que representan las subestructuras del tumor (WT, TC, ET)a lo largo de las épocas de entrenamiento del modelo.	74
7.3	Predicciones realizadas por el modelo, donde cada fila muestra un plano anatómico distinto de un mismo caso del conjunto de prueba. En la primera columna un corte de RM de la secuencia FLAIR, la segunda columna contiene la segmentación realizada por el especialista y por último en la tercera columna se muestra la predicción obtenida del modelo entrenado.	75

7.4	Predicciones realizadas por el modelo de los primeros 5 casos de estudio de RM, donde cada fila muestra un caso distinto del conjunto de evaluación del desafío <i>BraTS'23</i> , cada columna muestra una secuencia de captura diferente (T1,T1ce,T2,FLAIR). Las imágenes presentan la segmentación automática de GBM en casos de RM, resaltando con distinto color las subestructuras que componen al tumor: amarillo para Edema, cafe para el tumor realzado y azul para el núcleo del tumor.	79
7.5	Predicciones realizadas por el modelo de 5 casos de estudio de RM, donde cada fila muestra un caso distinto del conjunto de evaluación del desafío <i>BraTS'23</i> , cada columna muestra una secuencia de captura diferente (T1,T1ce,T2,FLAIR). Las imágenes presentan la segmentación automática de GBM en casos de RM, resaltando con distinto color las subestructuras que componen al tumor: amarillo para Edema, cafe para el tumor realzado y azul para el núcleo del tumor. En la práctica la secuencia FLAIR es la que consultan con más frecuencia los especialistas para armar un diagnóstico o tratamiento.	80
7.6	Reconstrucción 3D de la segmentaciones automática realizada por el modelo entrenado Swin UNETR de un caso de GBM del conjunto de prueba.	81
7.7	Reconstrucción 3D del cerebro y las subestructuras del tumor provenientes de las predicciones del plano axial realizadas por el modelo entrenado.	82
7.8	La figura muestra A) reconstrucción completa del tumor, B) vista coronal y C) vista sagital de la reconstrucción del tumor con el cerebro.	83
7.9	Comparación de A) un corte de RM del plano axial de un caso de prueba junto con B)la reconstrucción 3D del tumor completo y c) la reconstrucción 3D del tumor sobre el corte de RM. . .	84
7.10	La Figura muestra la reconstrucción de las diferentes subestructuras del tumor y un corte de plano axial del estudio de RM que sirve como referencia de la posición del tumor con respecto al cerebro. A) Reconstrucción 3D del edema, B) reconstrucción del tumor completo y el núcleo tumoral, C) reconstrucción con todas las subestructuras del tumor y D) Reconstrucción 3D del tumor completo y el cerebro.	85

Índice de tablas

6.1	Hiperparámetros y sus valores para entrenamiento del modelo Swin UNETR.	67
6.2	Partición de datos para entrenamiento, validación y prueba para implementación del modelo.	68
7.1	Cálculo de coeficientes DICE del conjunto de prueba mostrando los índices da cada una de las clases de salida del modelo, junto con los promedios y la desviación estándar de cada caso, y el promedio global del conjunto de datos.	77
7.2	Cálculo de coeficientes Hausdorff del conjunto de prueba mostrando los índices da cada una de las clases de salida del modelo, junto con los promedios y la desviación estándar de cada caso, y el promedio global del conjunto de datos	78

Acrónimos

ACs agentes de contraste. 15

FLAIR *Fluid attenuated inversion recovery*. 35

fRM resonancia magnética funcional. 3, 4

GBM glioblastoma multiforme. 1

IRM imágenes resonancia magnética. 4, 5

ITD imagen de tensor de difusión. 3

MET microambiente tumoral. 7

OMS Organización Mundial de la Salud. 1

RF radiofrecuencia. 31

RM resonancia magnética. 9

RT radioterapia. 8

SE eco de espín. 32

SLP supervivencia libre de progresión. 18

SNC sistema nervioso central. 9

T1 tiempo de relajación longitudinal. 29

T2 tiempo de relajación transversal. 29

TC tomografía computarizada. 9

TE tiempo de eco. 29, 31

TEP tomografía por emisión de positrones. 22

TMZ temozolomida. 8

TR tiempo de repetición. 31

Glosario

actividad mitótica División celular. 1

atipia nuclear Aspecto anormal de los núcleos de las células. 1

barrera hematoencefálica Barrera celular semipermeable que envuelve los vasos sanguíneos del cerebro y la médula espinal e impide que moléculas grandes, células inmunológicas y sustancias potencialmente dañinas o causantes de enfermedades (como los virus) pasen del riego sanguíneo al sistema nervioso central. 7

Biopsia En ocasiones es necesario tomar una muestra del tejido cerebral para confirmar el diagnóstico de glioblastoma. La biopsia se realiza mediante una cirugía y permite obtener una muestra del tumor para analizarla en el laboratorio. 14

células endoteliales Tipo principal de célula que se encuentra en el revestimiento interno de los vasos sanguíneos, los vasos linfáticos y el corazón. 7

disfunción cognitiva Es el término utilizado para describir los problemas de pensamiento, como la memoria y la concentración. 3

edema peritumoral una afección en la que los líquidos se acumulan en los tejidos del cuerpo, puede ser causado por algunos tipos de quimioterapia, por ciertos cánceres y por padecimientos no relacionados con el cáncer. 3

Gemistocítico Es una variante histológica de los astrocitomas difusos que se caracteriza por presentar tendencia a una rápida progresión hacia la malignidad. 2

glía Conjunto de células que acompañan a las neuronas en el sistema nervioso central y periférico. También se llama neuroglía. 7

necrosis Muerte patológica de un conjunto de células o de cualquier tejido del organismo. 1

oligodendrogliomas Tumor primario del sistema nervioso central (SNC). Esto significa que se origina en el cerebro o la médula espinal. 9

pleomorfismo Variación en el tamaño y la forma de las células o sus núcleos. 1

proliferación microvascular Crecimiento anormal de pequeños vasos sanguíneos. 1

Resonancia Magnética (RM) Es la técnica más utilizada para detectar glioblastomas. La RM permite obtener imágenes detalladas del cerebro y detectar la presencia de tumores. 14

Tomografía computarizada (TC) Esta técnica utiliza rayos X para obtener imágenes detalladas del cerebro. Aunque la TC no es tan sensible como la RM o la TEP, puede ser útil para detectar glioblastomas en algunos casos. 14

Tomografía por emisión de positrones (TEP) Técnica utiliza una pequeña cantidad de material radiactivo para detectar áreas de alta actividad metabólica en el cerebro. Los glioblastomas suelen tener una alta actividad metabólica, lo que los hace visibles en las imágenes de este tipo de estudio. 14

tromboembolismo venoso Afección que se produce cuando se forma un coágulo de sangre (trombo) en una o más venas del cuerpo. 3

trombosis vascular Afección que ocurre cuando se forma un coágulo de sangre en una vena profunda. 1

Capítulo 1

Introducción

1.1 Aspectos generales

Un astrocitoma es un tipo de tumor cerebral que se origina en las células gliales llamadas astrocitos, que son células que apoyan el intercambio de iones en la barrera hematoencefálica. Son el tipo más común de los tumores cerebrales y puede formarse en cualquier lugar del sistema nervioso central (Figura 1.1 muestra los diferentes tipos de células gliales). Según la Organización Mundial de la Salud (OMS) los astrocitomas se clasifican de I a IV grado con base en la apariencia de las células tumorales y la velocidad de crecimiento del tumor [1]. Dentro de la clasificación de la OMS el glioblastoma multiforme (GBM) está catalogado dentro de los tumores astrocíticos como grado IV debido a que presenta atipia nuclear, pleomorfismo, actividad mitótica, trombosis vascular, proliferación microvascular y necrosis. La mayoría se consideran glioblastomas primarios, los glioblastomas secundarios son aquellos que provienen de astro-

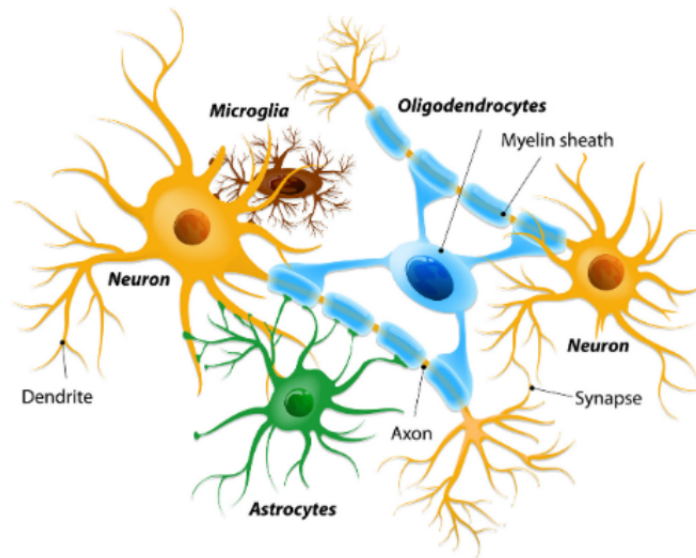


Figura 1.1: La imagen muestra diferentes tipos de células gliales las cuales dan soporte a las neuronas y conforman las vainas de mielina de los axones [3] (Imagen: “Illustration Of Neurons And Neuroglial Cells - Neuroglia@seekpng.com”).

citomas difusos (principalmente Gemistocítico) grado II o anaplásicos grado III con progresión a glioblastoma [2].

Los tumores grado I tienen bajo potencial de extenderse a otras regiones del cerebro y tienen posibilidad de curarse al ser resecados quirúrgicamente. Los tumores grado II son tumores infiltrantes pero de baja actividad proliferativa celular y en algunos casos tienden a progresar a grados superiores (III y IV). Los tumores grado III y IV son denominados de “alto grado” o “malignos”, los grado III presentan lesiones con evidencia histológica de malignidad y los grado IV tienen evidencia de malignidad citológica ya que presentan células anormales que se multiplican sin control e invaden los tejidos cercanos con predisposición a necrosis y están relacionados con una evolución rápida y fatal de la enfermedad [4].

En la mayoría de los casos el diagnóstico está basado en pruebas de imáge-

nes, específicamente de resonancias magnéticas y tomografías computarizadas, que ya tienen un alto grado de exactitud [5].

El tratamiento involucra manejo de los síntomas y signos. Es importante recordar que los problemas comunes asociados son edema peritumoral, crisis epilépticas, fatiga, tromboembolismo venoso y disfunción cognitiva. Así mismo el tratamiento establece manejo quirúrgico y terapia coadyuvante. El manejo quirúrgico tiene como fundamento tres objetivos: obtener material para diagnóstico, disminuir el efecto de masa y disminuir la carga tumoral. El grado de resección está relacionado con un aumento de la supervivencia y se recomienda ampliamente realizar cirugía para reseca la mayor cantidad de tumor posible con el fin de mejorar la calidad de vida y supervivencia, incluso en pacientes mayores y es considerada la primera opción terapéutica [6].

Actualmente este tipo de procedimiento es realizado utilizando distintos instrumentos y herramientas que apoyan al especialista a la toma de decisiones durante la planeación de un tratamiento, entre los que se encuentran la estereotaxía, la tractografía, resonancia magnética funcional intraoperatoria, equipo de ultrasonido y mapeo cortical. Todas estas herramientas son muy relevantes para poder otorgar ventajas en el tratamiento y la recuperación del paciente.

El tratamiento neuroquirúrgico ha ido evolucionando debido a mejoras en la tecnología operativa y la visualización en el quirófano. Con técnicas de imagen avanzadas como resonancia magnética funcional (fRM) o imagen de tensor de difusión (ITD), los médicos pueden obtener aproximaciones de regiones funcionales del cerebro [7]. La navegación sin marco como sistema de neuronavegación prácticamente se requiere de manera elemental en la cirugía cerebral

y puede ser una guía para el cirujano previo al procedimiento quirúrgico. Este sistema utiliza imágenes médicas preoperatorias de fRM del paciente para crear un mapa tridimensional y servir como guía de los instrumentos quirúrgicos durante la cirugía de forma segura, permitiendo un alto grado de precisión en la localización de áreas cerebrales específicas. Con este tipo de instrumentos se puede definir cuál es la mejor vía para abordar la lesión tumoral, qué riesgos se tienen que considerar si se abordan otras estructuras adyacentes, y en qué momento la remoción de tumor representa más un riesgo que un beneficio real para el paciente. La implementación de estas herramientas tiene como objetivo aumentar el grado de resección siempre y cuando se preserve al máximo las funciones cerebrales del paciente [8].

La segmentación de imágenes es un campo fundamental en la visión por computadora y una gran cantidad de métodos se han desarrollado con la finalidad de procesar imágenes y extraer información de interés contenida en ellas. La segmentación de tumores cerebrales consiste en separar los diferentes tejidos tumorales presentes en un conjunto de imágenes médicas, de los tejidos normales del cerebro, por lo general la detección de estos tejidos suele ser fácilmente detectable cuando se suministra un material de contraste al paciente previo al estudio de imagen.

La segmentación de imágenes con redes neuronales profundas es un campo en constante crecimiento en la visión por computadora y el aprendizaje profundo, estas redes son capaces de aprender patrones y características más abstractas en imágenes a partir de grandes conjuntos de datos de entrenamiento y pueden segmentar imágenes de manera precisa y eficiente. La segmentación de imágenes resonancia magnética (IRM) utilizando redes neuronales profundas para la detección de glioblastomas es una tarea importante en el campo de la

neurooncología y la imagen médica [9].

1.2 Estructura del trabajo

Con el objetivo de explorar la aplicación de tecnologías actuales para generar herramientas que puedan auxiliar a los especialistas de la salud, se propone la aplicación de modelos de redes neuronales profundas para realizar tareas de segmentación de imágenes y hacer frente a los desafíos de la interpretación de imágenes médicas. Se plantea la implementación de un sistema automatizado para realizar la tarea de segmentación de glioblastomas en imágenes de resonancia magnética y, partiendo de esta segmentación, realizar una reconstrucción en tres dimensiones del tumor completo, que ofrecerá una representación visual de los volúmenes donde se ubica el tumor y de las zonas que podrían sufrir afectación por su presencia.

Para la realización de este proyecto se plantea la siguiente estructura de trabajo, con una descripción de cada uno de los capítulos en los que se reparte.

En el capítulo 2 se presenta una descripción general sobre el caso de estudio de GBM, la incidencia de casos, su diagnóstico, técnicas de tratamiento y el seguimiento al paciente, que componen el tratamiento completo para este tipo de neoplasia.

En el capítulo 3 se realiza una revisión general de los principios básicos para la obtención de IRM y las secuencias de captura que suelen utilizarse para la segmentación de imágenes médicas.

En el capítulo 4 se revisan los fundamentos teóricos de las arquitecturas de redes neuronales convolucionales y de aprendizaje profundo. Así mismo se presentan algunas de las herramientas de *software* disponibles actualmente para realizar tareas de segmentación de imágenes médicas que implementan modelos de inteligencia artificial y aprendizaje profundo.

El capítulo 5 contiene una descripción de las distintas bases de datos públicas existentes de imágenes cerebrales. Se presentan también las características de la base de datos y la arquitectura del modelo de aprendizaje profundo que se implementan para la realización de este proyecto.

El capítulo 6 describe la metodología utilizada para la implementación del proyecto, el preprocesamiento y partición del conjunto de datos, y los hiperparámetros utilizados para el entrenamiento, validación y evaluación del modelo. Por último se describen las métricas de evaluación utilizadas para la implementación de este tipo de métodos de segmentación de GBM en IRM.

En el capítulo 7 se analizan los resultados obtenidos durante el proceso de entrenamiento del modelo, las predicciones obtenidas, así mismo se describe el proceso de reconstrucción en tres dimensiones de las predicciones obtenidas de la segmentación de GBM en IRM.

En el capítulo 8 se presentan las discusiones y conclusiones generales del proyecto, las posibles aplicaciones prácticas de los resultados obtenidos y los trabajos futuros donde se puede ampliar la aplicación de la segmentación de IRM utilizando modelos de aprendizaje profundo para el mapeo de áreas funcionales en el tejido cerebral y su relación con el diagnóstico y tratamiento de GBM.

Capítulo 2

Caso de estudio: Glioblastoma

2.1 Astrocitomas y Glioblastoma multiforme

Un astrocitoma es un tipo de tumor cerebral que se origina en las células gliales, llamadas astrocitos, que apoyan el intercambio de iones en la barrera hematoencefálica. Los astrocitomas se clasifican como tumores cerebrales de grado I a IV, siendo el grado IV el más agresivo conocido como glioblastoma multiforme o GBM [2]. Sin tratamiento, la supervivencia media de los pacientes es de aproximadamente 6 meses, que puede ampliarse a 15 meses con el uso de terapias multimodales. La baja eficacia de las terapias contra el GBM se debe principalmente a la infiltración tumoral en el tejido cerebral sano, que depende de la interacción de las células del GBM con el microambiente tumoral (MET). La interacción de las células del GBM con el MET implica componentes celulares como las células madre, la glía, las células endoteliales y componentes no celulares como la matriz extracelular, la hipoxia aumentada

y factores solubles como la adenosina, que promueven la capacidad de invadir tejidos del GBM [10].

En la actualidad se han realizado esfuerzos para identificar posibles biomarcadores que permitan mejorar el diagnóstico y tratamiento. A pesar de contar con datos específicos acerca de la incidencia de los tumores encefálicos y de que el glioblastoma multiforme es el tumor encefálico maligno más frecuente en el mundo, en nuestro país existen pocos reportes acerca de la incidencia de GBM.

Los síntomas del GBM pueden incluir dolor de cabeza, cambios en la visión, cambios en el estado de ánimo y problemas de memoria. El tratamiento actual es complejo y consiste en una terapia multimodal, que consiste inicialmente en una resección quirúrgica de la mayor cantidad de tumor posible preservando la función del paciente, seguida de radioterapia (RT) y quimioterapia, siendo la temozolomida (TMZ) el fármaco de referencia [11]. A pesar de la terapia multimodal, la supervivencia media de los pacientes no supera los 15 meses y sólo el 26 % de los pacientes vive hasta 2 años después del tratamiento [10]. La baja eficacia de las terapias contra el GBM se debe principalmente a la infiltración de células tumorales en el tejido cerebral sano, lo que dificulta la resección quirúrgica total del tumor y perjudica el efecto localizado de la radioterapia, conduciendo inevitablemente a la recurrencia del GBM con desenlaces fatales [11].

Al día de hoy el GBM se compone de tipos, primarios y secundarios, que constituyen distintas entidades de enfermedad que evolucionan a través de diferentes vías genéticas, afectan a pacientes de diferentes edades y probablemente difieren en el pronóstico y la respuesta al tratamiento. El GBM primario

representa más del 80 % del GBM, ocurre en pacientes de edad avanzada (edad media = 64 años). El GBM secundario se desarrolla a partir de astrocitoma o oligodendrogliomas de grado bajo, ocurre en pacientes más jóvenes (edad media = 45 años) [12]. Sin embargo, debido a la agresividad del tumor, la tasa de supervivencia a largo plazo es baja ya que tiene un mal pronóstico pues sólo una tercera parte de los pacientes sobreviven en el primer año y únicamente el 5 % a los 5 años, la recurrencia es frecuente. De acuerdo con el Registro Histopatológico de Neoplasias, en México, la proporción de casos nuevos fue del 1.3 %, con un 56 % en hombres y un 44 % en mujeres [13].

Los glioblastomas representan un porcentaje importante de las neoplasias intracraneales, si bien no hay una causalidad definida es el tipo de glioma más común que representa hasta 77-81 % de todos los tumores malignos primarios del sistema nervioso central (SNC) [14].

2.1.1 Incidencia

La incidencia de nuevos tumores cerebrales primarios en Estados Unidos se estima que es de 18 por cada 100,000 habitantes, lo que resulta en aproximadamente 40,000 nuevos tumores cerebrales primarios por año, 22,000 de los cuales son tumores de alto grado. Un total de 12,920 muertes en 2008 se atribuyeron a tumores cerebrales malignos primarios, a pesar de los esfuerzos continuamente renovados en el tratamiento del GBM. Los análisis anteriores mostraron una incidencia cada vez mayor de tumores cerebrales malignos durante las décadas de 1980 y 1990, pero se cree que esto es el resultado del sesgo de detección causado por el aumento en el acceso a tecnologías de imágenes médicas como la TC y RM [15] (Figura 2.1).

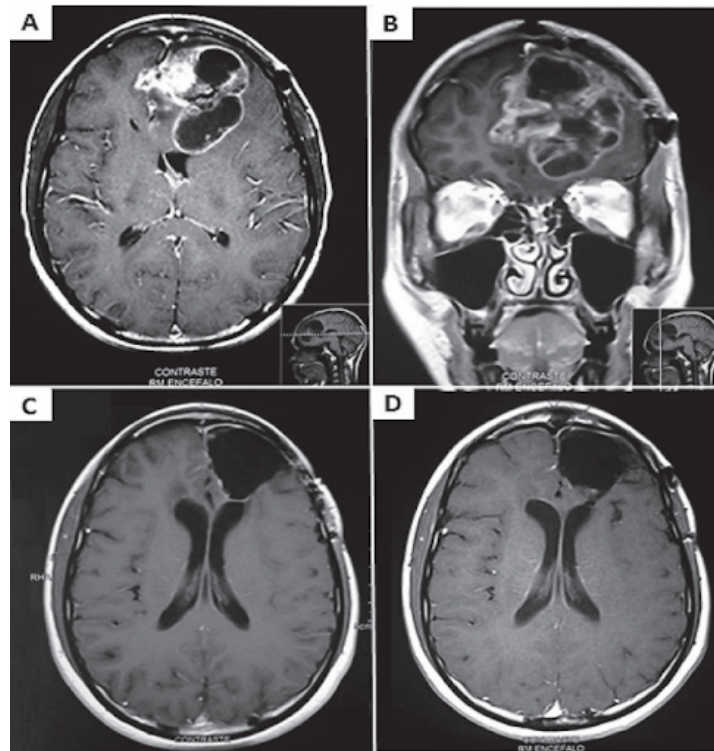


Figura 2.1: A) IRM de encéfalo con contraste preoperatorio, donde se evidencia el tumor sólido con bordes irregulares y el componente quístico frontal izquierdo que capta el medio de contraste . Desviación de la línea media cerebral y compresión del ventrículo lateral izquierdo con efecto de masa: A) imagen axial. B) Imagen coronal. C) RM encéfalo T1 con contraste posoperatorio luego de un mes que demuestra los cambios luego de la cirugía, no evidencia realce de contraste. D) RM encéfalo T1 con contraste posoperatorio, control a los 3 años de la cirugía, no evidencia realce de contraste [16] (Imagen: “Revista Peruana de Medicina Experimental y Salud Publica [16]”).

El GBM se puede presentar de dos maneras: primario, sin lesiones precursoras evidentes, o secundario, a partir de la transformación de un tumor glial difuso de grado menor progresando a un grado superior. Una característica de este tipo de tumor es que debido a su naturaleza de infiltrarse en el tejido sano que lo rodea y su alta capacidad de extenderse, no puede ser extraído del todo, por lo que la supervivencia es menor a un año en el 50 % de los casos. La supervivencia, una vez que se ha realizado el diagnóstico de GBM primario o secundario, es similar, aunque en estos últimos la supervivencia total depende del grado de lesión original que dio lugar al GBM [15].

Los GBM pueden observarse en la región subcortical, en cualquiera de los dos hemisferios. La localización del tumor varía según la región del cerebro: en el 31 % de los casos se encuentra en el lóbulo temporal; en el 24 %, en el lóbulo parietal; en el 23 %, en el lóbulo frontal, y en el 16 %, en el lóbulo occipital. La presencia del glioblastoma en el área frontotemporal también es frecuente [2]. Los GBM rara vez se localizan en el cerebelo y son muy raros en la médula espinal, con un comportamiento tumoral diferente en estos lugares. La ubicación de GBM en el cerebelo es más común en pacientes más jóvenes (50–56 años de edad); la ubicación supratentorial es prevalente en pacientes mayores (62-64 años de edad) y la ubicación en el cerebelo es rara (0.4-3.4 %) en este grupo de edad. Para los GBM de la médula espinal, la edad media es de 27 años, con predominio masculino; El 53 % de estos tumores se observan en menores de 18 años [11].

En México, de acuerdo con el Registro Histopatológico de Neoplasias, la proporción de casos nuevos fue del 1.3 %, con un 56 % en hombres y un 44 % en mujeres [12]. GLOBOCAN 2020 reportó en México 3,189 casos (1.6 %) y una mortalidad de 2,571 casos (2.8 %), así como una prevalencia de 8,806 casos [17].

2.2 Diagnóstico

Dependiendo de la ubicación y del tipo de tumor pueden aparecer diferentes síntomas, especialmente debido a que el tumor crece y afecta la médula espinal, los nervios que la rodean o los vasos sanguíneos. Estos tumores causan gran nivel de discapacidad y sufrimiento al paciente por su sintomatología y daño medular [18]. Su localización más común es supratentorial, la ubicación a ambos lados del tentorio del cerebelo (supra e infra) es todavía más inusual, de hecho, no hay una frecuencia consensuada de esta forma de presentación. En general se diferencian en lesiones multicéntricas y multifocales [19] (Figura 2.2). En cuanto a los síntomas que presenta el astrocitoma especialmente están en función de la localización del tumor. Si están en el cerebro suelen causar convulsiones, dolores de cabeza y vómitos. Los que están en la médula espinal pueden causar cierta discapacidad en la zona funcional afectada por el tumor. Por ello, el médico ha de realizar como primer paso un examen neurológico en el que habla con el paciente y obtenga información acerca de su comportamiento y estado general [20].

2.2.1 Evaluación clínica del paciente

La evaluación clínica de los pacientes afectados por glioblastoma se realiza de forma integral tomando en cuenta el examen físico general, neurológico y de las funciones mentales superiores. Cuando el paciente se presenta con datos de hipertensión intracraneal, se debe realizar un abordaje rápido y eficiente con el fin de asegurar el bienestar del paciente, ofreciéndole estudios auxiliares de diagnóstico y medidas terapéuticas tempranas [22].

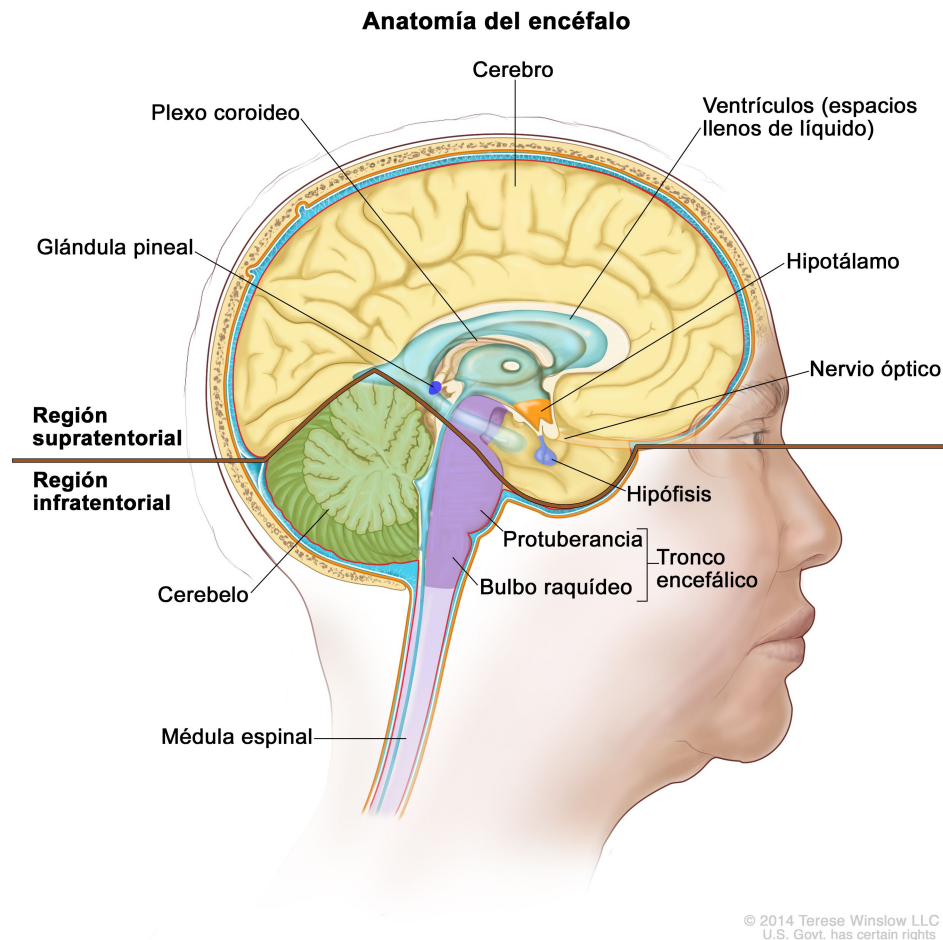


Figura 2.2: Anatomía del interior del encéfalo. La región supratentorial contiene el cerebro, los ventrículos, el plexo coroideo, el hipotálamo, la glándula pineal, la hipófisis y el nervio óptico. La región infratentorial contiene el cerebelo y el tronco encefálico [21] (Imagen: “Tratamiento del ependimoma infantil PDQ®—Versión para profesionales de salud [21]”).

Tras el establecimiento de estas medidas, puede observarse una mejoría del estado clínico del paciente tanto en los datos de hipertensión intracraneana como en los de focalización neurológica, por lo que la evaluación del paciente debe ser dinámica en función de la respuesta y no al manejo inicial. Con la mejoría del estado clínico del paciente se puede pasar a la indicación de estudios más complejos, como la valoración neuropsicológica, los estudios del lenguaje o los estudios en los que la participación del paciente sea definitiva, como la fRM. Cuanto más completo sea el estudio de un paciente, mejores datos se tendrán para hacer una propuesta terapéutica [22].

2.2.2 Diagnóstico con imagenología

Aunque el GBM puede presentarse con morfologías y ubicaciones diversas, casi todas las lesiones tienen características típicas de imagen que reflejan la fisiopatología de este tumor. De las diversas técnicas para la detección de glioblastomas, se encuentran:

- Resonancia Magnética (RM).
- Tomografía por emisión de positrones (TEP).
- Tomografía computarizada (TC)
- Biopsia

La detección temprana de glioblastomas no siempre es fácil y puede requerir la combinación de varias técnicas de diagnóstico.

La RM es una técnica de imagen muy potente que se utiliza frecuentemente en la detección clínica y la observación del pronóstico. Su utilidad se debe a su alta resolución espacial, su contraste con los tejidos blandos y que no es una técnica invasiva [23]. Aunque es muy eficaz en la detección de tejidos blandos, puede dar lugar a una mala representación de la imagen, haciendo que sea imposible diferenciar el tejido maligno del sano [24]. El contraste obtenido en la RM depende de la concentración de las moléculas de agua dentro del tejido, los tiempos de relajación (T1 y T2) de los protones de agua y la movilidad de las moléculas de agua; que reflejarán una diferencia de intensidad en la señal de los distintos tejidos [25]. La mayoría de los agentes de contraste (ACs) son complejos iónicos paramagnéticos o partículas de magnetita superparamagnética. La función de estos elementos es aumentar la intensidad de las imágenes ponderadas en T1 o por el contrario reducirla en el T2, aunque el primer caso es el más utilizado habitualmente en rutina clínica, denominados agentes de contraste positivos (La Figura 2.3 muestra la segmentación manual de una IRM en los distintos tiempos de relajación). Los ACs se usan en RM para mejorar la sensibilidad del diagnóstico, de manera que un 35 % de los diagnósticos por RM requieren de un ACs [26]. Son muchas las investigaciones que se han realizado sobre ACs con el fin de mejorar el diagnóstico clínico, basándose en la mejora de la calidad y contraste de las imágenes [22].

Es importante mencionar que la resonancia magnética por sí sola no puede confirmar el diagnóstico de glioblastoma, por lo que generalmente se requiere una biopsia para confirmar el diagnóstico. Sin embargo, la resonancia magnética es una herramienta valiosa en la detección, diagnóstico y seguimiento de glioblastomas y otros tumores cerebrales.

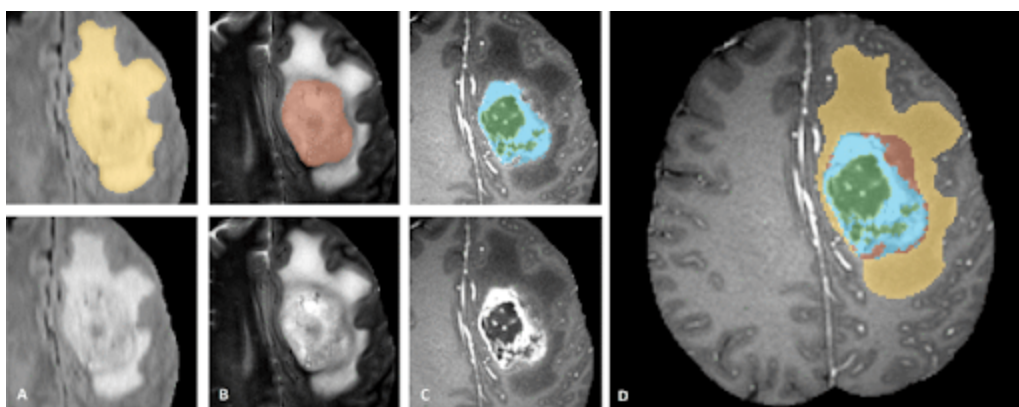


Figura 2.3: Anotación manual a través de evaluadores expertos. Se muestran parches de imagen con las estructuras tumorales que se anotan en las diferentes modalidades (arriba a la izquierda) y las etiquetas finales para todo el conjunto de datos (derecha). Los parches de imagen muestran de izquierda a derecha: todo el tumor visible en FLAIR, (A) el núcleo del tumor visible en T2 (B), las estructuras tumorales realzadas visibles en t1c (azul), que rodean los componentes quísticos/necróticos del núcleo (verde) (C). Las segmentaciones se combinan para generar las etiquetas finales de las estructuras tumorales (D): edema (amarillo), núcleo sólido sin realce (rojo), núcleo necrótico/quístico (verde), núcleo realzado (azul) [27] (Imagen: “The Multimodal Brain Tumor Image Segmentation Benchmark [27]”).

2.3 Tratamiento

El tratamiento de glioblastomas depende de varios factores que incluyen la ubicación y el tamaño del tumor, la edad y el estado de salud general del paciente. El tratamiento puede incluir una combinación de cirugía, radioterapia y quimioterapia. En el caso de glioblastomas, la tractografía se utiliza para evaluar la integridad de las vías de conexión que atraviesan el tumor y su efecto sobre la conectividad del cerebro. Esta técnica puede ayudar a los médicos a planificar la cirugía, identificar áreas críticas del cerebro para evitar daños durante la resección del tumor y evaluar el grado de invasión del tumor en áreas cerebrales funcionales cercanas [28].

2.3.1 Manejo preoperatorio

La RM funcional ofrece información sobre el dominio cerebral, con localización anatómica de las áreas críticas para el lenguaje, pero la RM no investiga realmente la bilateralidad potencial del lenguaje. La RM funcional es más útil que una prueba de Wada para lesiones del hemisferio dominante cerca de la corteza motora, lóbulos frontales, pars triangularis y opercularis (área de Broca) o área de Wernicke. Con los cambios en la actividad metabólica y la demanda de flujo sanguíneo, un área activa del cerebro durante un discurso silencioso o una tarea motora se infunde con más sangre oxigenada; este cambio se puede detectar ya que la sangre oxigenada transporta una señal paramagnética diferente a la desoxigenada (señal dependiente del nivel de oxígeno en sangre). Una limitación de la IRM funcional es que se vuelve menos útil en pacientes con un tumor recurrente debido a patrones vasculares alterados y artefactos de RM de la cirugía previa. Algunos pacientes requieren tanto una Wada como una IRM funcional como parte de la planificación preoperatoria [15].

2.3.2 Cirugía

La cirugía se utiliza para extirpar la mayor cantidad posible de tejido tumoral. En algunos casos, el glioblastoma se encuentra en una zona del cerebro que hace que la cirugía sea difícil o peligrosa, por lo que no siempre es una opción. La cirugía también puede ayudar a aliviar la presión en el cerebro causada por el tumor [11]. La resección quirúrgica es habitualmente la modalidad terapéutica inicial en pacientes de diagnóstico reciente, que tienen como factor un pronóstico alto de supervivencia a una resección amplia. Esto se demuestra

por el metanálisis que agrupó a los tres grandes estudios prospectivos multicéntricos aleatorizados y confirmó el efecto beneficioso de la resección extensa en los GBM. Así mismo, una resección extensa potencia la eficacia de las terapias adyuvantes y facilita el manejo clínico de dichos pacientes. La cirugía citorreductiva proporciona una muestra histológica más robusta, que permite un diagnóstico más preciso. La reducción de la carga tumoral tras la cirugía resectiva proporciona un alivio de la presión intracraneal y, por consiguiente, una reducción del edema peritumoral [29].

2.3.3 Técnicas intraoperatorias, mayor seguridad en el tratamiento

La craneotomía del paciente despierto, aunada al mapeo corticosubcortical y el monitoreo intraoperatorio, constituye actualmente el estándar de oro. Así mismo, la neuronavegación intraoperatoria se ha convertido en un complemento quirúrgico muy difundido; sin embargo, su precisión se ve comprometida una vez que se ha realizado la apertura dural [29].

Se ha demostrado un aumento en la tasa de resección completa y una mejora en la supervivencia supervivencia libre de progresión (SLP) [30], con el empleo de fluorescencia con ácido 5-aminolevulínico (5-ALA). La fluorescencia mediante 5-ALA tiene un valor predictivo alto para detectar realce del tejido afectado en los gliomas de alto grado como el Glioblastoma, así como para evidenciar focos anaplaésicos en gliomas de bajo grado. El efecto máximo de fluorescencia ocurre entre 6-9 horas después de su ingesta. Se enciende la luz ultravioleta del microscopio quirúrgico, en función de la fluorescencia las

zonas de más captación en rojo corresponden a tumor sólido, mientras que las zonas que se tiñen de rosa corresponden a células tumorales infiltradas o borde invasivo. El tejido sano se aprecia de color azul, el tejido necrótico no se tiñe y que algunas estructuras como los plexos coroides se tiñen de color rosáceo [31] (Figura 2.4).

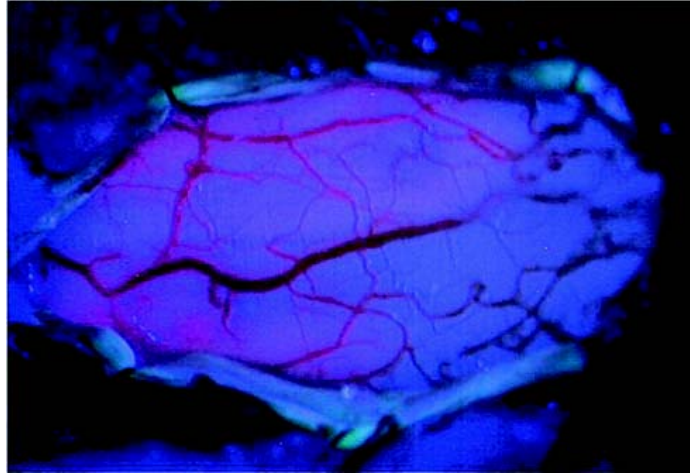


Figura 2.4: Médula cervical con luz violeta-azul. Se aprecia la fluorescencia rosada por transiluminación [32] (Imagen: “Resección de tumor intramedular guiada por fluorescencia con ácido aminolevulínico [32]”).

2.3.4 Biopsia estereotáctica

La técnica de estereotaxia es un método de neurocirugía que utiliza un marco estereotáxico para guiar la localización y el tratamiento de lesiones cerebrales. Se basa en la utilización de imágenes médicas para crear un mapa tridimensional del cerebro y proporciona una guía precisa para la inserción de instrumentos quirúrgicos en el cerebro. El marco estereotáxico se compone de tres partes principales: un soporte para la cabeza, un conjunto de varillas y un sistema de fijación que asegura que el conjunto de varillas no se mueva durante la cirugía. Una vez que se ha colocado el marco estereotáxico en la

cabeza del paciente, se realizan IRM o TC para localizar la estructura cerebral de interés en relación con el marco estereotáxico. La imagen se utiliza para crear un mapa tridimensional del cerebro que permite una orientación precisa de la estructura cerebral [33] (Figura 2.5 Muestra el abordaje quirúrgico con la ayuda de un sistema de marco estereotáxico).

Durante la cirugía, se utiliza el marco estereotáxico para guiar la colocación de instrumentos quirúrgicos en el cerebro, como agujas, electrodos o sondas. El marco estereotáxico proporciona una guía precisa para la inserción de estos instrumentos en el cerebro, lo que permite una cirugía menos invasiva y más precisa [33].

2.3.5 Radioterapia

La radioterapia utiliza radiación para destruir las células cancerosas. Se puede administrar después de la cirugía o como tratamiento principal si la cirugía no es posible. La radioterapia puede tener efectos secundarios, como fatiga, náuseas y pérdida de cabello [35]. Aunque no hay un consenso claro para el seguimiento de los pacientes postoperados de gliomas de alto grado, la tendencia global es que todo paciente sometido a cirugía en quien se haya realizado una resección significativa deberá evaluarse mediante RM posoperatoria dentro de las 24 a 48 horas posteriores a la intervención. Esto servirá como estudio base y de comparación con posteriores, seguido de estudio de seguimiento a las 3 semanas. Posteriormente deberá individualizarse de acuerdo con los hallazgos de biología molecular con un rango de frecuencia que varía de 2 a 3 meses [36].

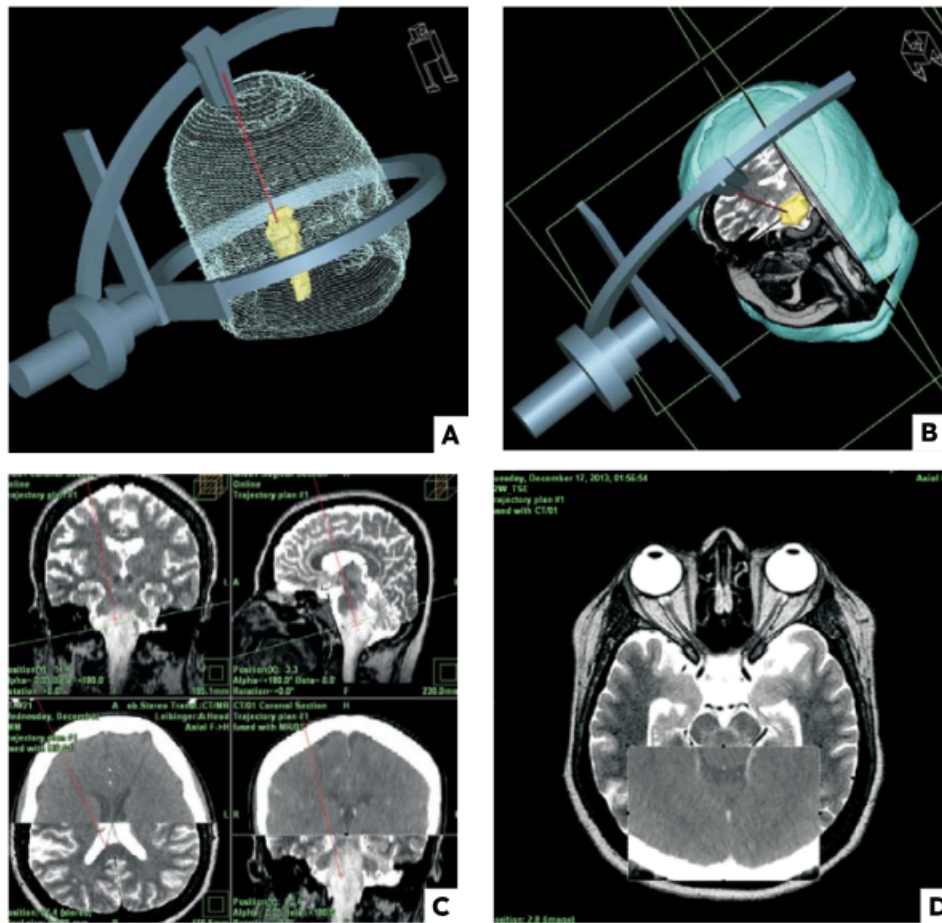


Figura 2.5: Realización de biopsia de alta precisión con posicionamiento de marco estereotáctico en lesiones del tallo cerebral (A, B) utilizando software de alta precisión para la planeación de trayectorias estereotácticas con fusión de imágenes (C, D) [34](Imagen: “Revista Gastrohnpup [34]”).

2.3.6 Quimioterapia y Terapia dirigida

La quimioterapia se utiliza para destruir las células cancerosas. Puede administrarse por vía oral o mediante una inyección en una vena. La quimioterapia puede tener efectos secundarios, como náuseas, vómitos y pérdida de cabello [10]. Los objetivos de la quimioterapia son: detener el crecimiento del tumor haciendo que las células tumorales no puedan copiarse a sí mismas e iniciar artificialmente el proceso normal de muerte celular (apoptosis) en el tumor. A la fecha existen cuatro medicamentos aprobados para el tratamiento de esta enfermedad: temozolomida, bevacizumab, carmustina y lomustina, así como el dispositivo OptuneTTFields (Campos de tratamiento tumoral), sin embargo, la esperanza de vida se incrementa entre 6 meses y 2 años [29].

La terapia dirigida es un tratamiento que utiliza medicamentos para atacar directamente las células cancerosas. Estos medicamentos suelen ser específicos para un tipo particular de cáncer y pueden ser más efectivos que la quimioterapia convencional. Algunas terapias dirigidas se administran por vía oral, mientras que otras requieren inyecciones [35].

2.4 Seguimiento del tratamiento

Los pacientes en tratamiento complementario deben ser seguidos por medio de revisiones clínicas, estudios de laboratorio y de RM, y en casos especiales estudios de TEP. El seguimiento se inicia con la RM de control postoperatorio, que es de suma utilidad para calcular el volumen residual de la lesión tras la realización de la cirugía. Posteriormente los pacientes deberán ser eva-

luados con RM no multimodal cada 2 a 3 meses para conocer la respuesta al tratamiento y evaluar el comportamiento del tumor. Siempre es recomendable realizar una evaluación multidisciplinaria de neurooncología en los casos de controversia para determinar, mediante consenso, el tipo de evolución del tumor y la conducta terapéutica a seguir [35].

El tratamiento neuroquirúrgico ha evolucionado con avances en la tecnología operativa y la visualización en el quirófano. M. Gazi Yasargil es una figura importante en neurocirugía debido a su desarrollo de técnicas y herramientas novedosas [37]. Con técnicas de imagen avanzadas como fRM o ITD los médicos pueden obtener aproximaciones de regiones funcionales del cerebro. El mapeo funcional despierto y los procedimientos de lateralización del lenguaje preoperatorio sirven como medios más precisos para determinar todas las regiones funcionales dentro del cerebro. Las técnicas de mapeo funcional de alta calidad han mejorado la seguridad de la resección neuroquirúrgica para pacientes con gliomas. La utilidad de tales técnicas varía junto con el espectro de la enfermedad patológica porque los pacientes con tumores de bajo grado y aquellos que hacen la transición de tumores de bajo grado han demostrado una capacidad creciente de plasticidad cerebral [35].

Capítulo 3

Imagen de resonancia magnética

3.1 Aspectos generales

Las IRM son una técnica de adquisición de imágenes que permiten obtener la anatomía interna del cuerpo de una manera segura y no invasiva. Se basa en los principios de la resonancia magnética nuclear y permite una amplia variedad de tipos de secuencias. Este medio de formación de imágenes ha sido de particular relevancia para la producción de imágenes cerebrales debido a la capacidad de las IRM para distinguir entre señales de diferentes tejidos blandos y así diferenciarlos (como la materia gris y blanca)[38]. En imágenes del cerebro, las secuencias tipo T1 y T2 son las más utilizadas, éstas se refieren a la señal dominante medida para producir el contraste observado en la imagen.

En la visualización de tumores cerebrales, una segunda secuencia T1 es adquirida después de la inyección de un agente de contraste que provoca una

disminución en el tiempo T1 del tejido que absorbe la sustancia. El realce provocado por esta modalidad es un fuerte indicador de la presencia de un tumor cerebral, sin embargo existen una gran variedad de tipos de tumores cerebrales y su aparición en el estudio de RM pueden variar considerablemente. El edema también se puede observar en muchos tipos de tumores primarios, y aparece como hiperintensidad en las imágenes T2 [39].

3.2 Propiedades microscópicas de la IRM

La imagen por resonancia magnética se basa en la absorción y emisión de energía, dentro del rango de la radiofrecuencia en el espectro electromagnético. Muchos científicos se formaron con la idea de que no se puede obtener una imagen de un objeto más pequeño que la longitud de onda de la energía utilizada para el estudio. La IRM supera esta limitación al producir imágenes basadas en variaciones espaciales de fase y frecuencia de la energía de radiofrecuencia que absorbe y emite el objeto evaluado.

El cuerpo humano está formado principalmente de grasa y agua que constituyen aproximadamente el 63 % del cuerpo humano. Los núcleos de hidrógeno de estos compuestos generan una señal de resonancia magnética nuclear. Cada voxel de una imagen del cuerpo humano contiene uno o más tejidos. Un zoom sobre el voxel revelaría que hay células y dentro de cada célula habría moléculas de agua, cada molécula de agua tiene un átomo de oxígeno y dos de hidrógeno (Figura 3.1). Sobre cada uno de los hidrógenos, pasando la nube de electrones, tenemos un núcleo compuesto por un único protón. El protón posee una propiedad llamada espín la cual puede ser considerada como un peque-

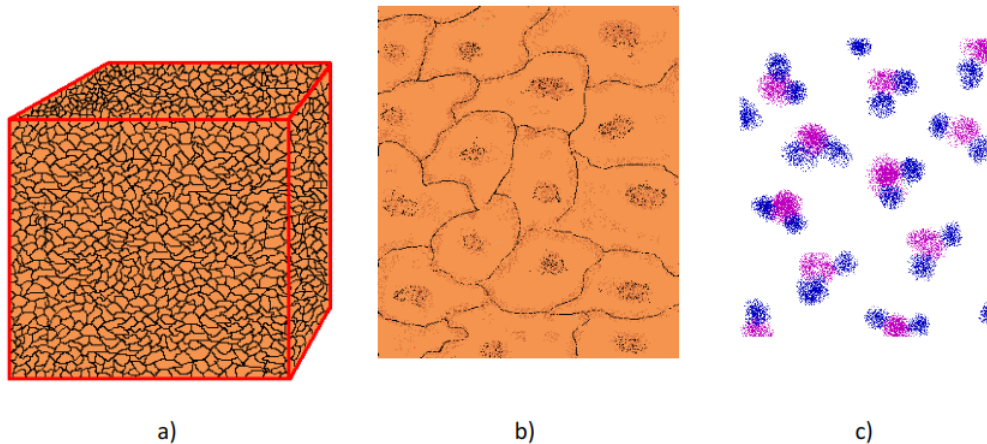


Figura 3.1: Propiedades microscópicas responsables de la MRI. a) Tejidos en un voxel b) Células en los tejidos c) Moléculas de agua en los tejidos [40] (Imagen: “MR pulse sequences: what every radiologist wants to know but is afraid to ask”, 2006 [40]).

ño campo magnético, y hace que el núcleo produzca una señal de resonancia magnética nuclear [40].

Los protones, al tener carga eléctrica y polarización magnética, se alinean en un campo magnético externo, estos campos pueden alinearse paralelamente a un campo magnético externo o de forma antiparalela al campo magnético. La alineación predominante es la que necesita menor uso de energía, así que se alinearán más protones en sentido paralelo. Sin embargo, la diferencia en el número de protones alineados paralelamente, en comparación con los que lo hacen forma antiparalela, es muy poca. Estos protones, además de alinearse paralela o de forma antiparalela al campo magnético, se mueven alrededor de él a manera de una pirinola, movimiento al que se denomina precesión [40] (Figura 3.2).

El movimiento de precesión, se define como el giro que realiza el protón alrededor del eje del campo magnético externo donde lo situemos, en este

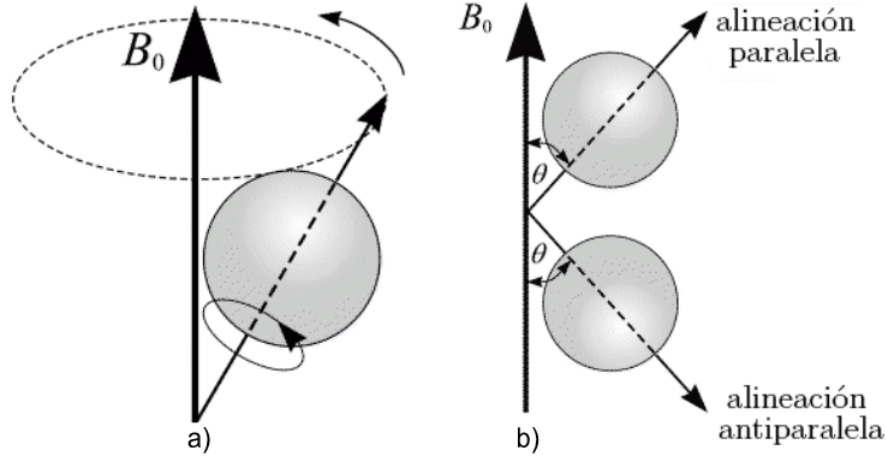


Figura 3.2: a) Precesión de un protón. La imagen corresponde al movimiento del eje de rotación sobre las curvas discontinuas. b) La alineación del espín con el campo magnético nunca llega a realizarse del todo, en virtud del origen cuántico del espín; lo más que puede lograrse es una alineación en el sentido de la dirección del campo magnético que llamamos “paralela”, o una alineación en sentido contrario que llamamos “antiparalela” [38] (Imagen: “MRI, basic principles and applications [38]”).

caso, el momento magnético μ (espín) precesará alrededor de la dirección del campo externo, B_0 , cuya frecuencia está determinada por la intensidad del campo magnético al que se somete al paciente. Cuanto más intenso es el campo magnético, mayor será la frecuencia de precesión (ω), que podrá calcularse según la ecuación de Larmor, donde B_0 es la intensidad del campo magnético externo (que se mide en *Teslas* o *T*), y γ es una constante giromagnética que es diferente para cada elemento (el valor para el protón de hidrógeno es de 42.5 MHz/T).

$$\omega = \gamma B_0 \quad (3.1)$$

De la Ecuación 3.1 se concluye que la frecuencia de precesión (ω) es igual a la constante giromagnética por la intensidad del campo magnético. Es de gran importancia conocer la frecuencia de precesión en la generación de imágenes de resonancia magnética [41].

El resultado de someter al paciente a un campo magnético es la obtención de una magnetización neta, a partir del exceso de protones que se orientan paralelamente al eje de un campo magnético externo. Como esta magnetización se encuentra en la dirección del campo magnético, se le denomina magnetización longitudinal. Esta magnetización longitudinal se usa para obtener señales de resonancia magnética; sin embargo, no aporta información sobre el contraste magnético de los tejidos (que es lo que nos interesa), por lo que se hace necesario el uso de pulsos cortos de ondas electromagnéticas, que se denominan pulsos de radiofrecuencia, cuyo propósito es perturbar ese exceso de protones que se encuentran en un menor nivel de energía y que conforman la magnetización longitudinal [42].

De los pulsos, no todos logran perturbar el vector de magnetización longitudinal; solamente un pulso de radiofrecuencia con la misma frecuencia de precesión puede hacerlo entrar en resonancia e interactuar. Al aplicar el pulso de radiofrecuencia a los tejidos, esto provoca dos efectos. El primero es proporcionar energía a los protones causando que algunos de ellos se dispongan en la posición antiparalela, el segundo hace que los protones entren en sincronismo y comiencen a precesar en fase, En resumen, el pulso de radiofrecuencia hace que disminuya la magnetización longitudinal y se establezca una nueva magnetización transversal, que puede ser medida y que aporta la información sobre el contraste magnético de los tejidos [43].

Un pulso de radiofrecuencia que inclina la magnetización generando un nuevo vector de magnetización transversal que se ubica a 90 grados con respecto al vector de magnetización longitudinal inicial, se denomina “pulso de 90 grados”. Son posibles otros pulsos de radiofrecuencia denominados según el ángulo que originan; sin embargo, los pulsos más usados en la generación

de secuencias de resonancia magnética son de 90 y de 180 grados. Al enviar el pulso de radiofrecuencia, algunos protones captan energía y pasan a un nivel de mayor energía, por lo que disminuye la magnetización longitudinal [44] (Figura 3.3).

Una vez obtenido el vector de magnetización transversal, el siguiente paso que permite obtener información sobre el contraste magnético de los tejidos, consiste en la interrupción del pulso de radiofrecuencia, con lo que los protones vuelven a su menor nivel de energía y pierden la coherencia de fase. De esta manera, liberan la energía (anteriormente tomada del pulso de radiofrecuencia) como una corriente eléctrica denominada señal de resonancia magnética o eco, que es captada por una antena para cuantificar los tiempos en que la magnetización transversal desaparece, conocido como tiempo de relajación transversal (T2), y los tiempos en que la magnetización longitudinal se restablece completamente, conocido como tiempo de relajación longitudinal (T1). Al intervalo de tiempo transcurrido entre la aplicación del pulso de radiofrecuencia y la captación de la señal de resonancia magnética o eco, se le conoce como tiempo de eco (TE), factor también importante en la generación de las diferentes secuencias [43] (Figura 3.4).

La relajación transversal y la longitudinal ocurren sincrónicamente. Al tiempo de relajación longitudinal se le denomina T1 y es cuando se ha recuperado el 63 % de la magnetización longitudinal original. Al tiempo de relajación transversal se le denomina T2 y es cuando ha decrecido la magnetización transversal en un 37 % de su valor. Los tiempos de relajación son propiedades de los tejidos que no se pueden modificar y que van a proporcionar información sobre la composición del tejido examinado. Así, al variar los parámetros técnicos de las secuencias, se puede observar mejor alguno de los dos tiempos de relajación

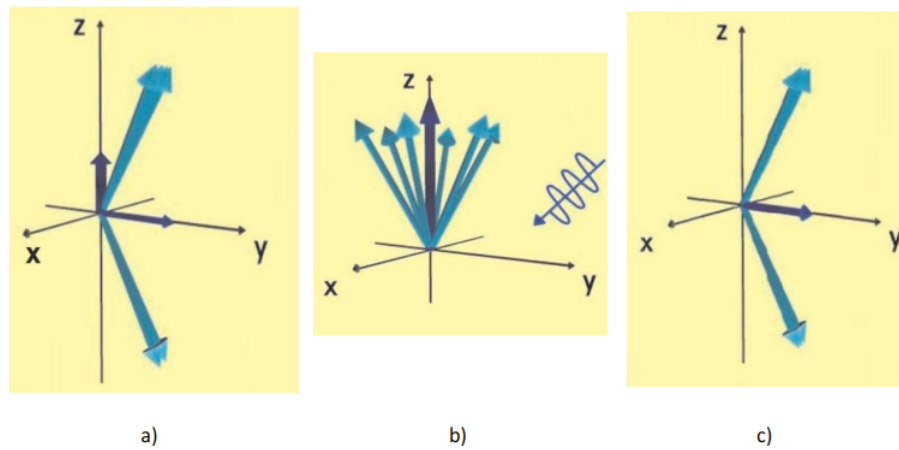


Figura 3.3: En la figura se muestra el efecto de los “pulso de 90 grados”, en a) se puede observar el efecto de un pulso de radiofrecuencia de 90 grados. En b) los protones entran en sincronismo y comienzan a precesar en fase. En c) sus vectores se suman en dirección transversal al campo magnético y se establece una magnetización transversal o un vector suma que representa el momento magnético total de un tejido [43] (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación, Universitas Médica” [43]).

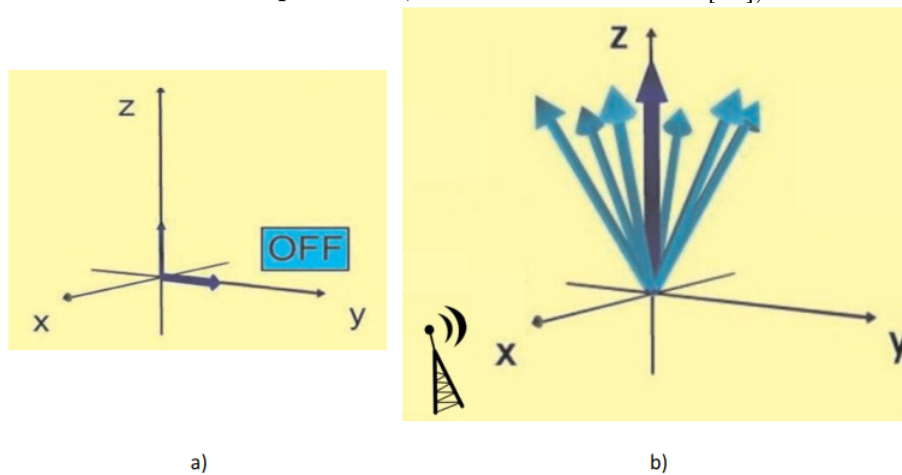


Figura 3.4: En la figura a) observamos que los protones regresan a su nivel de menor energía, disminuye la magnetización transversal T2 y se recupera la magnetización longitudinal T1. En b) se representa el proceso que genera la emisión de la señal de resonancia o eco que va a ser captado por una antena [43] (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación, Universitas Médica” [43]).

con el fin de modificar el contraste entre los tejidos [44].

Como la señal con la que se trabaja es baja, pues proviene de un bajo exceso de protones que inicialmente se alinearon paralelamente al campo magnético, es preciso repetir las mediciones varias veces para obtener una imagen que tenga utilidad clínica. Es por esto que se debe emplear más de un pulso de radiofrecuencia; a esto se le denomina secuencia de pulsos. Como pueden usarse varios tipos de pulsos (90 ó 180 grados) y los intervalos de tiempo entre pulsos sucesivos (tiempo de repetición o TR) pueden ser diferentes, el resultado es la posibilidad de generar diferentes tipos de secuencias. Así, la elección de una secuencia de pulsos determina la clase de señal que se obtiene de un tejido. Esta señal se digitaliza en una escala de grises y la terminología adecuada para describirla la califica como señal alta, intermedia, baja o nula [45].

3.3 Tipos de secuencia RM

Como los tiempos de relajación T1 y T2 son dependientes del tiempo, el momento del pulso de RF y la lectura de la energía radiada cambian la apariencia de la imagen. El TR describe el tiempo entre las emisiones de pulsos de RF. El TE describe el retardo antes de que se mida la energía del pulso RF radiada por el tejido en cuestión. Cada secuencia es una combinación de pulsos de RF y gradientes, dependiendo de estas combinaciones se puede obtener una secuencia de IRM diferente. A continuación se explican algunas de las secuencias utilizadas para analizar tumores cerebrales a partir de IRM [46].

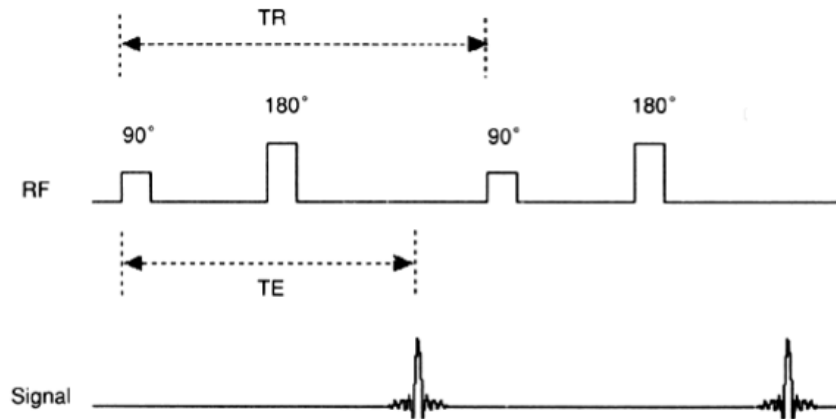


Figura 3.5: La figura muestra la secuencia SE de forma esquemática (Imagen: “Introducción biofísica a la resonancia magnética”. Centre Diagnòstic Pedralbes, 1993 [48]).

3.3.1 Secuencia eco de espín SE

La secuencia SE consiste en enviar un primer pulso de 90 grados que excita los núcleos de los átomos. Después de dejar relajar los núcleos durante un tiempo ($TE/2$) durante el cual los espines se desfazarán, se envía un pulso de 180 grados de tal forma que invierta la posición del espín respecto al campo magnético. Transcurrido exactamente un nuevo intervalo de tiempo $TE/2$ se recoge la señal. Después de un tiempo TR se vuelve a repetir el proceso [47] (Figura 3.5).

Una vez recogida la señal se deja relajar la magnetización y al cabo de un tiempo de repetición TR se repite la secuencia. Debido a que el tiempo TR es mucho mayor que el tiempo TE, una posibilidad muy útil en la práctica es obtener varios ecos dentro del mismo TR. Para ello después del primer eco se deja transcurrir de nuevo un tiempo $TE/2$ de desfase y se vuelve a enviar un pulso refasador de 180 grados, recogién dose un nuevo eco al cabo de un

tiempo $TE/2$ y así sucesivamente. Cada uno de estos ecos tiene un TE más largo. La calidad de imagen es ligeramente distinta en función de la exactitud de los pulso de radiofrecuencia [48].

3.3.2 Secuencia T1

IRM T1 se refiere a la adquisición de imágenes que representan las diferencias en el tiempo de relajación ($T1$) de los protones de los tejidos del cuerpo. Las secuencias T1 a menudo son adquiridas antes y después de la aplicación de medio de contraste. El agente de contraste en el cerebro proporciona una diferencia apreciable entre materia gris y blanca. En el cuerpo muestra una gran diferenciación entre grasa y agua (el agua se muestra oscura mientras que la grasa es más intensa [44]).

3.3.3 Secuencia T2

Empleando un tiempo de repetición largo, los tiempos de relajación longitudinal entre los tejidos serán aproximadamente iguales. Usando tiempos de eco largos, las diferencias en el tiempo de relajación transversal entre los tejidos aparecen como las diferencias en el contraste en las imágenes. Por lo tanto, usando tiempos de repetición y tiempos de eco largos, se elimina el efecto T1 y se obtienen imágenes potenciadas en T2. Esta secuencia deja ver la grasa como una señal de baja intensidad y el líquido, como una señal de alta intensidad, por lo que es útil en la identificación de lesiones patológicas que suelen caracterizarse por un aumento en el contenido de agua. Se produce también

↓ T1	↑ T2	
GRASA	AGUA LIBRE	BLANCO
HUESO MEDULAR		
SUBS. BLANCA	GRASA	
SUBS. GRIS	HUESO MEDULAR SUBS. GRIS	
MÚSCULO	SUBS. BLANCA	
AGUA LIBRE	MÚSCULO	
LIGAMENTOS	LIGAMENTOS	
HUESO CORTICAL	HUESO CORTICAL	
AIRE	AIRE	NEGRO
↑ T1	↓ T2	

Figura 3.6: La figura muestra una tabla donde se puede observar las diferencias en las intensidades de niveles de gris entre las secuencias T1 y T2 (Imagen: “Introducción biofísica a la resonancia magnética”. Centre Diagnòstic Pedralbes, 1993 [48]).

inversión en la intensidad de señal de la sustancia blanca, que se observa de menor intensidad de señal con respecto a la sustancia gris [44] (Figura 3.6).

3.3.4 Secuencia FLAIR

La recuperación de la inversión atenuada de fluido FLAIR (por sus siglas en inglés *Fluid attenuated inversion recovery*) es una secuencia de imágenes comparable con T2 pero en donde el líquido cefalorraquídeo se suprime [49]. Es usada comúnmente cuando las lesiones del cerebro aparecen muy cerca de los ventrículos. Se ha convertido en una herramienta de rutina para la evaluación de lesiones en los pacientes con trastornos neurológicos tales como la esclerosis múltiple [50] (Figura 3.7).

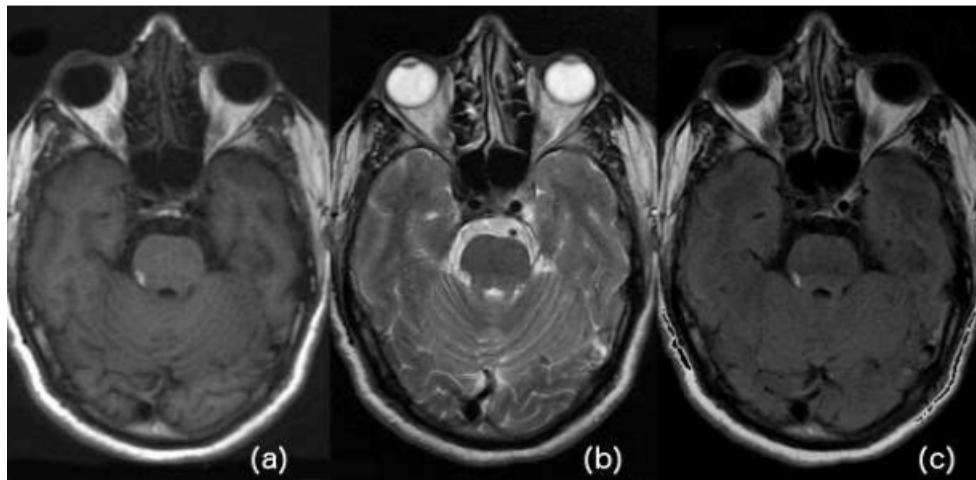


Figura 3.7: La figura compara la imagen del mismo caso de RM entre cada una de las secuencias *a)* T1, *b)* T2 y *c)* FLAIR (Imagen: “Resonancia magnética cerebral: secuencias básicas e interpretación”. Universitas Médica, 2011 [43]).

3.3.5 Medios de contraste

Son sustancias paramagnéticas que alteran los campos magnéticos locales, acortando o acelerando los tiempos de relajación, por lo que su principal efecto es un refuerzo de la relajación de los protones, acortando el T1 y el T2 de los tejidos en los que transcurren. En IRM cerebrales, el medio de contraste usado es el gadolinio y para su aplicación se prefieren las imágenes potenciadas en T1, en las que se ven de alta intensidad las lesiones patológicas que se realizan tras su administración [48].

Capítulo 4

Estado del arte de la segmentación automática de imágenes médicas

4.1 Segmentación y análisis de imágenes médicas

Tan pronto como fue posible obtener y guardar imágenes médicas en un equipo de cómputo, se han desarrollado sistemas para el análisis automático de las mismas. Desde la década de los setenta el análisis de imágenes médicas era realizado mediante la aplicación secuencial de un modelo matemático de procesamiento de píxeles de bajo nivel para construir componentes basados en reglas y así resolver ciertas tareas en concreto. A finales de los años noventa los métodos supervisados, entrenados para desarrollar sistemas de aprendizaje automático, fueron ganando cada vez más popularidad entre los métodos de

análisis de imagen médica.

Al día de hoy las técnicas de aprendizaje automático para el reconocimiento de patrones constituyen la base de muchos de los sistemas de análisis de imágenes médicas más exitosos. Por lo cual, se ha observado la evolución de sistemas diseñados en su totalidad por humanos, hacia sistemas entrenados por computadoras empleando datos de muestra para la extracción de características de la imagen. Aunque en la mayoría de los casos el etiquetado de las muestras es realizado por los investigadores, los sistemas avanzan rápidamente hacia el aprendizaje de las características directamente de los datos. [51].

4.2 Aprendizaje profundo en análisis de imágenes médicas

Como es sabido, hoy en día la aplicación de técnicas de aprendizaje profundo en diferentes ámbitos está en continuo crecimiento y actualización, más concretamente en estudios donde se trata de análisis de imágenes [51], arquitecturas como las redes neuronales convolucionales o recurrentes han abierto un amplio campo de estudio para el desarrollo de distintas técnicas de aplicación entre las cuales se encuentran:

- Clasificación, tanto de imágenes completas como de partes aisladas dentro de estas. El primer tipo corresponde a la clasificación de una imagen como positivo o negativo en una enfermedad. En este caso, las redes neuronales convolucionales (CNN por sus siglas en inglés *Convolutional*

Neural Network) son las técnicas empleadas actualmente para este fin, ya que en ciertas tareas llegan a superar la precisión de los expertos. El segundo tipo corresponde a la clasificación de una parte de la imagen en dos o más clases, para lo cual son muy importantes tanto la información local de la lesión como el contexto a su alrededor.

- Detección de órganos o lesiones, más centrado en la localización, tanto espacial como temporal, esta es una parte fundamental del diagnóstico y es de las más complejas para los especialistas. Se ha trabajado mucho en los sistemas de detección por computadora que son diseñados para detectar lesiones automáticamente mejorando la precisión de la detección así como el tiempo invertido por los expertos, como ejemplo se encuentra la FracNet [52] que es una CNN diseñada para la detección de fracturas y fisuras en estructuras óseas a partir del análisis de imágenes de radiografía.
- Segmentación de órganos y subestructuras, permitiendo la caracterización de parámetros relacionados con la forma o el volumen en análisis abdominales, cardíacos o cerebrales, entre otros. Esta segmentación también se puede aplicar al caso de lesiones, combinando los retos de la detección de objetos, así como la segmentación de órganos y subestructuras aplicando algoritmos de aprendizaje profundo. Actualmente las tareas de segmentación es un tópico relevante en artículos sobre la aplicación del aprendizaje profundo en imágenes médicas, por lo que parece lógico observar una gran variedad de metodologías estudiadas, incluyendo el desarrollo de arquitecturas CNN únicas, la aplicación de redes neuronales recurrentes (RNN por sus siglas en inglés *Recurrent Neural Network*) y la implementación de modelos basados en transformadores de visión (ViT por sus siglas en inglés *Vision Transformer*) que cada vez toman

más relevancia en tareas de segmentación automática.

- Registro o alineamiento espacial de imágenes médicas, una tarea muy común en el análisis de imágenes que se lleva a cabo para que todas las imágenes que conforman un conjunto de datos estén encuadradas respecto a la misma referencia. Esta técnica es empleada sobre todo cuando las imágenes pueden desalinearse debido a movimientos de los pacientes, o a movimientos involuntarios de órganos, como los pulmones o el corazón.
- Búsqueda automática de imágenes dentro de una base de datos masiva. Este proceso se realiza con la intención de identificar casos similares, enfermedades raras o incluso para mejorar los cuidados de los pacientes. Otras aplicaciones son la de generación y mejora de imágenes a partir de arquitecturas CNN, cada vez aplicado a más tareas de deep learning, o la de elaboración de informes médicos a partir de una imagen.

Las CNN se encuentran entre las redes neuronales profundas que han demostrado un rendimiento destacado en una amplia variedad de tareas de visión por computadora, como clasificación de imágenes, detección de objetos, segmentación semántica, entre otras. Las CNN funcionan mediante la aplicación de filtros convolucionales para extraer características locales de las imágenes, seguidos de operaciones de *pooling* para reducir la dimensionalidad. Las capas completamente conectadas se utilizan para combinar las características extraídas y realizar tareas de clasificación o regresión. La retropropagación y la optimización se utilizan para ajustar los pesos de la red durante el entrenamiento. Las CNN han demostrado un gran éxito en el procesamiento de imágenes y han sido fundamentales en el avance de la visión por computadora.

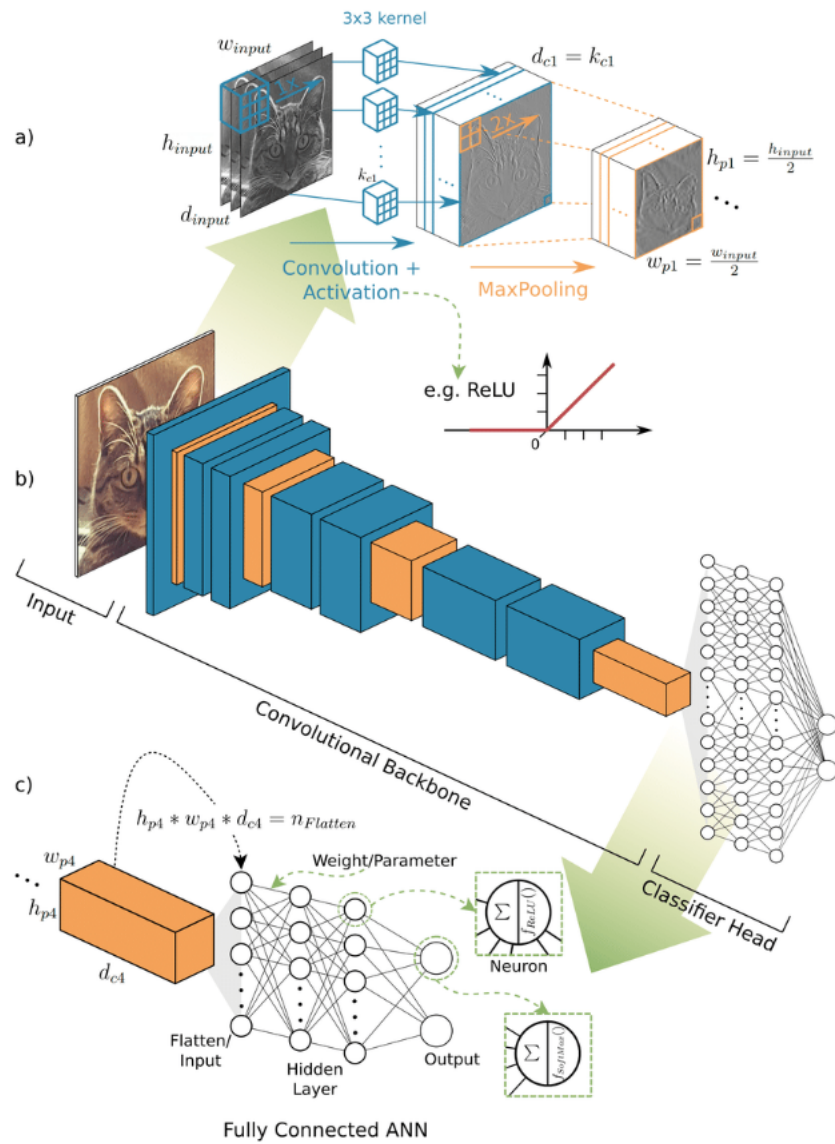


Figura 4.1: Descripción general y detalles de una arquitectura CNN para el reconocimiento de imágenes. *a)* Expande una entrada RGB de tres canales, convolución + función de activación. *b)* A partir de una imagen de entrada, se crean los mapas de características por convolución y se redimensionan mediante operaciones de agrupación máxima que reducen la resolución pero los mapas de características son más profundos hasta que llegan a un clasificador. *c)* Clasificación realizada con los pesos entrenados de la capa convolucional así como de la capa completamente conectada [53] (Imagen: “Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends”[53]).

4.3 Algoritmos de aprendizaje profundo para segmentación de imágenes médicas

Para el estudio de imágenes médicas ha surgido una gran necesidad de incluir nuevas técnicas para mejorar el seguimiento del paciente, buscando un diagnóstico temprano que atenda al pronóstico y que pueda monitorear la respuesta al tratamiento. Además, dados los avances en las tecnologías médicas utilizadas hoy en día en la práctica diaria, actualmente se dispone de multitud de parámetros y variables obtenidos directa e indirectamente del paciente los cuales en algunos casos son desaprovechados por la falta de herramientas adecuadas para la explotación de este tipo de información. En este contexto toma relevancia la implementación de técnicas de aprendizaje automático, de las cuales el aprendizaje profundo para el desarrollo de mejores herramientas que auxilien a los especialistas de la salud en la elaboración de diagnósticos cada vez más precisos.

El aprendizaje automático es la rama de la inteligencia artificial (IA) que comprende multitud de técnicas que permitirán a un algoritmo aprender características de su entorno y proporcionar soluciones ante un determinado problema, por ello se puede definir como el proceso de optimización de un criterio de desempeño usando datos o experiencia pasada, para finalmente obtener un modelo que pueda ser utilizado para solventar la tarea para la cual fue diseñado [54].

La mayoría de los algoritmos de aprendizaje automático requieren de una fase previa de extracción de características con las cuales entrena el modelo, este hecho convierte a los modelos extremadamente dependientes de esta fase

por lo que es necesario un buen procesamiento previo de los datos. Del mismo modo, existe un tipo de métodos que no requieren de esta etapa previa, los cuales reciben el nombre de aprendizaje por representación, donde el algoritmo utiliza las mejores características para clasificar los datos proporcionados [55]. De esta manera, el aprendizaje profundo es una especialización del llamado aprendizaje por representación, compuesto principalmente por el uso de redes neuronales diseñadas con multitud de capas junto con una serie de filtros convolucionales con parámetros libres. Logrando que el algoritmo sea capaz de aprender el conjunto de características que reflejan una jerarquía de estructuras en los datos que permiten expresar representaciones complejas mediante representaciones más simples.

4.3.1 Herramientas de software para segmentación automática de imágenes médicas

La implementación de algoritmos de aprendizaje automático en herramientas de segmentación ha experimentado un crecimiento significativo en los últimos años, lo cual ha generado que exista una gran variedad de programas y herramientas de software para realizar tareas de la segmentación automática de imágenes. Estas herramientas ayudan a los profesionales de la salud para identificar y delinear estructuras de interés en imágenes médicas. Algunos de los principales software y frameworks que se utilizan son:

3D Slicer y 3D SlicerAI

3D Slicer es una plataforma de código abierto ampliamente utilizada en radiología y medicina para la visualización y el procesamiento de imágenes médicas en tres dimensiones. Ofrece una variedad de herramientas para la segmentación, el análisis y la visualización de estructuras en imágenes médicas, como tomografías computarizadas y resonancias magnéticas. La extensión 3D SlicerAI integra capacidades de inteligencia artificial para la segmentación automática de imágenes médicas, permite a los usuarios utilizar algoritmos de aprendizaje automático para ayudar en tareas de segmentación, como la identificación de estructuras de interés en imágenes médicas de manera más eficiente y precisa. Ambas herramientas son útiles en la investigación y la práctica clínica para el análisis de imágenes médicas avanzadas [56].

ITK-SNAP

ITK-SNAP es una aplicación de código abierto utilizada en radiología y medicina para la segmentación y visualización de imágenes médicas. Permite a los profesionales médicos delinear y etiquetar estructuras de interés en imágenes de TC y RM, y también ofrece herramientas para el análisis cuantitativo de imágenes médicas [57].

Freesurfer

Freesurfer es una suite de software de procesamiento de imágenes médicas utilizada principalmente en la investigación neurocientífica. Está especializada en la segmentación y análisis de la corteza cerebral a partir de imágenes de resonancia magnética estructural. Freesurfer proporciona mediciones detalladas de la anatomía cerebral, lo que es valioso en estudios de neuroimagen y en la investigación de enfermedades neurológicas [58].

OsiriX

OsiriX es una aplicación de código abierto utilizada en el campo de la radiología para la visualización y el procesamiento de imágenes médicas. Su principal función es permitir a los profesionales médicos y radiólogos visualizar, analizar y diagnosticar a partir de imágenes médicas en un entorno de visualización avanzado, proporcionando una amplia gama de funcionalidades que mejoran la precisión y la eficiencia en la interpretación de imágenes [59].

BrainSuite

Es un conjunto de software de código abierto diseñado principalmente para el procesamiento y análisis de imágenes cerebrales, con un enfoque especial en IRM cerebral. Su objetivo principal es la segmentación y el análisis detallado de las estructuras cerebrales, lo que lo hace útil en la investigación neurocientífica y en aplicaciones clínicas. Es una herramienta para la segmentación, el

análisis y la visualización de imágenes cerebrales, lo que lo convierte en una herramienta esencial en la práctica clínica. Su capacidad para crear modelos de tres dimensiones (3D) del cerebro y realizar mediciones precisas lo hace especialmente útil en el estudio de la anatomía cerebral y el diagnóstico de trastornos neurológicos [60].

MONAI

MONAI (Medical Open Network for AI) es un framework de código abierto diseñado específicamente para aplicaciones de inteligencia artificial en imágenes médicas. Proporciona un conjunto de herramientas y bibliotecas de Python que facilitan el desarrollo, entrenamiento y despliegue de modelos de IA en el campo de la radiología y otras disciplinas médicas. Es una herramienta esencial para investigadores que trabajan en el desarrollo de aplicaciones de inteligencia artificial para el análisis de imágenes médicas, facilita la implementación de algoritmos de aprendizaje profundo y el procesamiento de datos médicos, lo que acelera la investigación y el desarrollo en el campo de la medicina basada en datos y la radiología computarizada [61].

DeepMedic

DeepMedic es una herramienta de procesamiento de imágenes médicas basada en inteligencia artificial con un enfoque particular en aplicaciones de imágenes cerebrales. Su principal objetivo es identificar y segmentar con precisión estructuras de interés en imágenes médicas, como tumores, lesiones o regiones anatómicas específicas. DeepMedic utiliza arquitecturas de redes neuronales

convolucionales profundas para realizar la segmentación automática de imágenes. Estas CNN están diseñadas para aprender patrones complejos en los datos de imagen y mejorar la precisión de la segmentación. Su capacidad para identificar y segmentar estructuras con precisión en imágenes 3D lo convierte en una herramienta valiosa en la investigación médica y en la práctica clínica para el diagnóstico y el tratamiento de enfermedades neurológicas [62].

El *software* de segmentación automática de imágenes médicas es una herramienta esencial que mejora la precisión, la eficiencia y la calidad de la atención médica, debido a esto la elección del *software* especializado dependerá de factores como, la experiencia en programación, los recursos disponibles, la calidad y la cantidad de datos, la aplicación clínica específica y la constante actualización sobre investigaciones y herramientas disponibles en el área de la salud y la computación. Esto con la finalidad ofrecer a los profesionales de la salud herramientas adicionales para realizar diagnósticos más precisos, planificar tratamientos personalizados y contribuir a la investigación médica, lo que finalmente beneficia a los pacientes y al avance de la investigación científica.

Capítulo 5

Base de datos y arquitectura Swin UNETR

En este capítulo se hará una descripción de los conjuntos de datos disponibles de imágenes médicas y del conjunto de datos empleado para el entrenamiento. Así mismo se hace una descripción del principio de funcionamiento de la arquitectura Swin UNETR utilizada para el desarrollo de este proyecto.

5.1 Repositorios de imágenes médicas cerebrales de libre acceso

El Proyecto Conectoma Humano (PCH) [63] tiene como objetivo proporcionar una compilación de datos neuronales, una interfaz para navegar gráficamente

estos datos y la oportunidad de lograr conclusiones nunca antes vistas sobre el cerebro humano vivo. Los datos de neuroimagen anatómica y de difusión del PCH están disponibles abiertamente para la comunidad científica para su examen y exploración. Estos incluyen imágenes cerebrales y volúmenes de resultados obtenidos del avanzado sistema de imágenes Siemens 3T Conectoma provenientes del Hospital General de Massachusetts [63].

La recopilación y el intercambio de datos para este proyecto estuvo a cargo del Proyecto conectoma humano. La financiación del proyecto fue proporcionada por diversos institutos de Estados Unidos, el Instituto Nacional de Investigación Dental y Craneofacial, el Instituto Nacional de Salud Mental y el Instituto Nacional de Trastornos Neurológicos y Accidentes Cerebrovasculares. Los datos de PCH son difundidos por el Laboratorio de Neuroimagen de la Universidad del Sur de California [63].

El ISMRM 2015 *Tractography Challenge* [64] buscaba evaluar y comparar las conexiones de tractografía de última generación. La tarea del desafío era proporcionar la mejor reconstrucción posible de las vías de las fibras en una réplica simulada de forma realista de una imagen de RM ponderada por difusión de todo el cerebro con propiedades típicas de una adquisición clínica. Se esperaba que los concursantes procesaran el conjunto de datos del desafío utilizando su mejor modelo de procesamiento, pidiendo que procesaran los datos de difusión en imagen de resonancia magnética (dIRM) de un paciente clínico para producir el mejor resultado de tractografía posible. El reto ya está cerrado, sin embargo, los datos continúan disponibles para descarga [64].

Adicionalmente repositorios como OASIS Brains (*Open Access Series of Images Studiens*), TCIA (*The Cancer Image Archive*) o ADNI (*Alzheimer's*

Disease Neuroimaging Initiative) son algunas otras bases de datos con imagen médica disponibles para realizar consultas de investigación.

5.1.1 Conjunto de datos BraTS 2023

Para la realización de este proyecto de tesis de Maestría, se ha empleado el conjunto de datos de imagen médica proporcionada por el *Data Challenge BraTS 2023 (Brain Tumor Segmentation)* [65][66]. Para poder obtener acceso a los datos es necesario registrarse en el portal *Synapse*[67], plataforma de código abierto para el análisis colaborativo de datos científicos. Puede almacenar datos, código, resultados y descripciones de trabajos de investigación, indicando el propósito para el cual desea solicitar emplearse el conjunto de datos.

El desafío de la segmentación internacional de tumores cerebrales (BraTS). BraTS, desde 2012, se ha centrado en la generación de un entorno de evaluación comparativa y un conjunto de datos para la delimitación de gliomas cerebrales en adultos. El objetivo del desafío de este año sigue siendo la generación de un entorno de evaluación comparativa común, pero su conjunto de datos se ha ampliado sustancialmente a 4500 casos para abordar poblaciones adicionales, tumores de meningioma, preocupaciones clínicas como datos faltantes y consideraciones técnicas de aumento de imagen. Específicamente, el enfoque de BraTS 2023 es identificar los algoritmos de última generación actuales para abordar la misma población de gliomas adultos que en el desafío RSNA-ANSR-MICCAI BraTS, así como la Población de pacientes con glioma cerebral subsahariano desatendidos, meningioma intracraneal, metástasis cerebral, pacientes con tumores cerebrales pediátricos, datos faltantes globales y locales,

útiles técnicas de aumento de datos [65] [67].

5.1.2 Descripción de datos de entrenamiento

El conjunto de datos de entrenamiento se compone de 1251 archivos, cada uno de los cuales es un estudio de RM multimodal de un paciente distinto. El conjunto de datos consta de exploraciones de imágenes por resonancia magnética multiparamétrica (mpMRI) de sitios múltiples adquiridas clínicamente de rutina de pacientes con tumores cerebrales, con diagnóstico patológicamente confirmado y estado de metilación del promotor del gen MGMT disponible. Cada archivo corresponde al estudio de RM de un paciente distinto, donde todas las exploraciones BraTS mpIRM están disponibles como archivos NIfTI (.nii.gz) y constan de las siguientes secuencias capturadas por cada uno de los archivos.

Los datos de entrenamiento y validación son los mismos que se utilizaron en *BraTS'21* [68]; sin embargo, el conjunto de datos de prueba publicado en el desafío *BraTS'23* se ha actualizado con muchas más exploraciones de mpMRI.

- Secuencias ponderadas en T1.
- Secuencias ponderadas en T1 tras la administración de un medio contraste (T1ce).
- Secuencias ponderadas en T2.
- Secuencias FLAIR, que son una prolongación de las T2, con supresión del líquido cefalorraquídeo.

- Imágenes segmentadas, en las cuales se muestra el tumor diferenciando las tres partes más relevantes: el tumor realzado, el edema peritumoral, y el núcleo del tumor necrótico.

El conjunto de cortes de cada caso generan una imagen en volumen de la que se puede elegir el plano anatómico para su visualización. Por la naturaleza del modelo de aprendizaje profundo implementado en este proyecto, el análisis de las muestras durante el proceso de entrenamiento se realiza con volúmenes aleatoriamente recortados del conjunto original, y se trabaja con los tres planos anatómicos durante todo el entrenamiento. Todas las imágenes que constituyen el conjunto de datos disponible para entrenamiento, validación y prueba, cuentan con un archivo de segmentaciones reales de las subregiones tumorales, creadas y aprobadas por neurorradiólogos expertos. Adicionalmente, todas las imágenes han sido proporcionadas con supresión del cráneo, de modo que no ha sido necesaria su eliminación en el preprocesamiento (Figura 5.1).

5.1.3 Segmentación de tumor

Las regiones segmentadas consideradas son las siguientes: 1) el tumor resaltado (ET por sus siglas en inglés *Enhancing Tumor*), 2) el núcleo del tumor (TC del inglés *Tumor Core*) y 3) el tumor completo (WT del inglés *Whole Tumor*) (Figura 5.2). El ET lo conforman las áreas que se muestran hiperintensas, o de alta intensidad en la señal, en secuencias T1ce al compararlas con la T1, pero también al compararlo con materia blanca “sana” en la T1ce. El TC representa el tumor completo que engloba el ET, así como la parte necrótica (con fluido) y las partes no resaltadas (sólidas). Estas partes suelen mostrarse hipointensas, lo que quiere decir que muestran una intensidad de señal menor

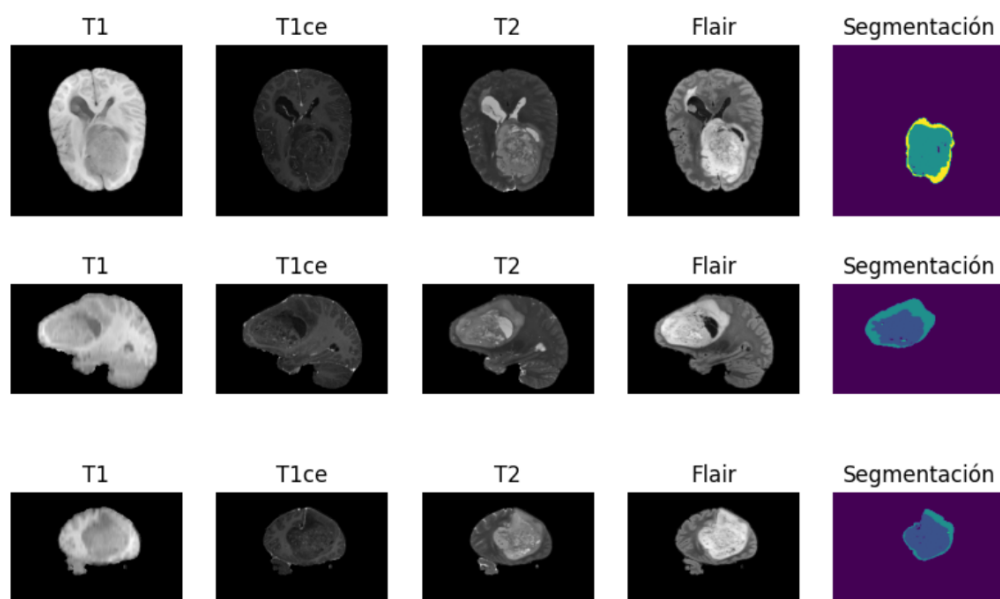


Figura 5.1: Estudio de un paciente con diagnóstico confirmado de GBM, donde se presentan las diferentes secuencias capturadas de RM y la segmentación del área donde se encuentra el tumor. Cada fila representa un plano anatómico distinto (axial, sagital y coronal, de arriba hacia abajo respectivamente). Las columnas muestran cada una de las secuencias capturadas (T1, T1ce, T2, Flair) adicionando en la última columna la segmentación de las áreas donde está presente el tejido tumoral. Conjunto de datos procedentes de BraTS 2023 [68].

que la considerada de referencia para el tejido en la T1ce, en comparación con la T1. Por último, el WT describe la lesión completamente englobando las partes descritas anteriormente junto con el edema peritumoral (ED), el cual suele mostrarse hiperintenso en las secuencias FLAIR [68].

5.1.4 Formato de imágenes

El formato de las imágenes proporcionadas por BraTS es NIfTI (.nii). NIfTI (*Neuroimaging Informatics Technology Initiative*) es una iniciativa impulsada por el Instituto Nacional de la Enfermedad Mental y el Instituto Nacional

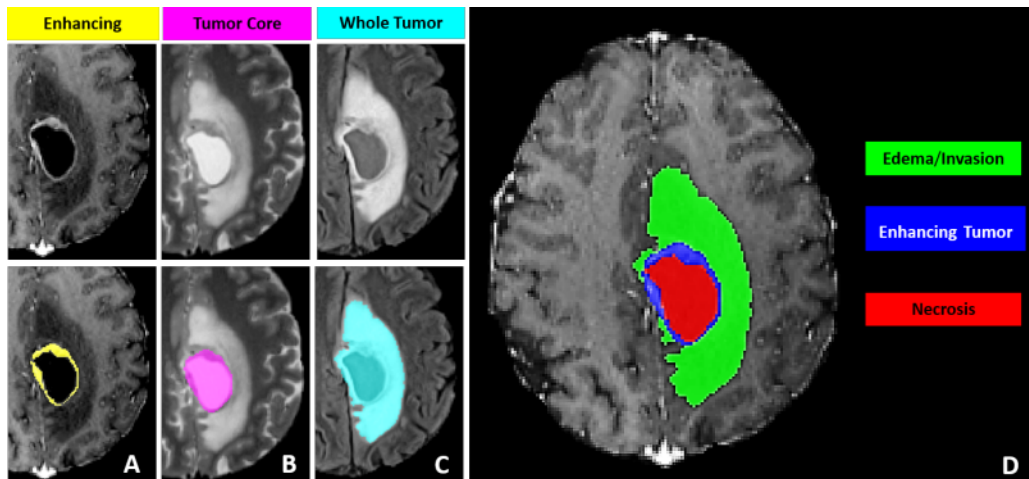


Figura 5.2: Segmentación manual de una imagen de RM realizada por los expertos en anotación. Los paneles de imágenes *A* – *C* indican las regiones consideradas de la segmentación: *A*) el tumor realzado (ET - amarillo), *B*) el núcleo del tumor (TC - magenta) y *C*) todo el completo (WT - cian). *D*) muestra las segmentaciones combinadas que generan las etiquetas finales de las subregiones tumorales, núcleo mejorado (amarillo), núcleo necrótico/quístico (rojo) y edema/invasión (verde) (Imagen: The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification [68]).

de Desórdenes Neurológicos y Derrame Cerebral de Estados Unidos, y cuyo principal objetivo es ofrecer los servicios necesarios para desarrollar herramientas informáticas relacionadas con neuroimagen [69]. Este formato permite almacenar tanto los datos de una imagen del cerebro en 3 dimensiones, como metadatos sobre la misma (tipo de dato, tamaño del archivo, dimensiones, número de canales). La ventaja del formato .nii respecto a otros, como DICOM, es que los metadatos no incluyen información sobre el paciente, lo cual es muy importante para preservar su privacidad.

5.2 Arquitectura del modelo

Swin UNet Transformers (Swin UNETR)

El modelo Swin UNETR es un tipo de red neuronal ViT (ViT por sus siglas en inglés *Vision Transformers*) que combina la eficiencia de extracción de características de Swin Transformer con la capacidad de segmentación de U-Net. Esto permite un rendimiento superior en tareas de segmentación de imágenes, especialmente en imágenes de alta resolución, donde la arquitectura Swin Transformer destaca en la extracción de características y U-Net se destaca en la segmentación precisa.

A continuación se comentan las bases de este tipo de modelos ViT y las arquitecturas en las que se basa el funcionamiento del modelo Swin UNETR.

5.2.1 Transformador de Visión ViT

Los modelos de transformación de visión o ViT surgieron como una alternativa prometedora a las CNN para el procesamiento de imágenes. Si bien las CNN han sido dominantes en el campo de la visión por computadora, las ViT han demostrado un rendimiento notable en diversas tareas y han logrado resultados competitivos [70].

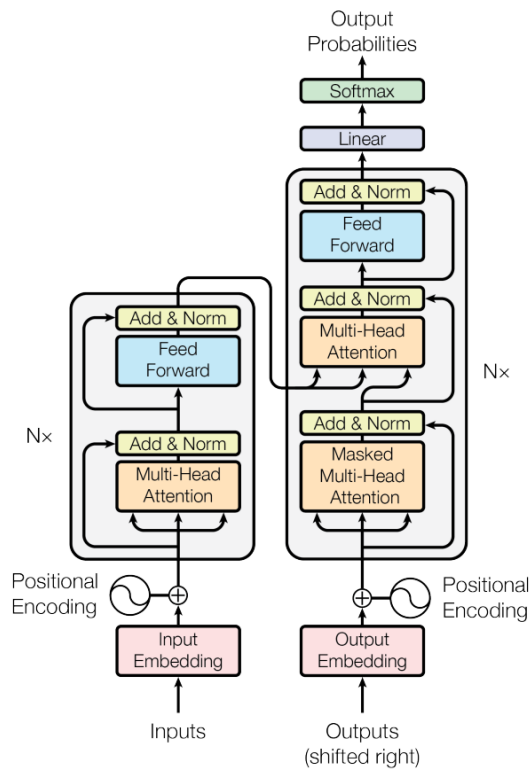


Figura 5.3: El Transformador sigue esta arquitectura general utilizando autoatención apilada y capas puntuales y completamente conectadas tanto para el codificador como para el decodificador, como se muestra en las mitades izquierda y derecha de la figura, respectivamente [71] (Imagen: Advances in neural information processing systems [71]).

La arquitectura original ViT utiliza transformadores (Figura 5.3) para el procesamiento de imágenes sin convoluciones, que se basan en mecanismos de atención, eliminando la necesidad de capas convolucionales. El funcionamiento de las arquitecturas basadas en ViT se puede resumir en las siguientes etapas: preprocesamiento de datos, codificación con atención, codificación posicional, transformaciones lineales, clasificación o regresión [72][71].

- La primera etapa, preprocesamiento de datos, las imágenes de entrada se dividen en parches de tamaño fijo y cada parche se aplanan en un vector y se aplica una transformación lineal para obtener un vector embebido. En la etapa de codificación con atención los vectores embebidos se utilizan como entradas para las capas de atención que calculan la importancia relativa de los vectores embebidos y ajusta sus pesos en función de la relación entre ellos [72].
- La atención permite que la red se enfoque en partes específicas de la imagen y capture características relevantes en diferentes regiones [71].
- Posteriormente en la codificación posicional, se agrega información posicional a los vectores embebidos para preservar la relación espacial entre los parches. Esto permite que la red comprenda la disposición y estructura espacial de los elementos en la imagen. Se continúa con la etapa de transformaciones no lineales realizando que cada capa de atención, se aplican transformaciones no lineales, como capas de activación (por ejemplo, ReLU, Tanh, sigmoide, MSE), para introducir no linealidad en el modelo y capturar relaciones más complejas en los datos [72].
- Por último la salida de la última capa de atención se utiliza para realizar tareas específicas, como clasificación o regresión. Para la clasificación, se puede usar una capa densa seguida de una función de activación softmax para asignar probabilidades a diferentes clases. Para la regresión, se puede utilizar una capa densa con una función de activación lineal para predecir valores continuos [72].

Dada la capacidad de procesamiento y almacenamiento que pueden alcanzar los modelos ViT para realizar tareas de segmentación de imágenes médicas,

se ha tomado la arquitectura del modelo Swin UNETR como base en la realización de este proyecto. La arquitectura SwinUNETR combina dos modelos principales: Swin Transformer y UNETR. A continuación se describen brevemente cada uno de ellos.

Swin Transformer

Swin Transformer es un tipo de ViT que construye mapas de características jerárquicas y tiene una complejidad computacional lineal al tamaño de la imagen. Como se ilustra en la Figura 5.4, el modelo Swin Transformer construye una representación jerárquica partiendo de parches de pequeño tamaño (perfilados en gris) y fusionando gradualmente los parches vecinos en capas Transformer más profundas. Con estos mapas de características jerárquicas, el modelo Swin Transformer puede aprovechar convenientemente técnicas avanzadas de predicción densa, como las redes piramidales de características (FPN) o U-Net.

La complejidad computacional lineal se consigue calculando la autoatención localmente dentro de ventanas no solapadas que dividen una imagen (perfiladas en rojo). El número de parches en cada ventana es fijo, por lo que la complejidad es lineal con el tamaño de la imagen. Estos méritos hacen que Swin Transformer sea adecuado como columna vertebral de propósito general para diversas tareas de visión, a diferencia de las arquitecturas anteriores basadas en ViT que producen mapas de características de una única resolución y tienen una complejidad cuadrática [73].

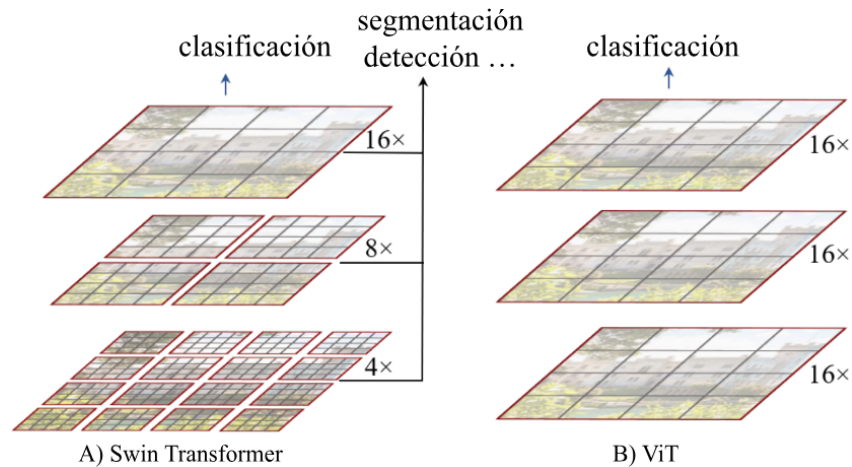


Figura 5.4: A) Swin Transformer es un tipo de ViT que crea mapas de características jerárquicas fusionando parches de imágenes en capas más profundas y tiene una complejidad de cálculo lineal debido al cálculo de la autoatención solo dentro de cada ventana local. B) Los modelos ViT producen mapas de características de una única resolución baja y tienen una complejidad de cálculo cuadrático (Imagen: Swin transformer: Hierarchical vision transformer using shifted windows [73]).

UNETR

UNETR, o UNet Transformer, es una arquitectura basada en ViT para la segmentación de imágenes médicas, que reformula la tarea de segmentación 3D como un problema de predicción secuencia a secuencia 1D y utiliza un transformador como codificador para aprender información contextual de los parches de entrada incrustados. Las representaciones extraídas del transformador codificador se fusionan con el decodificador basado en CNN a través de conexiones de salto en múltiples para predecir los resultados de la segmentación. En lugar de transformadores en el decodificador, UNETR utiliza un decodificador basado en CNN. Esto se debe a que los transformadores son incapaces de captar adecuadamente la información localizada, a pesar de su gran capacidad para aprender información global [74] (Figura 5.5).

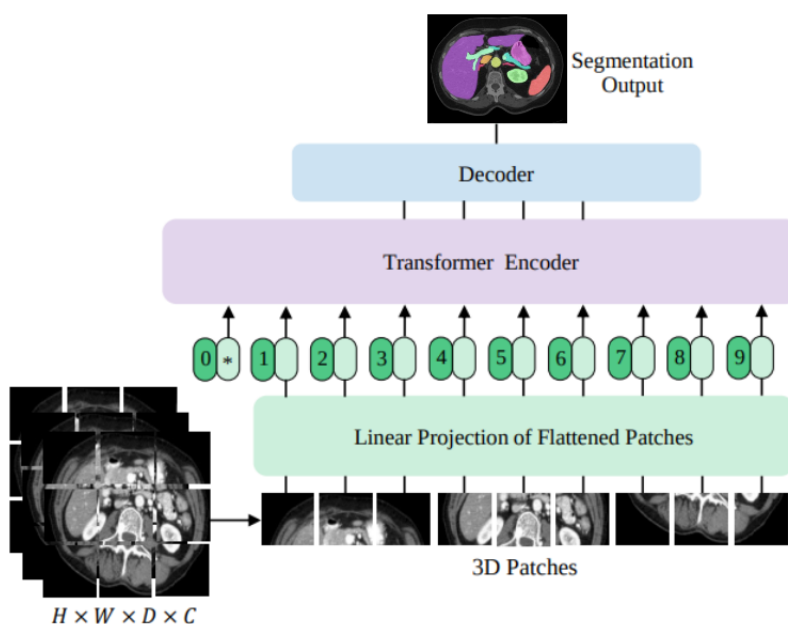


Figura 5.5: El modelo UNETR consiste de un transformador codificador que utiliza directamente parches 3D y está conectado a un decodificador basado en CNN mediante una conexión de salto (Imagen: Unetr: Transformers for 3d medical image segmentation [74]).

5.2.2 Swin UNETR

En el modelo Swin UNETR Transformers, o Swin UNETR, la tarea de segmentación semántica de tumores cerebrales 3D se reformula como un problema de predicción de secuencia a secuencia en el que los datos de entrada multimodales se proyectan en una secuencia 1D de incrustación y se utilizan como entrada a un transformador Swin jerárquico como codificador. El codificador Swin Transformer extrae características en cinco resoluciones diferentes mediante el uso de ventanas desplazadas para calcular la autoatención y está conectado a un decodificador basado en una red neuronal totalmente convolucional (FCNN por sus siglas en inglés *Fully Convolutional Neural Network*) en cada resolución mediante conexiones de salto [70] 5.6.

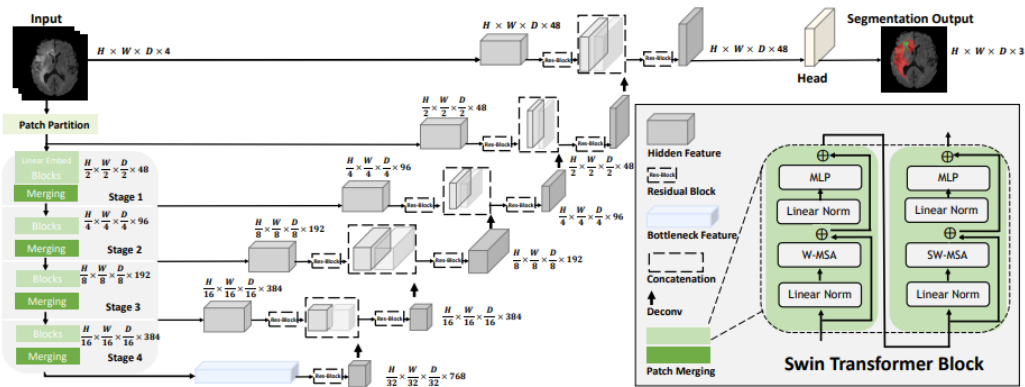


Figura 5.6: Descripción general de la arquitectura Swin UNETR. La entrada del modelo son IRM multimodal en 3D con 4 canales. El Swin UNETR crea parches sin traslape de los datos de entrada y utiliza una capa de partición de parches para crear ventanas con el tamaño deseado para calcular la autoatención. Las representaciones de características codificadas en Swin Transformer se alimentan a un decodificador CNN a través de una conexión de salto en múltiples resoluciones. La salida de segmentación final consta de 3 canales de salida correspondientes a las subregiones tumor completo (WT), tumor realzado (ET) y núcleo tumoral (TC) [70](Imagen:Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images [70]).

El modelo Swin UNETR se ha entrenado específicamente para tareas de segmentación de imágenes médicas y ha demostrado un alto rendimiento en comparación con otros enfoques de ViT y CNN, debido a su capacidad para capturar relaciones espaciales en imágenes de estudios de TC ó RM, y su adaptabilidad a una variedad de tipos de imágenes y patologías, lo que es esencial en aplicaciones de diagnóstico médico y de planificación del tratamiento.

Capítulo 6

Metodología

En este capítulo se describen brevemente las herramientas de software y bibliotecas empleadas para desarrollar el proyecto, desde la carga de datos proveniente del conjunto de datos como el modelo de segmentación y su implementación.

6.1 Herramientas de software

6.1.1 Bibliotecas de Python

Python es uno de los lenguajes de programación más populares y ampliamente utilizado para desarrollar modelos de aprendizaje profundo. Hay varias bibliotecas y marcos de trabajo en Python que facilitan la implementación y el entrenamiento de redes neuronales profundas.

MONAI

MONAI es un marco de trabajo de código abierto diseñado específicamente para el procesamiento de imágenes médicas y el desarrollo de aplicaciones de IA en el ámbito de la salud. Proporciona herramientas y utilidades para el preprocesamiento de datos, entrenamiento y evaluación de modelos de IA, así como la visualización y análisis de resultados en el dominio de la imagen médica [61].

PyTorch

Es otro marco de trabajo popular y de código abierto para Deep Learning. PyTorch ofrece una estructura dinámica y computacionalmente eficiente que permite un flujo de trabajo flexible y ágil en el desarrollo de modelos. PyTorch también proporciona herramientas para la investigación y experimentación con modelos de Deep Learning [75].

6.1.2 Visualization Toolkit (VTK)

VTK es un sistema de software de código abierto para el procesamiento de imágenes, gráficos 3D, renderizado de volumen y visualización. VTK incluye muchos algoritmos avanzados, por ejemplo reconstrucción de superficies, modelado implícito, destrucción, y técnicas de renderizado (renderizado de volumen acelerado por hardware, control LOD). VTK es utilizado por académicos para la enseñanza y la investigación; por instituciones gubernamentales

de investigación como Los Alamos National Lab en los EE. UU. o CINECA en Italia; y por muchas firmas comerciales que usan VTK para construir o ampliar productos [76].

6.1.3 Insight Toolkit (ITK)

ITK es una biblioteca multiplataforma de código abierto que proporciona a los desarrolladores un amplio conjunto de herramientas de software para el análisis de imágenes multidimensionales. Desarrollado a través de metodologías de programación extremas, ITK se basa en una arquitectura probada y especialmente orientada para el procesamiento, la segmentación y el registro de imágenes científicas en dos, tres o más dimensiones [77].

6.2 Preprocesamiento de imagen

Después de solicitar y obtener los datos que vamos a utilizar para el entrenamiento del modelo se realiza un preprocesamiento para ajustar sus características como tamaño, resolución ó tonalidad. Se inicia ajustando el número de casos para trabajar en el entrenamiento del modelo, ya que el repositorio de BraTS'23 contiene más de 1,251 casos de pacientes distintos, y considerando las características de hardware y el limitado poder de procesamiento disponible, se seleccionó un subconjunto de 280 casos distintos. Esta selección toma en cuenta que cada secuencia capturada en el estudio de RM almacena un promedio de 635 cortes, o imágenes, repartidos en 155 cortes del plano axial, 240 del plano coronal y 240 del plano sagital, considerando todas las secuencias

de captura de RM proporcionadas (T1,T1ce,T2, Flair) da un total de 2,540 imágenes por cada caso, y para el subconjunto seleccionado de 280 casos se tienen 711,200 imágenes distintas con sus máscaras, para realizar el entrenamiento, validación y evaluación del modelo. En la práctica no todas las IRM son utilizadas debido a que no todos los cortes capturan tejido tumoral, al aplicar las transformaciones recortan las imágenes vacías que no necesitamos del conjunto de entrenamiento, obteniendo 502,740 en total para implementar el modelo.

A continuación se describe el procedimiento seguido para la realización del proyecto.

1. Crear un archivo de texto plano en notación de objetos JavaScript (JSON) con los nombres de cada uno de los casos seleccionados de IRM, que contiene el nombre y la ruta de cada uno de los archivos según su secuencia de captura junto con el archivo de segmentación para cada caso.
2. Realizar una conmutación aleatoria de los nombres de los archivos utilizando la función *random.permutation()* de *numpy* para garantizar la aleatoriedad en la selección de casos de distinto grado de afección de GBM y crear los conjuntos de entrenamiento, el validación.
3. Cargar los conjuntos de entrenamiento y validación de los casos seleccionados en dos *DataLoaders* (herramienta de software que permite gestionar datos con inserciones, modificaciones, borrados y exportacione) con una resolución 128 x 128 x 128 con 4 canales NHWC (del inglés *Number, Height, Width, Channels*, que corresponde a el número de ejemplos, altura, ancho y número de canales).
4. Posteriormente se aplican las transformaciones *CropForegroundd*, *RandFlipd*,

RandSpatialCropd, *NormalizeIntensityd*, *RandScaleIntensityd* para el *DataLoader* de entrenamiento, y *NormalizeIntensityd* en el conjunto de validación. Estas transformaciones ayudan a estandarizar el tamaño y la escala de las imágenes, normalizar los valores de píxeles, aumentar la cantidad de datos de entrenamiento y mejorar la capacidad de generalización del modelo. Aplicar estas transformaciones de manera adecuada, se puede lograr mejorar el rendimiento del modelo durante en el entrenamiento.

5. Los *DataLoader* creados de validación y de entrenamiento eliminan la agrupación por paciente o caso clínico, de modo que todas las imágenes forman dos conjuntos globales.
6. Por último, las imágenes segmentadas, como se describe en la sección Capítulo 5.1.2, contienen cuatro etiquetas diferentes, 0, 1, 2 y 3 (0 fondo de la imagen, 1 núcleo del tumor, 2 tumor con realce, 3 edema); comúnmente se aplica una binarización para que los píxeles se agrupen en dos categorías 1 (tumor) o 0 (no tumor), pero en este caso generamos una salida multicategorica donde cada clase será una de las diferentes estructuras que conforman al tumor (edema, tumor realzado y núcleo del tumor) (Figura 6.1).

6.3 Implementación y entrenamiento del modelo Swin UNETR

Para la implementación de la arquitectura Swin UNETR se probaron distintos valores para los hiperparámetros, seleccionando los que obtuvieron mejor

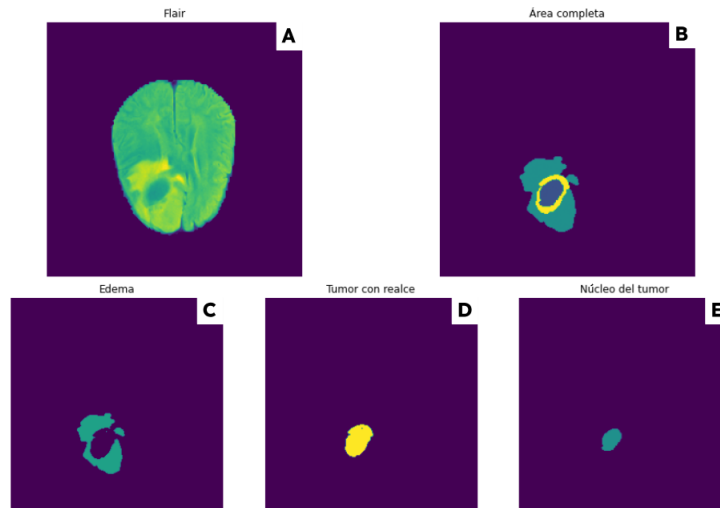


Figura 6.1: La imagen muestra un corte del estudio de RM de la secuencia FLAIR y la segmentación de las subestructuras que conforman el tumor. *A*) Pertenece a un corte de la secuencia FLAIR de un caso confirmado de GBM, *B*) muestra el área completa del tumor que contiene las tres subestructuras que lo componen, *C*) es el área segmentada que presenta edema peritumoral, *D*) área segmentada donde se detecta el tumor con realce y *E*) es la segmentación del núcleo del tumor. Conjunto de datos procedentes de BraTS'23 [68].

resultado en las métricas de exactitud y pérdida en los entrenamientos del modelo, estos valores están mostrados en la Tabla 6.1 donde la columna izquierda muestra los nombres de los hiperparámetros y en la columna derecha el valor asignado a cada uno.

El entrenamiento fue realizado durante 16 horas 23 minutos en un equipo de cómputo del Laboratorio de Bioinstrumentación del Instituto de Ciencias Aplicadas y Tecnología ICAT que cuenta con una tarjeta GPU NVIDIA Titan RTX de 24 GB, 48 GB de RAM y procesador Intel Core i7 de décima generación a 3.80GHz. Cabe resaltar que este tipo de arquitecturas de ViT son implementadas necesariamente en una infraestructura lo suficientemente robusta debido a que se requiere gran capacidad de procesamiento y memoria RAM. La ejecución de arquitecturas de Transformadores en este tipo de

Hiperparámetros de Entrenamiento	
Hiperparámetro	Valor
Transformaciones entrenamiento	RandFlipd RandSpatialCropd NormalizeIntensityd RandScaleIntensityd CropForegroundd
Transformaciones validación	NormalizeIntensityd Spacingd
Dimensión de imagen de entrada	(128, 128, 128)
Optimizador	AdamW
Tasa de aprendizaje	1e-4, 1e-5
Métricas de exactitud	Coefficiente DICE
Función de pérdida	Entropía cruzada binaria
Canales de entrada	4
Clases de salida	3
Épocas	300
Épocas de validación	30
Tamaño de batch	16

Tabla 6.1: Hiperparámetros y sus valores para entrenamiento del modelo Swin UNETR.

hardware aprovecha la aceleración de GPU y los recursos de supercomputación disponibles en las arquitecturas de Nvidia, lo que permite un procesamiento más rápido, una mayor eficiencia y una capacidad para manejar modelos y conjuntos de datos más grandes.

La partición total de los datos para los conjuntos de entrenamiento, validación y prueba quedó definida como se muestra en la tabla 6.2, que presenta el nombre del conjunto de datos en la primer columna, el número de imágenes en la segunda columna y por último el número de casos de RM seleccionados para cada conjunto de datos, al final de cada columna se muestra el total de imágenes y casos de RM utilizados.

Partición del Conjunto de Datos		
Conjunto	Imágenes	Casos de RM
Entrenamiento	358,256	200
Validación	90,396	50
Prueba	54,088	30
Total	502,740	280

Tabla 6.2: Partición de datos para entrenamiento, validación y prueba para implementación del modelo.

6.4 Validación de predicciones

Una vez completado el entrenamiento del modelo, se realizan las predicciones con el conjunto de prueba obteniendo la segmentación de cada una de estas. Para la salida obtenida en la predicción de cada una de las clases que comprenden las subregiones o subestructuras que conforman al tumor (tumor completo (WT), núcleo tumoral (TC) y tumor realzado (ET)), se busca la correspondencia con los mapas de segmentación de la imagen de entrada, y se han comparado con las segmentaciones reales para evaluar la calidad de la predicción.

6.4.1 Coeficiente DICE

Para realizar esta tarea de evaluación de las predicciones se ha utilizado el coeficiente DICE, también conocido como índice de similitud de DICE o coeficiente de Sørensen-Dice, es una métrica comúnmente utilizada para evaluar la similitud entre dos conjuntos de píxeles en problemas de segmentación de imágenes médicas. Es particularmente útil para evaluar la superposición o la concordancia entre una máscara de segmentación predicha por un modelo y

una máscara de segmentación de referencia o *ground truth* [78].

$$DICE = \frac{2 * (A \cap B)}{|A| + |B|} \quad (6.1)$$

De la ecuación 6.1 se obtiene que A y B son los dos conjuntos de píxeles que se comparan, por ejemplo A puede representar el conjunto de píxeles que corresponden a la segmentación realizada por el experto dentro de una imagen y B la predicción segmentada por modelo de aprendizaje profundo de la misma imagen. $(A+B)$ representa la cantidad de píxeles comunes o superpuestos entre A y B , mientras que $|A|$ y $|B|$ representan el número total de píxeles en los conjuntos A y B , respectivamente [79].

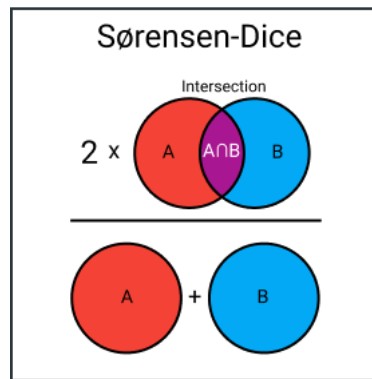


Figura 6.2: El índice de Sørensen-Dice mide la similitud y diversidad de conjuntos de muestras, porcentaje de superposición entre dos conjuntos, que es un valor entre 0 y 1 (Imagen: “9 Distance Measures in Data Science: The advantages and pitfalls of common distance measures” [79]).

El coeficiente varía de 0 a 1, donde 0 indica una falta total de superposición entre las dos máscaras de segmentación y 1 indica una superposición perfecta o una concordancia completa (Figura 6.2).

6.4.2 Distancia de Hausdorff

La distancia de Hausdorff es una métrica utilizada en el campo de la segmentación de imágenes para medir la discrepancia o la diferencia entre dos conjuntos de puntos. En el contexto de la segmentación de imágenes médicas, se utiliza para evaluar la similitud entre una máscara de segmentación producida por un algoritmo o modelo y una máscara de referencia o máscara de segmentación de expertos [80].

La distancia de Hausdorff se define como la distancia máxima entre un punto en uno de los conjuntos y su punto más cercano en el otro conjunto. Formalmente, si A y B son dos conjuntos de puntos en un espacio métrico, la distancia de Hausdorff entre A y B se calcula con la ecuación 6.2:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (6.2)$$

Donde $h(A, B)$ es la distancia de Hausdorff directa y se define como la distancia máxima de un punto en A a su punto más cercano en B . La distancia de Hausdorff inversa $h(B, A)$, se define como la distancia máxima de un punto en B a su punto más cercano en A . En el contexto de la segmentación de imágenes, A y B representan los conjuntos de puntos que corresponden a las regiones segmentadas en las máscaras de entrada y en las predicciones realizadas por el modelo de aprendizaje profundo, respectivamente. La distancia de Hausdorff proporciona una medida de la máxima discrepancia entre las dos máscaras y se utiliza para evaluar la precisión y la calidad de la segmentación [81].

Estas métricas son ampliamente utilizadas en la evaluación cuantitativa

de algoritmos de segmentación de imágenes médicas, ya que proporciona una medida de la similitud y discrepancia espacial entre las máscaras de segmentación y las predicciones de los modelos. Es especialmente útil en casos donde la precisión de la segmentación es crucial, como en aplicaciones médicas donde se segmentan estructuras anatómicas o regiones de interés específicas en imágenes médicas.

Capítulo 7

Resultados, reconstrucción y análisis

7.1 Resultados

En este Capítulo se presentan los resultados obtenidos del entrenamiento del modelo Swin UNETR y la segmentación de glioblastomas en IRM.

7.1.1 Introducción

El entrenamiento se realiza con los datos obtenidos del repositorio de BraTS'23 [68]. Este conjunto de datos está constituido de 250 casos de estudio de GBM de distinto grado que contienen un archivos para cada una de las distintas se-

cuencias capturadas de RM (T1, T1ce, T2, Flair) y la segmentación realizada por especialistas con de las subestructuras que componen el tumor (WT, TC, ET). Este subconjunto contiene 448,652 cortes de RM para el entrenamiento y validación del modelo utilizando las características de los hiperparámetros mencionados en el Capítulo 6.3, tras probar con diferentes parámetros de entrenamiento, se tomó como definitivo el modelo con las mejores métricas de exactitud y pérdida.

7.1.2 Resultados de entrenamiento

La etapa de entrenamiento se completó después de 16 horas y 23 minutos de ejecución, obteniendo como mejor índice de pérdida 36% y de exactitud 92%, posteriormente se grafica el progreso del entrenamiento donde se puede visualizar el descenso de la pérdida y el incremento de la exactitud a través de las 300 épocas (Figura 7.1), donde se puede observar un avance uniforme y constante por cada una de estas métricas obtenidas del proceso de validación durante el entrenamiento.

Del mismo modo, se generan los gráficos que muestran el incremento de la exactitud promedio de validación para cada una de las clases de salida que representan las subestructuras del tumor (WT, TC, ET)(Figura 7.2), observando el ascenso constante del coeficiente DICE promedio durante el proceso de entrenamiento. Estos gráficos muestran el comportamiento del modelo producto de la configuración de hiperparámetros comentada en le Capítulo 6.3, dichos valores se obtiene variando los hiperparámetros y entrenando repetidas ocasiones el modelo hasta alcanzar los mejores índices en las métricas de pérdida y exactitud.

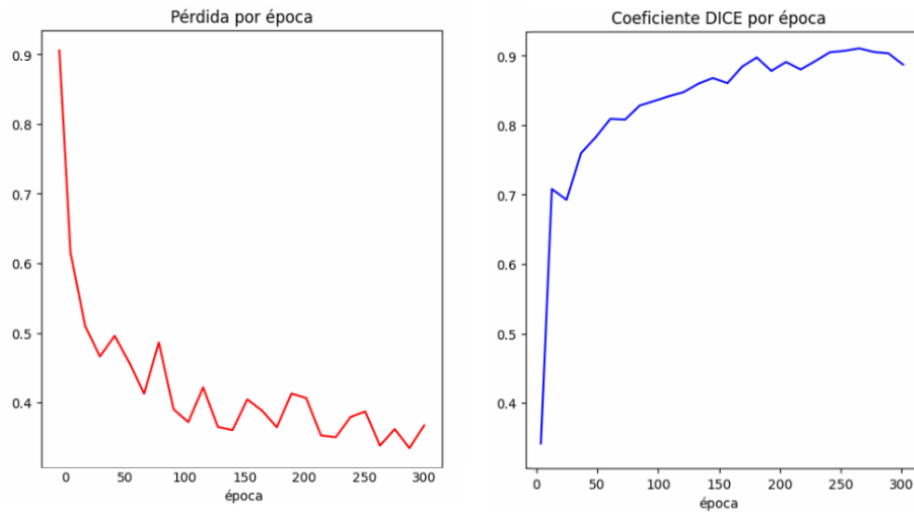


Figura 7.1: Gráficas que muestran el descenso de la pérdida y el incremento de la exactitud a través de las épocas de entrenamiento del modelo.

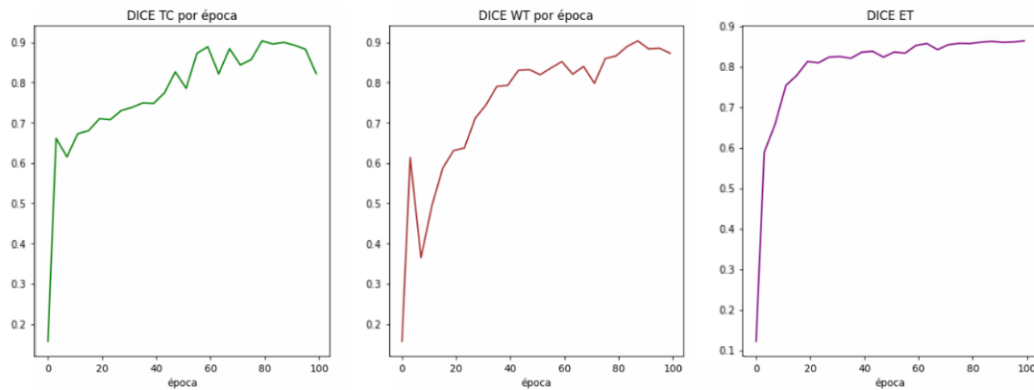


Figura 7.2: Gráficas que muestran el incremento de la exactitud promedio para cada una de las clases de salida que representan las subestructuras del tumor (WT, TC, ET) a lo largo de las épocas de entrenamiento del modelo.

7.1.3 Resultados de prueba

Una vez completado el entrenamiento se realizaron las evaluaciones con el conjunto de prueba que comprende 10 casos de RM ajenos a los subconjuntos de entrenamiento y validación, con finalidad de obtener de calcular los índices de similitud y diferencia entre las segmentaciones esperadas y las predichas

por el modelo.

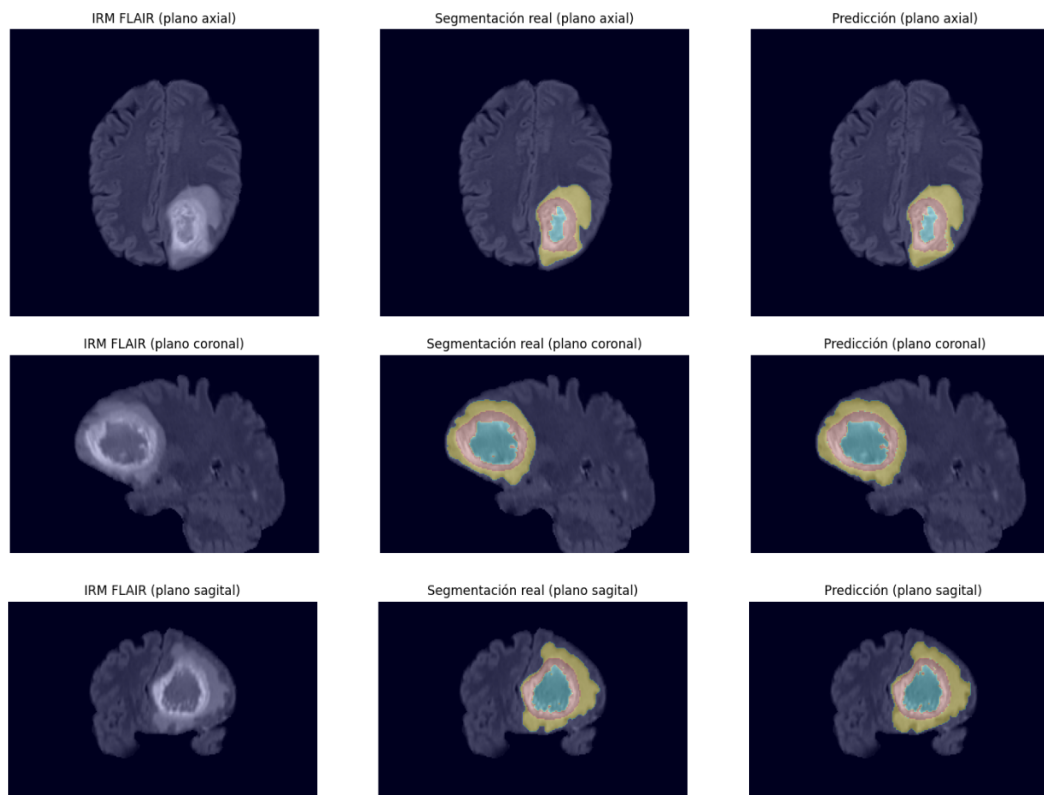


Figura 7.3: Predicciones realizadas por el modelo, donde cada fila muestra un plano anatómico distinto de un mismo caso del conjunto de prueba. En la primera columna un corte de RM de la secuencia FLAIR, la segunda columna contiene la segmentación realizada por el especialista y por último en la tercera columna se muestra la predicción obtenida del modelo entrenado.

Al obtener las predicciones de los diferentes casos de prueba, cada uno muestra una gran similitud entre lo esperado y lo predicho por el modelo, comparando las subestructuras del tumor que se pueden visualizar con colores diferentes dentro de cada imagen (Figura 7.10). Con las segmentaciones obtenidas se calcula el coeficiente DICE y la distancia de Hausdorff para obtener una evaluación general de desempeño del modelo como se observa en la Tabla 7.1 y la Tabla 7.2, donde se muestran los coeficientes, sus promedios y la desviación estándar obtenidos para cada una de las clases de salida (WT, TC, ET) del modelo repartidos por columna, cada fila representa un estudio

de RM distinto del conjunto de prueba, y por último se muestra un promedio global y su desviación estándar para cada una de las métricas obtenidas.

Aunado al cálculo de los coeficientes para evaluar el rendimiento y exactitud del modelo, se realizaron las predicciones de segmentación de 10 casos de RM del conjunto de evaluación para el desafío *BraTS'23* disponible para la fase de evaluación de los modelos participantes. Como estos casos de RM son para probar y evaluar la segmentación del modelo implementado no contienen un archivo con máscaras segmentadas para tomarlas como referencia, aun así las predicciones realizadas por el modelo entrenado muestran segmentaciones que contienen las tres subestructuras del tumor en la mayoría de los casos independientemente (Figura 7.4 y Figura 7.5).

Para cuantificar la calidad de estas predicciones se ha calculado el promedio del coeficiente DICE de todas las muestras del conjunto de prueba, obteniendo un índice promedio global de 0.882 ± 0.04 . Tomando en cuenta el número de casos seleccionados para el entrenamiento y evaluación, la similitud de las predicciones y que en estas no se muestran artefactos o ruido se considera que el modelo obtuvo un gran desempeño en la tarea de segmentación, comprobando que la implementación de este tipo de tecnologías es viable para proyectos locales incluso los desarrollados con recursos de hardware limitados y aún así obtener predicciones precisas y de buena resolución, lo que sienta la bases sólidamente para la reconstrucción 3D de las predicciones que se realizan a partir de las segmentaciones obtenidas por el modelo entrenado.

Evaluación del conjunto de Prueba				
Coeficiente DICE				
Prueba	TC	WT	ET	Avg.
1	0.315	0.961	0.844	0.707±0.28
2	0.545	0.948	0.883	0.792±0.17
3	0.689	0.934	0.899	0.841±0.10
4	0.756	0.937	0.899	0.864±0.07
5	0.763	0.927	0.876	0.856±0.06
6	0.795	0.927	0.876	0.866±0.05
7	0.795	0.907	0.876	0.859±0.04
8	0.822	0.913	0.888	0.875±0.03
9	0.842	0.918	0.899	0.886±0.03
10	0.853	0.920	0.906	0.893±0.02
11	0.839	0.918	0.895	0.884±0.03
12	0.853	0.922	0.889	0.888±0.02
13	0.864	0.927	0.895	0.895±0.02
14	0.868	0.929	0.894	0.897±0.02
15	0.874	0.928	0.896	0.899±0.02
16	0.879	0.929	0.892	0.900±0.02
17	0.882	0.927	0.891	0.900±0.01
18	0.889	0.930	0.893	0.904±0.01
19	0.892	0.933	0.895	0.907±0.01
20	0.888	0.934	0.887	0.903±0.02
21	0.890	0.932	0.886	0.903±0.02
22	0.894	0.934	0.890	0.906±0.01
23	0.926	0.924	0.883	0.911±0.01
24	0.945	0.921	0.910	0.925±0.01
25	0.954	0.931	0.927	0.937±0.01
26	0.939	0.932	0.897	0.923±0.01
27	0.943	0.937	0.902	0.927±0.01
28	0.899	0.941	0.777	0.872±0.06
29	0.908	0.945	0.798	0.884±0.06
30	0.903	0.942	0.710	0.852±0.10
Avg.	0.837±0.12	0.931±0.01	0.879±0.04	0.882±0.04

Tabla 7.1: Cálculo de coeficientes DICE del conjunto de prueba mostrando los índices de cada una de las clases de salida del modelo, junto con los promedios y la desviación estándar de cada caso, y el promedio global del conjunto de datos.

Evaluación del conjunto de Prueba				
Coeficiente Hausdorff				
Prueba	TC	WT	ET	Avg.
1	30.298	103.696	19.052	51.015±37.53
2	25.457	58.848	13.431	32.578±19.21
3	18.385	43.500	10.407	24.097±14.10
4	14.907	36.883	8.740	20.177±12.07
5	21.939	36.344	17.052	25.112±8.18
6	18.949	36.094	14.877	23.307±9.19
7	18.949	33.293	14.877	22.373±7.89
8	17.050	30.904	13.451	20.468±7.52
9	15.169	28.182	12.049	18.467±6.98
10	14.050	26.353	10.959	17.121±6.64
11	14.486	25.877	11.621	17.328±6.15
12	13.260	24.675	11.125	16.353±5.94
13	12.405	22.949	10.364	15.239±5.51
14	11.919	22.079	10.002	14.666±5.29
15	11.726	21.818	9.850	14.465±5.25
16	11.303	21.057	9.552	13.971±5.06
17	11.089	20.930	9.447	13.822±5.07
18	10.603	22.701	9.240	14.181±6.04
19	10.966	21.934	9.690	14.197±5.49
20	10.860	21.252	9.651	13.921±5.20
21	10.681	22.455	9.507	14.214±5.84
22	10.624	21.848	9.514	13.995±5.57
23	6.073	21.035	5.894	11.001±7.09
24	4.794	27.625	4.745	12.388±10.77
25	4.942	39.833	4.973	16.583±16.44
26	16.315	42.191	16.340	24.948±12.19
27	14.122	36.883	14.143	21.716±10.72
28	14.766	40.216	16.380	23.787±11.63
29	13.493	35.802	14.638	21.311±10.25
30	14.379	46.528	18.277	26.395±14.32
Avg.	14.132±5.34	33.127±16.12	11.662±3.60	19.640±7.68

Tabla 7.2: Cálculo de coeficientes Hausdorff del conjunto de prueba mostrando los índices de cada una de las clases de salida del modelo, junto con los promedios y la desviación estándar de cada caso, y el promedio global del conjunto de datos .

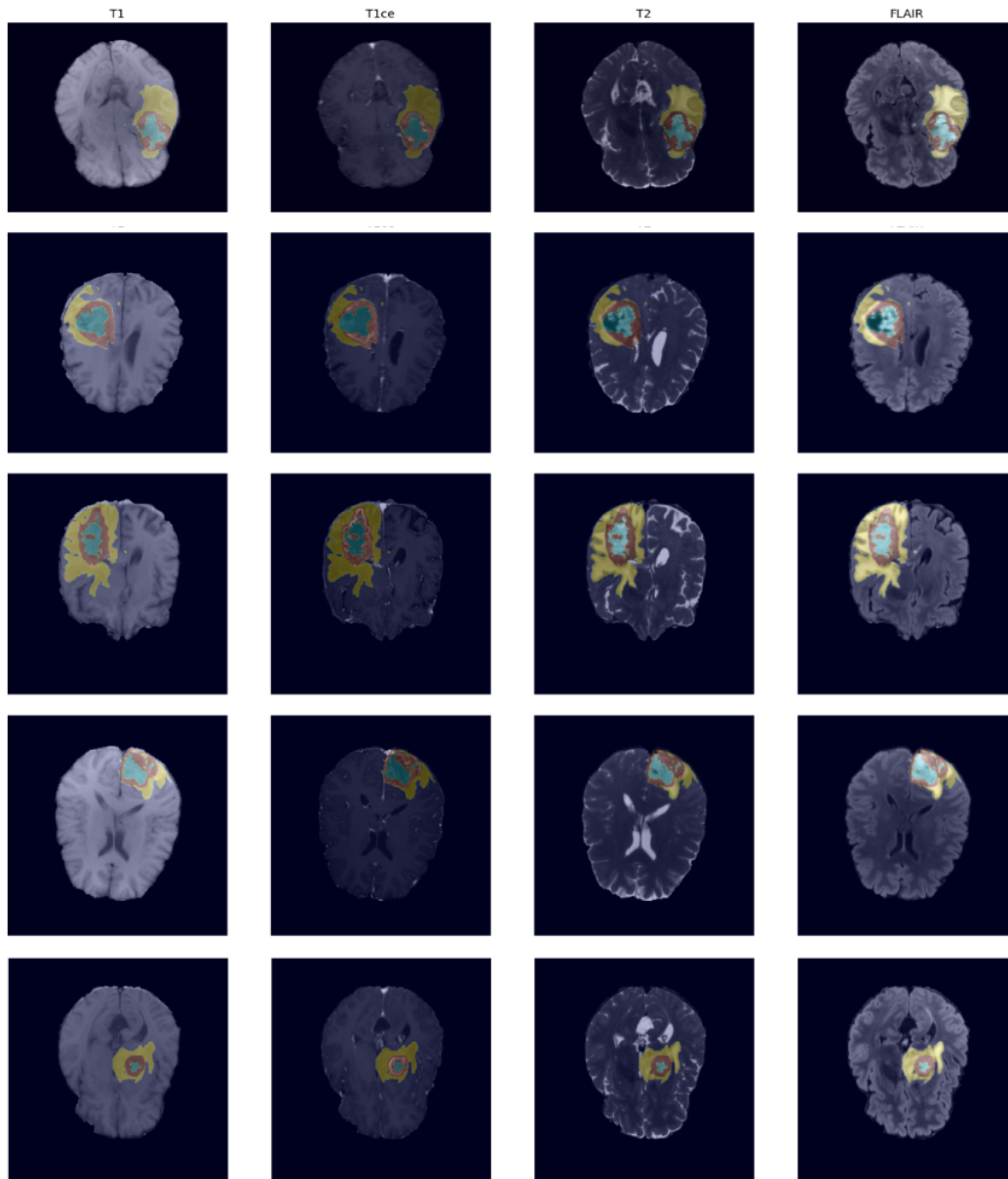


Figura 7.4: Predicciones realizadas por el modelo de los primeros 5 casos de estudio de RM, donde cada fila muestra un caso distinto del conjunto de evaluación del desafío *BraTS'23*, cada columna muestra una secuencia de captura diferente (T1,T1ce,T2,FLAIR). Las imágenes presentan la segmentación automática de GBM en casos de RM, resaltando con distinto color las subestructuras que componen al tumor: amarillo para Edema, café para el tumor realzado y azul para el núcleo del tumor.

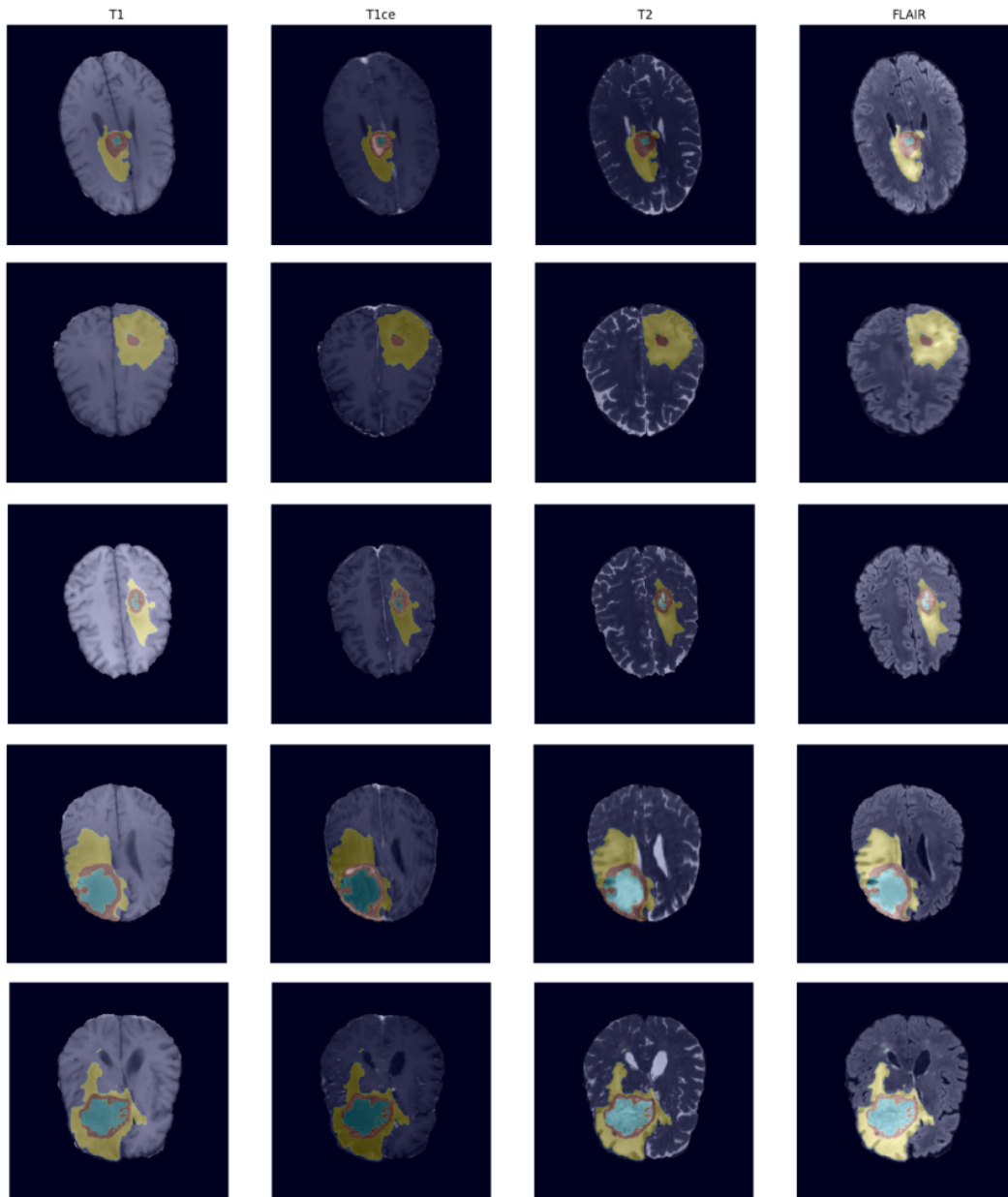


Figura 7.5: Predicciones realizadas por el modelo de 5 casos de estudio de RM, donde cada fila muestra un caso distinto del conjunto de evaluación del desafío *BraTS'23*, cada columna muestra una secuencia de captura diferente (T1,T1ce,T2,FLAIR). Las imágenes presentan la segmentación automática de GBM en casos de RM, resaltando con distinto color las subestructuras que componen al tumor: amarillo para Edema, cafe para el tumor realzado y azul para el núcleo del tumor. En la práctica la secuencia FLAIR es la que consultan con más frecuencia los especialistas para armar un diagnóstico o tratamiento.

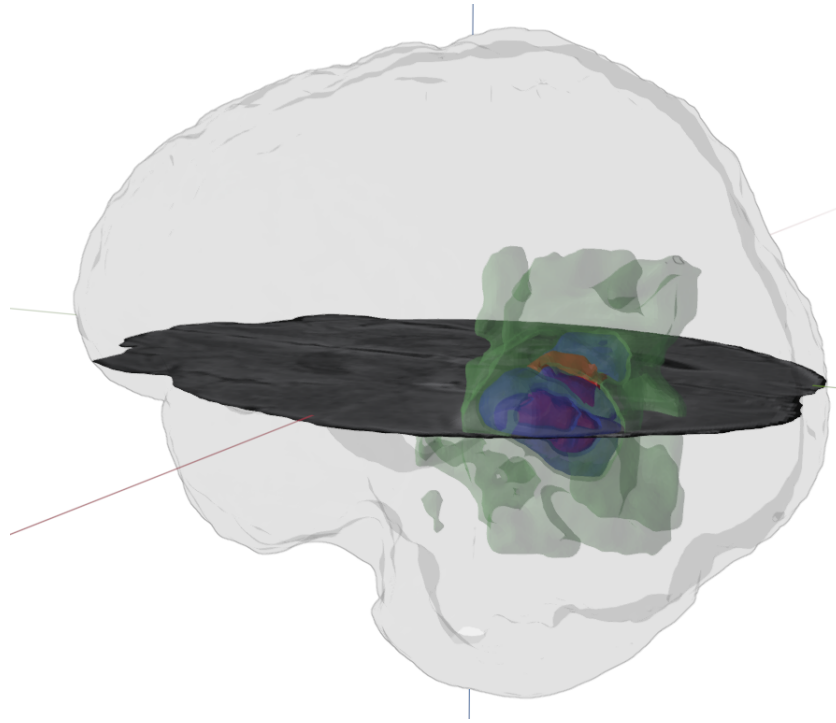


Figura 7.6: Reconstrucción 3D de la segmentaciones automática realizada por el modelo entrenado Swin UNETR de un caso de GBM del conjunto de prueba.

7.2 Reconstrucción

Para realizar la reconstrucción del tumor a una representación en 3D vamos hacer uso principalmente de la herramienta de software VTK, una herramienta versátil y poderosa utilizada en una variedad de aplicaciones, desde la visualización médica hasta la simulación científica y la realidad virtual.

Tras obtener las predicciones del modelo como un tensor ordenado de cortes con las subestructuras del tumor segmentadas, se utilizan los cortes segmentados del plano axial para iniciar la reconstrucción de 3D tanto de las áreas segmentadas como del resto del tejido cerebral, dando como resultado un archivo *.STL* para cada una de las clases de salida del modelo o subestructuras

del tumor y un archivo adicional con la reconstrucción del cerebro obtenida del estudio de RM (Figura 7.7).

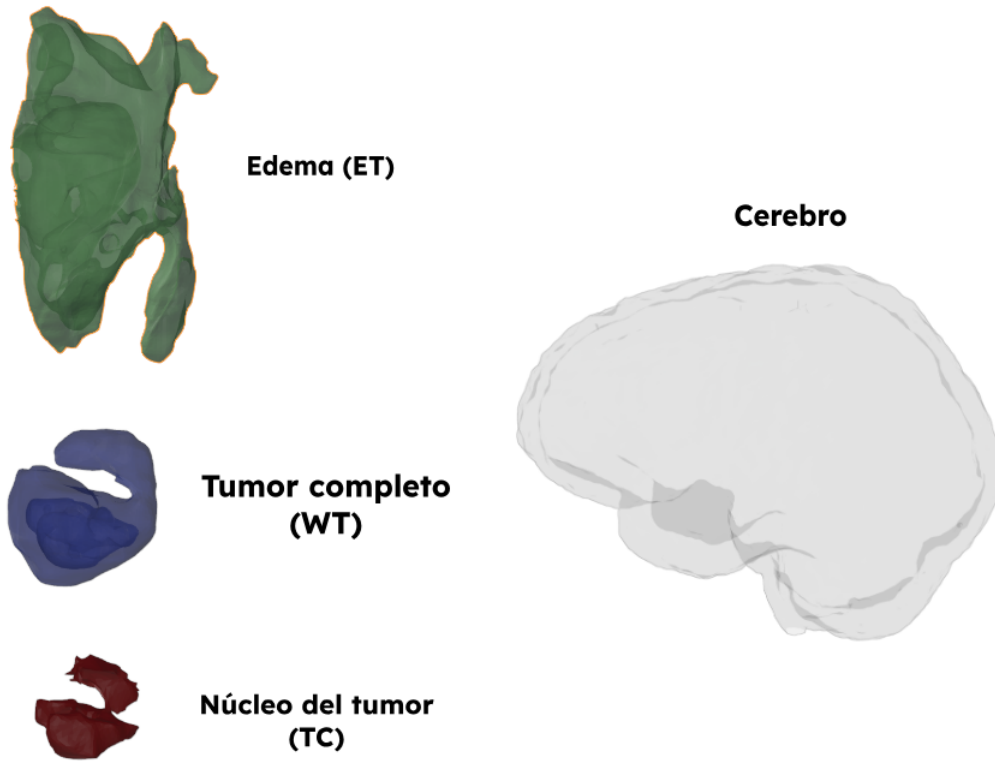


Figura 7.7: Reconstrucción 3D del cerebro y las subestructuras del tumor provenientes de las predicciones del plano axial realizadas por el modelo entrenado.

Para visualizar las reconstrucción 3D completa se importan los archivos .STL al software de diseño Blender (Figura 7.8). Cabe señalar que la resolución de la reconstrucción dependerá directamente de la resolución de las imágenes con las que entrenamos el modelo Swin UNETR.

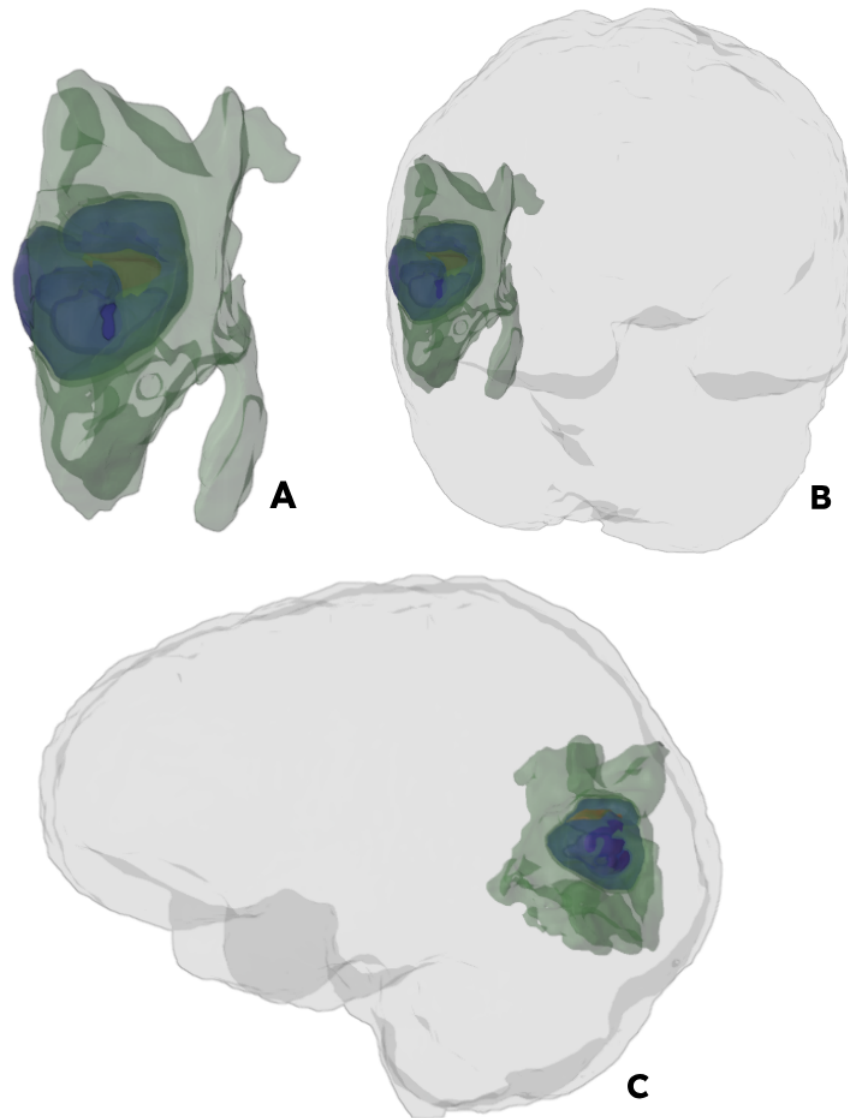


Figura 7.8: La figura muestra A) reconstrucción completa del tumor, B) vista coronal y C) vista sagital de la reconstrucción del tumor con el cerebro.

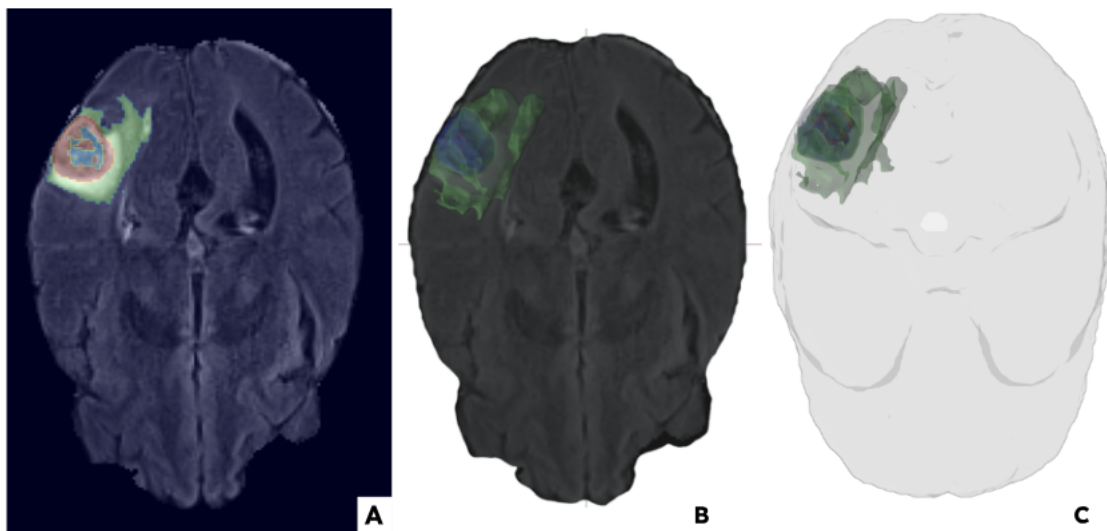


Figura 7.9: Comparación de A) un corte de RM del plano axial de un caso de prueba junto con B) la reconstrucción 3D del tumor completo y c) la reconstrucción 3D del tumor sobre el corte de RM.

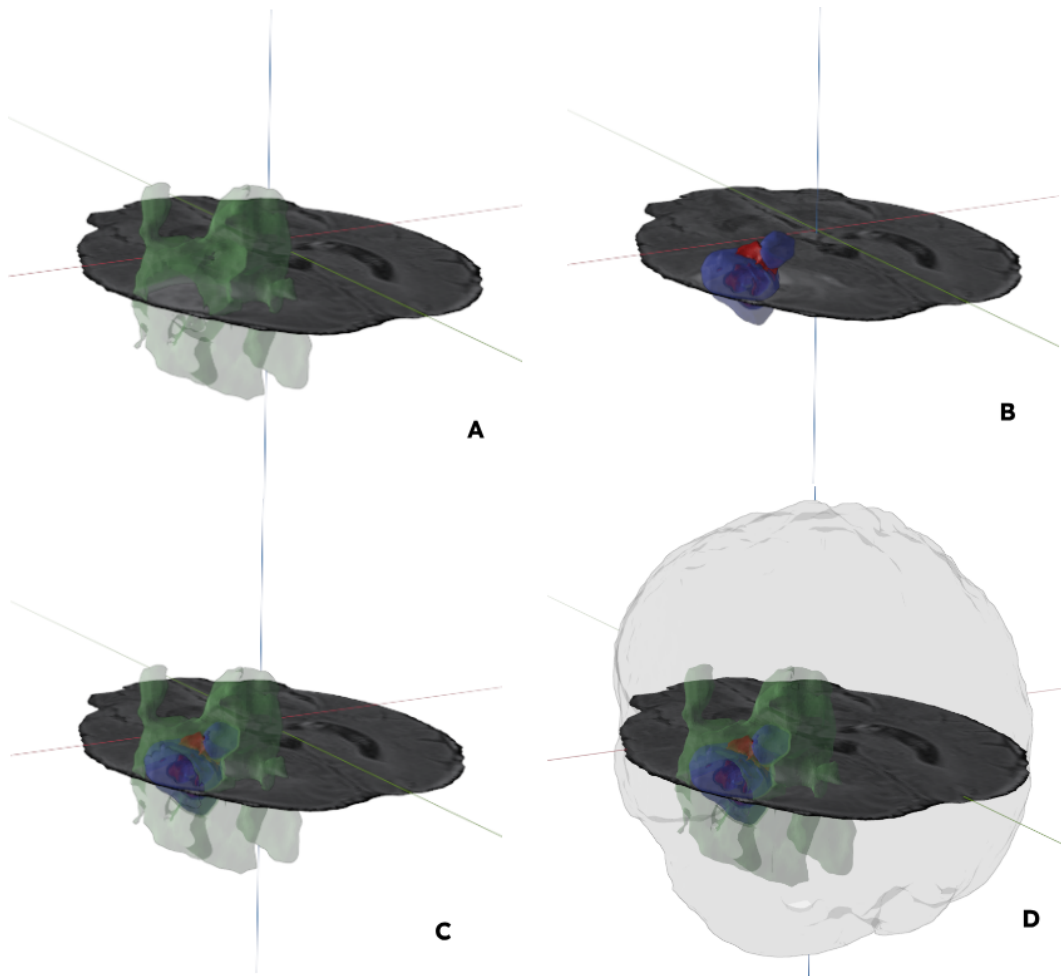


Figura 7.10: La Figura muestra la reconstrucción de las diferentes subestructuras del tumor y un corte de plano axial del estudio de RM que sirve como referencia de la posición del tumor con respecto al cerebro. A) Reconstrucción 3D del edema, B) reconstrucción del tumor completo y el núcleo tumoral, C) reconstrucción con todas las subestructuras del tumor y D) Reconstrucción 3D del tumor completo y el cerebro.

Capítulo 8

Discusión y Conclusiones

8.1 Discusiones

Con base en los resultados obtenidos del entrenamiento de las métricas de los coeficientes DICE y Hausdorff del conjunto de prueba, se mostró que los modelos basados en ViT tienen un buen desempeño procesando grandes conjuntos de datos. No obstante, esta misma capacidad puede convertirse en una limitante si no se cuenta con las bases de datos adecuadas para el entrenamiento. En trabajos relacionados de la literatura, con modelos de esta arquitectura, entrenados en ambientes de hardware más robusto, se obtiene un valor promedio Dice de 90.2 en predicciones de múltiples categorías con un conjunto de datos de 1251 estudios de RM. Aunque el modelo implementado en este proyecto utiliza un menor número de imágenes, se logró obtener un coeficiente DICE promedio de 88.2, lo que apunta que se debe de tener un mínimo de datos para que el modelo Swin-UNETR pueda entrenarse de manera adecuada y obtener

buenos resultados en sus predicciones.

La segmentación multiclase que realiza este modelo en las predicciones es otra de las ventajas que otorga esta arquitectura en comparación con otros modelos convolucionales, entregando información sobre las áreas que abarca cada una de estas subestructura del tumor. Aunado a esto, la reconstrucción en tres dimensiones incrementa la utilidad del proyecto al mostrar en volumen el cómo está distribuido el tumor en el tejido cerebral en cada una de estas subestructuras.

Como ya se ha mencionado, el diagnóstico y tratamiento de GBM hace uso de distintas técnicas de análisis como el estudios de imágenes médicas. En este proyecto centramos la atención en el caso específico de IRM de contraste para implementar la segmentación y reconstrucción tridimensional de tumores, y proponemos una herramienta adicional, en la que los especialistas puedan apoyarse para el diagnóstico de este tipo de padecimientos.

Aun después de las mejores métricas de similitud y de la reconstrucción de las subestructuras del GBM esta información es sólo un referente visual y podría no ser suficiente para lograr un diagnóstico definitivo ya que faltaría integrar otros datos como la cercanía del GBM a zonas de alta funcionalidad o las vías que puede estar afectando por la presencia del tumor dentro del tejido cerebral. Aunque parece ser mucha la información que se debe de integrar, la evolución de los modelos de aprendizaje profundo en específico ofrecen la posibilidad de ir escalando funcionalidades para lograr obtener mejores resultados. La integración de estas nuevas tecnologías abre la posibilidad para el desarrollo de mejores herramientas cada vez más precisas para la planeación y navegación quirúrgica.

8.2 Conclusiones y trabajo a futuro

El desarrollo de este proyecto propone la segmentación y la reconstrucción de GBM implementando arquitecturas actuales de aprendizaje profundo basado en transformadores de visión como lo es la Swin-UNETR. Así también, propone el uso de distintas bibliotecas de software actualizadas que son útiles para realizar la reconstrucción de imágenes médicas, con la finalidad de generar una herramienta que integre estas tecnologías para realizar tareas con mayor complejidad enriqueciendo los datos generados en cada etapa, desde el preprocesamiento de los datos de entrenamiento, validación y reconstrucción 3D, se produce una herramienta que ofrece mayor información sobre la distribución y localización de este tipo de tumores.

La tarea de segmentación y reconstrucción no bastan por sí solas para apoyar a la obtención de un diagnóstico confiable, pero la integración de más datos sobre padecimientos relacionados a esta neoplasia, así como el análisis de otro tipo de secuencias en IRM con otros enfoques de IA, generar reconstrucciones 3D con atributos biomecánicos con el que puedan interactuar los usuarios, pueden generar en conjunto características que aumentan el valor al realizar proyectos con este enfoque. Y es a través de la madurez alcanzada por las tecnologías de la información cómo nos damos cuenta de la viabilidad de estos proyectos, abriendo la ventana a nuevas formas de integración de las tecnologías en las áreas donde son más necesitadas.

Como trabajo a futuro proponemos extender el procesamiento de datos para realizar una reconstrucción más amplia con un modelo de construcción de conectomas implementando IA y aprendizaje profundo que planea generar la identificación precisa de las estructuras neuronales y sus conexiones, logrando

proporcionar una comprensión más detallada y precisa de la organización funcional del cerebro. Así mismo se busca mejorar el diagnóstico y tratamiento de enfermedades neurológicas permitiendo una mejor comprensión de las alteraciones de la conectividad en presencia de GBM, lo que podría ayudar a mejorar el diagnóstico y tratamiento de estas lesiones. Los modelos de prueba que utilizan información de conectomas en el quirófano parten de la aproximación de modelos entrenados con sujetos sanos para predecir conexiones cerebrales de pacientes con esta neoplasia, por lo que es relevante la contribución de crear una base de datos de pacientes con este padecimiento e investigar cómo se mejora un modelo, a priori, entrenado en sujetos sanos pero reajustado con datos de casos reales de pacientes diagnosticados con GBM.

La construcción de un modelo de conectividad cerebral podría proporcionar información valiosa preoperatoria para planificar y llevar a cabo procedimientos quirúrgicos, lo que mejoraría la reducción de riesgo al paciente. De la misma manera el desarrollo de modelos de conectividad cerebral utilizando aprendizaje profundo proporciona nuevos aportes a los sistemas de navegación y de diagnóstico existentes para el desarrollo de nuevas investigaciones con el objetivo de mejorar los resultados en cirugías cerebrales.

Bibliografía

- [1] David N Louis et al. “The 2016 World Health Organization classification of tumors of the central nervous system: a summary”. En: *Acta neuropathologica* 131 (2016), págs. 803-820.
- [2] Qinchuan Liang et al. “Inhibition of STAT3 reduces astrocytoma cell invasion and constitutive activation of STAT3 predicts poor prognosis in human astrocytoma”. En: *PLoS One* 8.12 (2013), e84723.
- [3] Xiaoyang Li. *Investigate the Impact of Traumatic Brain Injury (TBI) on glia neighborhoods of neuron*. Mar. de 2018.
- [4] Agil Vazquez Eric Jesús, De la Cruz Chávez Francisca y Corriéa Milán Ivanis Idael. “Astrocitoma fibrilar difuso grado II en edad pediátrica. A proposito de un caso”. En: *cibamanz2021*. 2021.
- [5] René David Astudillo Mancero et al. “Astrocitoma, diagnóstico y tratamiento”. En: *RECIMUNDO* 5.4 (2021), págs. 53-64.
- [6] Andrew Conger y Jason Wilson. “Controversies in Neuro-Oncology: Best Evidence Medicine for Brain Tumor Surgery: A Review Quinones-Hinojosa A, Raza SM. Edited by: Alfredo Quinones-Hinojosa, MD, and Shaan M. Raza, MD New York, NY: Thieme, 2014.” En: *Otology & Neurotology* 35.6 (2014), pág. 1016.

- [7] Jayesh P. Thawani y Steven Brem. “Chapter 1 - The Story of Glioblastoma: History and Modern Correlates”. En: *Glioblastoma*. Elsevier, 2016, págs. 1-9.
- [8] Sonia Tejada Solhés et al. “Glioblastoma treatment guidelines: consensus by the Spanish society of neurosurgery tumor section”. En: *Neurocirugía (English Edition)* 31.6 (2020), págs. 289-297.
- [9] M Havaei et al. “Brain Tumor Segmentation with Deep Neural Networks. Cornell University Library (2016)”. En: *Medical Image Analysis* 35 (2017), págs. 18-31.
- [10] José Ignacio Erices et al. “Glioblastoma Microenvironment and Invasiveness: New Insights and Therapeutic Targets”. En: *International Journal of Molecular Sciences* 24.8 (2023), pág. 7047.
- [11] Mary Elizabeth Davis. “Glioblastoma: overview of disease and treatment”. En: *Clinical journal of oncology nursing* 20.5 (2016), S2.
- [12] Christina L Appin et al. “Glioblastoma with oligodendroglioma component (GBM-O): molecular genetic and clinical characteristics”. En: *Brain pathology* 23.4 (2013), págs. 454-461.
- [13] A Armento et al. *Glioblastoma (de Vleeschouwer S., ed.)* 2017.
- [14] Hao Xu et al. “Geographic variations in the incidence of glioblastoma and prognostic factors predictive of overall survival in US adults from 2004–2013”. En: *Frontiers in aging neuroscience* 9 (2017), pág. 352.
- [15] Alfredo Quinones-Hinojosa. *Schmidek and Sweet: Operative Neurosurgical Techniques E-Book: Indications, Methods and Results (Expert Consult-Online and Print)*. Elsevier Health Sciences, 2012.

- [16] Pamela Garcíea-Corrochano et al. “Resección microquirúrgica de glioblastoma guiada con fluoresceína intraoperatoria: evaluación retrospectiva”. En: *Revista Peruana de Medicina Experimental y Salud Pública* 32 (2015), págs. 471-478.
- [17] J Ferlay et al. “Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Published 2020”. En: (2020).
- [18] Sergio Rojas Carmenathy, Yorkiel Castellanos Bertot y Ayansi Massó Rodríguez. “Astrocitoma medular de alto grado no infiltrante”. En: *Revista Información Científica* 98.5 (2019), págs. 640-647.
- [19] José A Chávez-López et al. “Astrocitoma anaplásico multifocal sincrónico con localización lobar atípica”. En: *Revista de Especialidades Médico-Quirúrgicas* 23.3 (2019), pág. 145.
- [20] Génesis Carolina Pérez Cruz, Elizabeth Stefania Molina Vega y Raquel Alexandra Colcha González. “Astrocitoma fibrilar difuso”. En: *RECIMUNDO* 3.3 ESP (2019), págs. 853-875.
- [21] PDQ Adult Treatment Editorial Board. “Adult central nervous system tumors treatment (PDQ®)”. En: *PDQ Cancer Information Summaries [Internet]*. National Cancer Institute (US), 2022.
- [22] Xiaoxia Cai et al. “Manganese oxide nanoparticles as MRI contrast agents in tumor multimodal imaging and therapy”. En: *International journal of nanomedicine* (2019), págs. 8321-8344.
- [23] Conroy Sun, Jerry SH Lee y Miqin Zhang. “Magnetic nanoparticles in MR imaging and drug delivery”. En: *Advanced drug delivery reviews* 60.11 (2008), págs. 1252-1265.

- [24] Hyon Bin Na, In Chan Song y Taeghwan Hyeon. “Inorganic nanoparticles for MRI contrast agents”. En: *Advanced materials* 21.21 (2009), págs. 2133-2148.
- [25] Ashish Avasthi et al. “Magnetic nanoparticles as MRI contrast agents”. En: *Surface-modified Nanobiomaterials for Electrochemical and Biomedicine Applications* (2020), págs. 49-91.
- [26] Ming Lei et al. “Activated surface charge-reversal manganese oxide nanocubes with high surface-to-volume ratio for accurate magnetic resonance tumor imaging”. En: *Advanced Functional Materials* 27.30 (2017), pág. 1700978.
- [27] H Menze Bjoern et al. “The multimodal brain tumor image segmentation benchmark (BRATS)”. En: *IEEE Trans. Med. Imaging* 34.10 (2015), págs. 1993-2024.
- [28] M Touat et al. “Glioblastoma targeted therapy: updated approaches from recent biological insights”. En: *Annals of Oncology* 28.7 (2017), págs. 1457-1472.
- [29] Aura A Erazo-Valle-Solés et al. “Recomendaciones institucionales ISSSTE en el tratamiento multidisciplinario de los gliomas de alto grado”. En: *Revista de Especialidades Médico-Quirúrgicas* 22.3 (2018), págs. 110-130.
- [30] R Stupp et al. “High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up”. En: *Annals of oncology* 25 (2014), págs. iii93-iii101.
- [31] Helena Benito Naverac. “Consideraciones anestésicas con el uso del GLIO-LAN® para la resección de gliomas malignos”. En: *Revista electrónica AnestesiaR* 10.7 (2018), pág. 3.

- [32] LM Bernal-Garcíea et al. “Resección de tumor intramedular guiada por fluorescencia con ácido aminolevulínico”. En: *Neurocirugía* 21.4 (2010), págs. 312-316.
- [33] Ray M. Chu y Keith L. Black. “Chapter 8 - Current Surgical Management of High-Grade Gliomas”. En: *Schmidek and Sweet Operative Neurosurgical Techniques (Sixth Edition)*. Ed. por Alfredo Quiñones-Hinojosa. Sixth Edition. W.B. Saunders, 2012, págs. 105-110.
- [34] Oscar Andrés Escobar Vidarte, Jairo Alberto Espinoza Martiénez y René Julián Varela Osorio. “Aproximación diagnóstica a las lesiones del tallo cerebral en niños y adultos a través de biopsia guiada por estereotaxia.” En: *Gastrohnap* 17.3 (2015), págs. 144-154.
- [35] Miguel Ángel Celis et al. “Primer consenso mexicano sobre recomendaciones de la atención multidisciplinaria del paciente con glioblastoma multiforme (GBM). Grupo Interdisciplinar Mexicano de Investigación en Neurooncología (GIMINO)”. En: *Gaceta Médica de México* 151.3 (2015), págs. 403-415.
- [36] Gabriela Núñez-Guardado et al. “Recomendaciones institucionales ISSSTE en el tratamiento multidisciplinario de los gliomas de alto grado”. En: *Revista de Especialidades Médico-Quirúrgicas* 22.3 (2017), págs. 110-130.
- [37] John M Tew Jr. “M. Gazi Yaşargil: Neurosurgery’s man of the century”. En: *Neurosurgery* 45.5 (1999), pág. 1010.
- [38] Brian M Dale, Mark A Brown y Richard C Semelka. *MRI: basic principles and applications*. John Wiley & Sons, 2015.
- [39] Lynn M Fletcher-Heath et al. “Automatic segmentation of non-enhancing brain tumors in magnetic resonance images”. En: *Artificial intelligence in medicine* 21.1-3 (2001), págs. 43-63.

- [40] Nadine Barrie Smith y Andrew Webb. *Introduction to medical imaging: physics, engineering and clinical applications*. Cambridge university press, 2010.
- [41] Jerrold T Bushberg y J Anthony Seibert. *The Essential Physics of Medical Imaging Study Guide*. Lippincott Williams & Wilkins, 2022.
- [42] AD Elster y JH Burdette. *Questions & Answers in Magnetic Resonance Imaging*. 2001.
- [43] Diego Miguel Rivera, Sofíea Puentes y Ligia Caballero. “Resonancia magnética cerebral: secuencias básicas e interpretación”. En: *Universitas Médica* 52.3 (2011), págs. 292-306.
- [44] Robert A Pooley. “Fundamental physics of MR imaging”. En: *Radiographics* 25.4 (2005), págs. 1087-1099.
- [45] Richard Bitar et al. “MR pulse sequences: what every radiologist wants to know but is afraid to ask”. En: *Radiographics* 26.2 (2006), págs. 513-537.
- [46] Michael A Jacobs, Tamer S Ibrahim y Ronald Ouwerkerk. “MR imaging: brief overview and emerging applications”. En: *Radiographics* 27.4 (2007), págs. 1213-1229.
- [47] Joseph P Hornak. “The basics of MRF”. En: <http://www.cis.rit.edu/htbooks/mri> (2006).
- [48] Jaume Gili. “Introducción biofísica a la resonancia magnética”. En: *Centre Diagnòstic Pedralbes* 5 (1993).
- [49] Donald W McRobbie et al. *MRI from Picture to Proton*. Cambridge university press, 2017.
- [50] Mohit Neema et al. “Normal findings on brain fluid-attenuated inversion recovery MR images at 3T”. En: *American Journal of Neuroradiology* 30.5 (2009), págs. 911-916.

- [51] Jeroen Van der Laak, Geert Litjens y Francesco Ciompi. “Deep learning in histopathology: the path to the clinic”. En: *Nature medicine* 27.5 (2021), págs. 775-784.
- [52] Liang Jin et al. “Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet”. En: *EBioMedicine* 62 (2020).
- [53] Thorsten Hoeser y Claudia Kuenzer. “Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends”. En: *Remote Sensing* 12.10 (2020), pág. 1667.
- [54] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.
- [55] Gabriel Chartrand et al. “Deep learning: a primer for radiologists”. En: *Radiographics* 37.7 (2017), págs. 2113-2131.
- [56] Andriy Fedorov et al. “3D Slicer as an image computing platform for the Quantitative Imaging Network”. En: *Magnetic resonance imaging* 30.9 (2012), págs. 1323-1341.
- [57] Paul A. Yushkevich et al. “User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability”. En: *Neuroimage* 31.3 (2006), págs. 1116-1128.
- [58] Bruce Fischl. “FreeSurfer”. En: *NeuroImage* 62.2 (2012). 20 YEARS OF fMRI, págs. 774-781. ISSN: 1053-8119.
- [59] Antoine Rosset, Luca Spadola y Osman Ratib. *OsiriX: an open-source software for navigating in multidimensional DICOM images*. 2004.
- [60] Takfarinas Medani et al. “Brainstorm-DUNEuro: An integrated and user-friendly Finite Element Method for modeling electromagnetic brain activity”. En: *Neuroimage* 267 (2023), pág. 119851.

- [61] M Jorge Cardoso et al. “Monai: An open-source framework for deep learning in healthcare”. En: *ArXiv* (2022).
- [62] Konstantinos Kamnitsas et al. “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”. En: *Medical image analysis* 36 (2017), págs. 61-78.
- [63] Van J Wedeen et al. “The geometric structure of the brain fiber pathways”. En: *Science* 335.6076 (2012), págs. 1628-1634. URL: <http://www.humanconnectomeproject.org>.
- [64] Klaus H Maier-Hein et al. “Tractography-based connectomes are dominated by false-positive connections”. En: *BioRxiv* (2016), pág. 084137. URL: <https://tractometer.org/>.
- [65] Dominic LaBella et al. “The ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge 2023: Intracranial Meningioma”. En: *ArXiv* (2023).
- [66] Anahita Fathi Kazerooni et al. “The Brain Tumor Segmentation (BraTS) Challenge 2023: Focus on Pediatrics (CBTN-CONNECT-DIPGR-ASNR-MICCAI BraTS-PEDs)”. En: *ArXiv* (2023).
- [67] Anna K Greenwood et al. “The AD Knowledge Portal: A Repository for Multi-Omic Data on Alzheimer’s Disease and Aging”. En: *Current protocols in human genetics* 108.1 (2020), e105.
- [68] Ujjwal Baid et al. “The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification”. En: *ArXiv* (2021).
- [69] Data Format Working Group et al. *NIfTI:—neuroimaging informatics technology initiative*. 2013.

- [70] Ali Hatamizadeh et al. “Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images”. En: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 7th International Workshop, BrainLes 2021, Held in Conjunction with MICCAI 2021, Virtual Event, September 27, 2021, Revised Selected Papers, Part I*. Springer. 2022, págs. 272-284.
- [71] Ashish Vaswani et al. “Attention is all you need”. En: *Advances in neural information processing systems* 30 (2017).
- [72] Alexey Dosovitskiy et al. “An image is worth 16x16 words: Transformers for image recognition at scale”. En: *ArXiv* (2020).
- [73] Ze Liu et al. “Swin transformer: Hierarchical vision transformer using shifted windows”. En: *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, págs. 10012-10022.
- [74] Ali Hatamizadeh et al. “Unetr: Transformers for 3d medical image segmentation”. En: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2022, págs. 574-584.
- [75] Adam Paszke et al. “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. En: *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, págs. 8024-8035.
- [76] W Schroeder, K Martin y B Lorensen. “The visualization toolkit, 4th edn. Kitware”. En: *New York* (2006).
- [77] Matthew McCormick et al. “ITK: enabling reproducible research and open science”. En: *Frontiers in neuroinformatics* 8 (2014), pág. 13.
- [78] Aaron Carass et al. “Evaluating white matter lesion segmentations with refined Sørensen-Dice analysis”. En: *Scientific reports* 10.1 (2020), pág. 8242.

- [79] Maarten Grootendorst. “9 Distance Measures in Data Science: The advantages and pitfalls of common distance measures”. En: *Towards Data Science* (9).
- [80] Min Tang, Minkyoung Lee y Young J Kim. “Interactive Hausdorff distance computation for general polygonal models”. En: *ACM Transactions on Graphics (TOG)* 28.3 (2009), págs. 1-9.
- [81] Daniel Ortiz Nieto. “Paralelización de un algoritmo de búsqueda de patrones en imágenes basado en la distancia de Hausdorff”. Tesis de maestría. 2012.