



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Informática

Diseño de una herramienta de clasificación semántica de  
imágenes de gliomas de alto grado mediante redes  
convolucionales profundas

Trabajo Fin de Grado

Grado en Ingeniería Informática

AUTOR/A: Juanes Busolo, Rodrigo

Tutor/a: García Gómez, Juan Miguel

Cotutor/a: Fuster i Garcia, Elies

CURSO ACADÉMICO: 2022/2023



# Resumen

---

Las redes convolucionales profundas han demostrado su gran valía para segmentar semánticamente imágenes complejas. Específicamente, las redes basadas en la arquitectura *UNet* han dado grandes resultados para la clasificación de imágenes prequirúrgicas multiparamétricas de glioma. En este proyecto realizaremos una comparación de arquitecturas propuestas recientemente con el objetivo de comprobar su rendimiento de forma comparada, por lo que se probarán arquitecturas como *R2U-Net*, *SegNet*, *X-Net* y *Multires-UNET* entre otras.

**Palabras clave:** redes neuronales convolucionales, imagen médica, glioma, segmentación

# Abstract

---

Deep convolutional networks have proven their excellent value for semantically segmenting complex images. Specifically, networks based on the *UNet* architecture have yielded remarkable results for the classification of multiparametric pre-surgical glioma images. In this project we will perform a comparison of recently proposed architectures with the aim of evaluating their performance in a comparative way, so architectures such as *R2U-Net*, *SegNet*, *X-Net* and *Multires-UNET*, among others, will be assessed.

**Key words:** convolutional neural networks, medical image, glioma, segmentation

# Resum

---

Les xarxes convolucionals profundes han demostrat la seua gran vàlua per a segmentar semànticament imatges complexes. Específicament, les xarxes basades en l'arquitectura UNet han donat grans resultats per a la classificació d'imatges prequirúrgiques multiparamétricas de glioma. En este projecte realitzarem una comparació d'arquitectures proposades recentment amb l'objectiu de comprovar el seu rendiment de forma comparada, per la qual cosa es provaran arquitectures com *R2U-Net*, *SegNet*, *X-Net* i *Multires-UNET*, entre altres.

**Paraules clau:** xarxes neuronals convolucionals, imatges mèdiques, glioma, segmentació



# Índice general

<b>Índice general</b> .....	<b>6</b>
<b>Índice de ilustraciones</b> .....	<b>9</b>
<b>Índice de tablas</b> .....	<b>10</b>
<b>Agradecimientos</b> .....	<b>11</b>
<b>1 Introducción</b> .....	<b>13</b>
1.1 Motivación .....	13
1.2 Objetivos.....	14
1.3 Estructura de la memoria .....	15
<b>2 Estado del Arte</b> .....	<b>17</b>
<b>3 Conceptos previos</b> .....	<b>20</b>
3.1 Tumores cerebrales y gliomas .....	20
3.2 Imágenes por Resonancia Magnética (IRM) .....	22
3.3 ITK-SNAP: Herramienta de segmentación y visionado de IRM .....	26
<b>4 Metodología</b> .....	<b>30</b>
4.1 Redes neuronales artificiales .....	30
4.2 Redes neuronales convolucionales .....	32
4.3 Redes convolucionales <i>UNet</i> .....	34
4.4 Redes ya existentes .....	36
4.4.1 <i>X-Net</i> .....	36
4.4.2 <i>SegNet</i> .....	37
4.4.3 <i>Multires-UNet</i> .....	38
4.4.4 <i>R2U-Net</i> .....	39
<b>5 Marco experimental</b> .....	<b>42</b>
5.1 Descripción de los datos .....	42
5.2 Preprocesado, técnicas de aumento y regularización de datos .....	44
5.2.1 Normalización .....	44
5.2.2 <i>RandomRotation</i> .....	45
5.3 Hiperparámetros.....	45
5.3.1 Función de pérdida .....	45
5.3.2 Optimizador del planificador de aprendizaje.....	46
5.4 Métricas de evaluación.....	46
5.5 Tecnologías utilizadas.....	47

5.5.1 Equipo, sistema operativo, entorno de desarrollo y lenguaje de programación .....	47
5.5.2 Librerías utilizadas .....	47
<b>6 Resultados.....</b>	<b>50</b>
6.1 Resultados cuantitativos.....	50
6.1.2 <i>Dataset</i> ONCO_TCGA.....	51
6.2 Resultados cualitativos .....	51
6.2.1 <i>Dataset</i> ONCO_Ivy.....	52
6.2.2 <i>Dataset</i> ONCO_TCGA.....	53
<b>7 Conclusiones.....</b>	<b>55</b>
7.1 Futuro trabajo .....	56
7.2 Relación con los estudios cursados .....	56
<b>Bibliografía .....</b>	<b>58</b>



# Índice de ilustraciones

Ilustración 1 – Tumores más comunes por rango etario.....	22
Ilustración 2 – Comparativa de resolución entre dos imágenes tomadas con distintas tecnologías. A la izquierda mediante TC y a la derecha mediante IRM.....	23
Ilustración 3 – Representación sagital del cerebro sobre 6 voluntarios. Siendo a el volumétrico T1 y b T2.....	24
Ilustración 4 – Imagen axial usando la técnica T1-w.....	24
Ilustración 5 – Imagen IRM axial usando la técnica T2-w.....	25
Ilustración 6 – Imagen IRM axial de tipo <i>FLAIR</i> .....	25
Ilustración 7 – Imagen IRM axial de tipo T1c o T1gd.....	26
Ilustración 8 – ITK-SNAP (versión 3.6) con una IRM de tipo T1. ....	27
Ilustración 9 – ITK-SNAP (versión 3.6) con una IRM de tipo T1 con su correspondiente máscara de segmentación, divida en distintos colores según el tipo de tejido. ....	28
Ilustración 10 – Modelo de una neurona artificial. ....	30
Ilustración 11 – Funciones de activación más comunes.....	31
Ilustración 12 – Perceptrón multicapa con dos capas ocultas. Delimitada en verde, la capa de entrada; en rojo, las ocultas y en verde, la de salida. ....	32
Ilustración 13 – Convolución de un matriz de 5x5 con un filtro de 3x3 con un <i>stride</i> de 1.....	33
Ilustración 14 – Ejemplo de dos tipos de submuestreo. A la derecha <i>Average-Pooling</i> y a la izquierda <i>Max-Pooling</i> .....	34
Ilustración 15 – Arquitectura del modelo <i>UNet</i> . En azul las matrices de características. Las flechas de color representan las distintas operaciones y las grises el proceso de recortado. ....	35
Ilustración 16 – Arquitectura del modelo <i>X-Net</i> . Incluye de imagen de entrada y posterior imagen de segmentación de salida. ....	36
Ilustración 17 – Arquitectura del modelo <i>SegNet</i> . Incluye ejemplo de imagen de entrada y posterior segmentación de salida. ....	38
Ilustración 18 – Arquitectura <i>Multires-UNet</i> .....	38
Ilustración 19 – Visualización del camino de recortado <i>Res</i> . ....	39
Ilustración 20 – Arquitectura <i>R2U-Net</i> . ....	40
Ilustración 21 – Imagen IRM T1-gd con máscara de segmentación correspondiente al dataset ONCO_Ivy. En verde, tejido correspondiente a <i>edema</i> ; en azul a <i>enhancing tumor</i> y en rojo a <i>necrosis</i> . ....	43
Ilustración 22 – Ejemplo de <i>RandomRotation</i> , para -90° y 90° respectivamente.....	45
Ilustración 23 – Resultados comparativos entre máscara original y predicha por modelo <i>R2U-Net</i> para ONCO_Ivy. ....	52
Ilustración 24 – Resultados comparativos entre máscara original y predicha por el modelo <i>R2U-Net</i> para ONCO_TCGA. ....	53

# Índice de tablas

Tabla 1 – Tabla comparativa de resultados para <i>dataset</i> ONCO_Ivy. ....	50
Tabla 2 – Tabla comparativa de resultados para <i>dataset</i> ONCO_TCGA.....	51

# Agradecimientos

A mis padres y a mi hermana, por enseñarme con amor y esmero a ser una persona capaz y responsable, no los cambiaría por nada del mundo.

Agradecer otra vez más madre, Doctora en Química Sostenible por la UJI. Por ser una cotutora más que sin saber un ápice de informática, su inteligencia, planificación y experiencia han sido fundamentales en el desarrollo de este trabajo. Sin ella este TFG no hubiera sido posible.

A mis amigos de la universidad. A Javi, Asier, Cristóbal, Alejandro y Luis. Me llevo conmigo estos 4 años de risas, lágrimas, logros y fracasos. Sois una parte importante de mi vida y parte de como soy hoy en día es gracias a vosotros.

Y aunque parezca egoísta, a mí mismo. Por no dar marcha atrás, por seguir en momentos de duda y desesperación.

*<<Si luchas por tu propia voluntad, cualquier castigo resultante es de tu propia creación. Aceptarlos es parte de ese ideal. >> - Archer, Fate/Stay Night: Unlimited Blade Works.*



# Introducción

El rampante crecimiento y desarrollo de las Inteligencias Artificiales y las nuevas metodologías relacionadas con el *machine learning* han supuesto a nivel global un gran impacto y avance en multitud de sectores de la sociedad actual. Sin embargo, para este trabajo nos centraremos en un sector en concreto, el de la salud.

La implantación de tecnologías relacionadas con la Inteligencia Artificial ha supuesto un antes y un después en un sector tan importante y relevante como es el de la salud [1]. Claros ejemplos son los asistentes robóticos para cirugías o *robot-assisted surgery*, que posibilita a los cirujanos y profesionales médicos la posibilidad de realizar operaciones sin la necesidad de estar en la misma sala que el propio paciente. O los sistemas basados en reglas, complejas redes basadas en *pattern-matching* junto a reglas creadas específicamente por especialistas de distintas áreas que permiten, en un ejemplo concreto, a partir del historial clínico de un paciente y su diagnóstico decidir qué tratamiento es el más adecuado para su condición o si ya está bajo uno adaptarlo.

Sin embargo, los mayores esfuerzos llevados a cabo por la IA en el sector de la salud son realizados en la clasificación y asimismo segmentación de órganos y tejidos. Esto se lleva a cabo mediante la implementación de Redes Neuronales Convolucionales o *Convolutional Neural Networks* en inglés y técnicas de procesado de imágenes.

El tema de este trabajo de fin de grado se centra en el glioma, un tumor que se localiza en el cerebro. Mediante la segmentación se procederá a la diferenciación de los diferentes tejidos que componen al tumor. Es por ello por lo que se trata de una segmentación multiclase.

## 1.1 Motivación

Hoy en día el cáncer supone una de las enfermedades más mortales. Dentro de esta enfermedad, el glioblastoma multiforme es de los más comunes y agresivos de los tumores cerebrales primarios. Según reportes, alrededor de diez millares de estadounidenses son detectados con casos de glioblastomas (tumor cerebral similar al glioma de naturaleza más agresiva) [2]. De media, solo la mitad consigue sobrevivir un año y un promedio del 25% de ellos sobrevive a los dos años.

Para poder realizar distintos procesos médicos relacionados con este cáncer, como puede ser el seguimiento monitorizado del tumor, delimitación de los tejidos cancerígenos o intervenciones quirúrgicas se precisan de IRM (imágenes de resonancia magnética).

No obstante, el procesado de esas imágenes por parte de expertos y las consecuentes delimitaciones de las diferentes regiones suponen una tarea complicada y no exenta de errores. Por otra parte, debido a la naturaleza del glioma, es necesario la toma de distintas imágenes a lo largo de diferentes puntos temporales. Además, para la correcta localización de los distintos tejidos que conforman el glioma, se deben tener en cuenta diferentes canales de la imagen. Esto hace que la segmentación “tradicional” suponga una carga de trabajo extensa y laboriosa, la cual choca con la agresividad y el tiempo de esperanza de los pacientes diagnosticados con esta enfermedad.

Con los últimos avances en el ámbito tecnológico, en específico en el campo del aprendizaje automático se ha demostrado con grandes resultados la capacidad de resolución para problemas de distinta índole, incluyendo la salud. Es por ello, que el estudio y la puesta en práctica de estos modelos de manera eficaz puede suponer una detección temprana de la enfermedad, concepto clave para aumentar la posibilidad de supervivencia del paciente.

## 1.2 Objetivos

El principal objetivo de este trabajo consiste en la implementación de una herramienta que a través de distintos modelos consiga una correcta segmentación del tumor y sus distintos tejidos. Para ello se han desglosado los siguientes subobjetivos:

- Selección precisa de imágenes IRM para la distinción de los diferentes tejidos tumorales y correspondiente procesado.
- Implementación de distintas redes convolucionales para la segmentación multiclase de imágenes.
- Probar la calidad de las redes y su desempeño mediante distintas métricas y análisis de resultados.

## 1.3 Estructura de la memoria

El documento presente constatará de un total de 6 capítulos, los cuales se presentan y describen de la siguiente manera:

- En el **capítulo 2** se comentan de manera breve y concisa aproximaciones y resultados de distintos trabajos respecto al problema planteado.
- En el **capítulo 3** se procede a detallar conceptos teóricos necesarios para el entendimiento del trabajo.
- En el **capítulo 4** se ahonda en la metodología empleada. Los modelos basados en la arquitectura *UNet*, su funcionamiento y la estructura que presentan.
- En el **capítulo 5** se discute el marco experimental del trabajo. Se comentan los conjuntos de datos, especificaciones de hiperparámetros utilizados, métricas implementadas para evaluar los modelos y las tecnologías usadas para implementarlos.
- En el **capítulo 6** se visualizan los resultados de los modelos entrenados y se compara su desempeño en base a lo explicado en el capítulo anterior.
- En el **capítulo 7** se procede a concluir con el trabajo, se plantean trabajos futuros y se hace una relación de los conocimientos expuestos con aquellos adquiridos a lo largo de la carrera.



# Estado del Arte

La segmentación de imágenes juega un rol importante en el entendimiento de estas, la extracción de características y análisis en diversas aplicaciones. Su uso es variado en el campo de la salud por su utilidad en la clasificación de tejidos, detección de tumores o incluso el estudio volumétrico de estos. Para el caso de la segmentación de tumores cerebrales, la finalidad es separar los diferentes tejidos tumorales (*edema*, *enhancing core* y *necrosis*) de los tejidos sanos del cerebro: sustancia blanca, sustancia gris y líquido cefalorraquídeo [44].

En los últimos años, diferentes investigadores han propuesto diferentes esquemas de segmentación bajo distintas técnicas: basadas en umbrales, en límites y en regiones. El primer esquema estipula que los píxeles que se encuentran a cierta distancia pertenecen a una clase concreta [45]. El segundo estipula que los píxeles que se encuentran en los bordes cambian radicalmente en referente a sus características [46]. La última especifica que las regiones se componen de píxeles aledaños con características similares [47].

Dentro de la segmentación supervisada (aquella que emplean datos etiquetados en la construcción del modelo), se han experimentado con métodos sencillos, entre los que destacan:

- *k-Nearest Neighbour*: esta técnica supervisada basada en memoria compara nuevas muestras de datos con una colección de datos ya clasificados. La elección de la clase de la nueva muestra se decide a partir del voto de los  $k$  vecinos más cercanos a esta. Este método se usó en el trabajo [48] para segmentar anomalías claras y oscuras entre los valores medios y bajos de gris en imágenes IRM de tipo *FLAIR* (explicado en el apartado 3.2)
- Máquinas de vectores de soporte: este modelo divide las imágenes en dos clases al encontrar un hiperplano que mejor se ajuste a la división de estas dos clases. Un estudio [49] empleó métodos de dependencia espacial de materia gris para extraer características que, tras una criba por parte de un algoritmo genético, se utilizan para alimentar una máquina de vectores de soporte para la clasificación de tejidos cerebrales entre normales, benignos o malignos.

No obstante, las aproximaciones realizadas en los últimos años para esta tarea se han centrado alrededor de las redes neuronales artificiales (explicadas en el apartado 4.1). Una de las redes más simples empleadas es la *FCNN* (*Fully Convolutional Neural Network*). Este tipo de red es capaz de segmentar patrones no lineales y aproximar funciones continuas. Asimismo, para facilitar su

entrenamiento se disponen de algoritmos que facilitan la tarea como es el caso del algoritmo de propagación hacia atrás (explicado en el apartado [4.1](#)). En este estudio [\[50\]](#) se implementa una *FCNN* para la detección eficiente de tumores cerebrales.

Por último, en la actualidad la arquitectura que sienta la base para la segmentación de imágenes médicas es *UNet* (explicada en el apartado [4.3](#)). En los últimos años, la cantidad de trabajos que implementan esta arquitectura o similares ha crecido exponencialmente [\[24\]](#). No obstante, se siguen ofreciendo variantes basadas en el modelo base que buscan mejorar las prestaciones ya obtenidas, siendo ejemplos los modelos *X-Net*, *SegNet* o *R2U-Net* (explicados en el apartado [4.4](#))



# Conceptos previos

En este tercer capítulo se procede a explicar y tratar los diferentes aspectos teóricos y las tecnologías empleadas a lo largo de este trabajo. Con ello se pretende dar una explicación completa y amena para que el lector entienda en su totalidad el desarrollo del proyecto. Este apartado contará con tres bloques principales: tumores cerebrales y definición de gliomas, imágenes IRM y por último tecnologías de segmentación y visionado de imágenes, en concreto, la aplicación ITK-Snap.

### 3.1 Tumores cerebrales y gliomas

En la sociedad actual, el cáncer supone uno de los factores de mortalidad más importantes en cuanto se refiere a enfermedades. Según informes de la IARC [3], un total de 8.2 millones de defunciones en 2012. Sin embargo, las predicciones auguran un aumento de los casos en un 70% para los próximos 20 años.

Dentro de esta enfermedad se encuentran los tumores cerebrales, afectando principalmente al sistema nervioso central. Dado que se trata de un tumor de carácter irregular en cuanto a grupo, se dividen en dos secciones principales; los de carácter primario (lesiones primarias), originados por las propias células pertenecientes al sistema nervioso central y las de lesiones secundarias, cuyo origen se encuentra en otra parte del cuerpo y se extienden hasta el cerebro mediante el proceso de la metástasis, donde el cual las células cancerígenas obtienen la capacidad de proliferar a otras partes del cuerpo (tanto tejidos como órganos) desde el tumor principal a través de canales como el sistema circulatorio (tanto sangre como linfa) [4]. De los tumores primarios destacan por frecuencia el glioblastoma, el tumor de grado alto más frecuente, cuyo rango de supervivencia para los casos detectados varía entre uno y cinco años [7].

En referencia a los tumores primarios del sistema nervioso central (TPNSC), la Organización Mundial de la Salud creó en el año 1979 un sistema de discriminación teniendo en cuenta dos factores; en primera instancia la célula primigenia que origina el proceso tumoral y por otra parte las características morfológicas en cuanto al pronóstico de su desarrollo. En base a estas métricas, se crearon un total de 4 grados de distinción [5]:

1. Grado I: Tumores de lento crecimiento y cuya posibilidad de evolución a uno de mayor malignidad es baja. También se caracterizan por ser circunscritos.

2. Grado II: Tumores de patrón difuso. Mayor posibilidad de evolución a tumor maligno, pero con crecimiento lento.
3. Grado III: Tumores de mayor riesgo, con localización de células atípicas y en algunos casos anaplásicas [6] (lo cual dificulta la diferenciación de células de un tumor). Asimismo, presentan un alto grado de mitosis.
4. Grado IV: Tumores caracterizados por rapidez de crecimiento. Mayor tasa de mitosis, incluso llegando a presentar vasos de neoformación o áreas necróticas.

A través de esta clasificación, los tumores de grado I y II (como son los gliomas o meningiomas) se catalogan como “benignos” o “bajo grado”. Los de carácter I son los más fáciles a tratar, pues mediante un proceso quirúrgico pueden ser sanados los pacientes. De otra manera, los de grado II presentan una mayor recurrencia y posibilidad de entrar en grado III o IV.

Los tumores de grado III y IV son los más complejos y difíciles de tratar. Ambos grados presentan un mayor historial de malignidad y los de tipo IV se caracterizan por la aparición de zonas necróticas (como es el glioblastoma). Esto se traduce en un rápido desarrollo y nula esperanza de supervivencia. Es por ello por lo que se clasifican como “alto grado” o “malignos”.

Aunque la mayoría de los tumores cerebrales son clasificados como benignos (66%), existen casos de tumores no histológicamente confirmados como benignos. Esto hace que el porcentaje de malignos llegue a aumentar hasta más del 44%. En otros enfoques, los hombres tienen mayor tasa de detección de tumores cerebrales de carácter maligno (55%) mientras que las mujeres poseen un mayor porcentaje de tumores, pero en este caso de carácter benigno (64%) [7].

Por último, con relación a la edad con la detección de tumores cerebrales es más frecuente en personas mayores de 85 años y menor en casos de personas de 0 a 14 años, tal y como puede observarse en la ilustración 1.

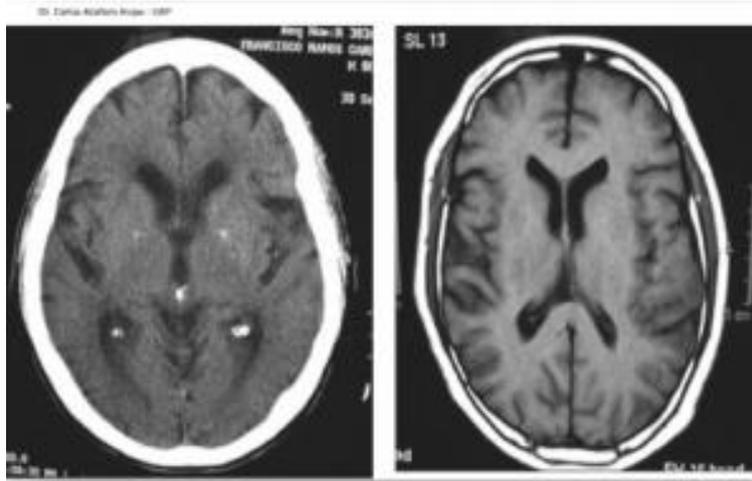
Edad	HISTOLOGÍA MÁS COMÚN		SEGUNDA HISTOLOGÍA MÁS COMÚN		TERCERA HISTOLOGÍA MÁS COMÚN		CUARTA HISTOLOGÍA MÁS COMÚN	
	HISTOLOGÍA	INCIDENCIA*	HISTOLOGÍA	INCIDENCIA*	HISTOLOGÍA	INCIDENCIA*	HISTOLOGÍA	INCIDENCIA*
0-4	Tumor Embrionario	1.24	Astrocitoma Pilocítico	1.03	Glioma Maligno	0.93	Tumor ependimario	0.48
5-9	Astrocitoma Pilocítico	1.01	Glioma Maligno	0.88	Tumor Embrionario	0.72	Tumor mixto glioneural	0.31
10-14	Astrocitoma Pilocítico	0.85	Glioma Maligno	0.51	Tumor pituitario	0.49	Tumor mixto glioneural	0.47
15-19	Tumor pituitario	1.66	Astrocitoma Pilocítico	0.6	Tumor mixto glioneural	0.48	Schwannoma	0.35
20-34	Tumor pituitario	3.11	Meningioma	1.39	Schwannoma	0.83	Astrocitoma G II	0.49
35-44	Meningioma	4.82	Tumor pituitario	4.36	Schwannoma	1.81	Glioblastoma	1.21
45-54	Meningioma	9.02	Tumor pituitario	4.64	Glioblastoma	3.54	Schwannoma	2.85
55-64	Meningioma	14.77	Glioblastoma	8.08	Tumor pituitario	5.37	Schwannoma	4.01
65-75	Meningioma	25.96	Glioblastoma	13.05	Tumor pituitario	7.30	Schwannoma	4.55
75-84	Meningioma	38.7	Glioblastoma	15.24	Tumor pituitario	7.32	Neoplasia no especificada	4.58
84 y más	Meningioma	51.31	Neoplasia no especificada	10.91	Glioblastoma	9.12	Tumor pituitario	4.69
Total	Meningioma	7.86	Tumor pituitario	3.49	Glioblastoma	3.20	Schwannoma	1.76

ILUSTRACIÓN 1 - TUMORES MÁS COMUNES POR RANGO ETARIO. FUENTE: [7]

### 3.2 Imágenes por Resonancia Magnética (IRM)

Las Imágenes por Resonancia Magnética se han convertido en un estándar en el tratamiento de distintas patologías en diversos ámbitos de la salud, siendo un ejemplo el ámbito neurológico. Al ser un proceso no invasivo que permite obtener imágenes de alta calidad que habilitan poder observar los distintos tejidos blandos y órganos del paciente y a su vez no emplea radiación ionizante (capaz de alterar la estructura de los átomos), se suele priorizar por otras técnicas de índole quirúrgica.

En el campo de la neurología, las IRM proporcionan en general mejores resultados en relación con resolución que las Tomografías Computarizadas (TC) a la hora de observar estructuras nerviosas. Entre otras, las IRM son capaces de captar edemas cerebrales, tumores e incluso infartos cerebrales. Al estar relacionadas estas patologías con las alteraciones de la masa de agua, cualquier cambio en ese volumen es vital para detectarlos.



*ILUSTRACIÓN 2 - COMPARATIVA DE RESOLUCIÓN ENTRE DOS IMÁGENES TOMADAS CON DISTINTAS TECNOLOGÍAS. A LA IZQUIERDA MEDIANTE TC Y A LA DERECHA MEDIANTE IRM. FUENTE: [8]*

Para la obtención de estas imágenes, los átomos de hidrógeno son fundamentales. En primera instancia, porque abundan dentro del cuerpo humano, el cual está conformado por un 70% de agua. Y en segundo lugar porque posee la estructura atómica más simple, un protón (carga positiva) en el núcleo y un electrón (carga negativa) en la corteza [8].

Dentro del cuerpo humano, los átomos de hidrógeno se encuentran en libre albedrío. Sin embargo, al ser expuestos a un campo magnético, los protones absorben esta energía y cambian su rotación. Al apagarse el campo, los protones comienzan a emitir energía para volver de nuevo a su rotación original en un proceso denominado precesión. Cuando esto ocurre, los protones emiten ondas de radiofrecuencia que son captadas por receptores de radiofrecuencia y que luego son usadas para formar la imagen.

Dentro de este contexto, las técnicas T1 y T2 simbolizan los tiempos de relajación a la hora de escanear tejidos entendido como un intervalo entre secuencia de pulsos. Por un lado, T1 controla el ratio en el que los protones de hidrógeno vuelven de nuevo a su rotación original, mientras que T2 discrimina el tiempo en el que el protón alcanza el estado de equilibrio o se establece a distintas fases de tiempo.

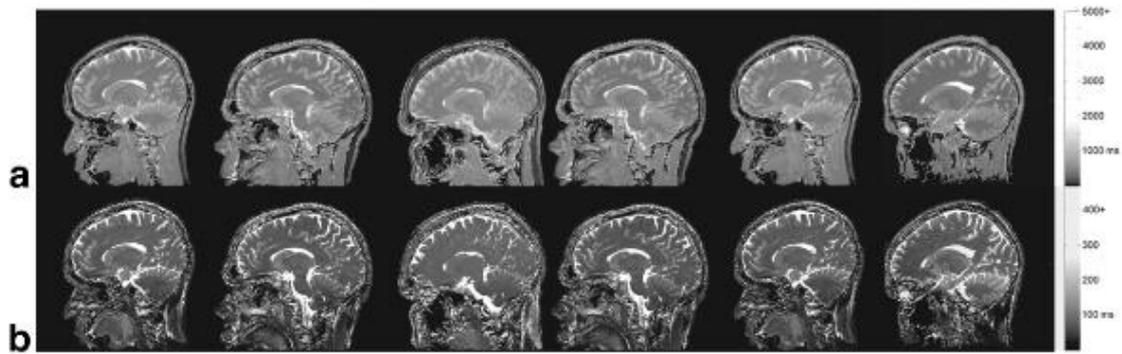


ILUSTRACIÓN 3- REPRESENTACIÓN SAGITAL DEL CEREBRO SOBRE 6 VOLUNTARIOS. SIENDO A EL VOLUMÉTRICO T1 Y B T2. FUENTE: [9]

De estas dos primeras técnicas se extienden otras, entre ellas *T1-weighted* y *T2-weighted*. Las imágenes de tipo *T1-weighted* requieren un menor tiempo de eco (TE) y también de repeticiones (TR), mientras que T2 toma más tiempo de ambas. Estas dos se refieren al periodo de duración entre las ondas de radiofrecuencia a las que se someten los protones y la recepción del eco producido y por otro lado al tiempo medido entre pulsos secuenciales aplicados a un área específica.

Las imágenes de tipo *T1-weighted* hace que los elementos se destaquen, tanto apareciendo como más luminosos u oscuros. Entre los elementos que aparecen iluminados son la sangre, la grasa o líquidos ricos en proteínas entre otros. Del otro grupo entre algunos son los edemas, los tumores o tendones, por mencionar algunos.



ILUSTRACIÓN 4 – IMAGEN AXIAL USANDO LA TÉCNICA T1-w. FUENTE: [10]

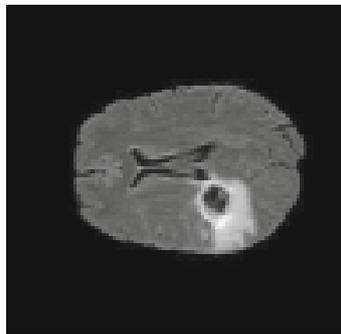
Por similitud, *T2-weighted* también destaca diferentes elementos, oscureciéndolos o iluminándolos. Algunos de los elementos que oscurece son el calcio, aire, tendones o ligamentos. Por otro lado, aquellos que son iluminados incluye el páncreas, fluidos o músculos.



*ILUSTRACIÓN 5 – IMAGEN IRM AXIAL USANDO LA TÉCNICA T2-w. FUENTE [10]*

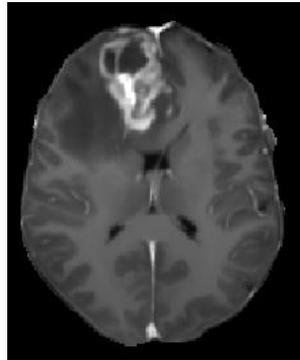
Aparte de estos dos tipos de imágenes, también existen varias bastante útiles las cuales serán utilizadas en este trabajo.

La primera se denomina por sus siglas en inglés *FLAIR* (*Fluid-Attenuated Inversión Recovery*). La característica de este tipo de imagen es la capacidad de suprimir los efectos del líquido cefalorraquídeo y a su vez reducir el contraste entre la sustancia gris y la blanca.



*ILUSTRACIÓN 6 - IMAGEN IRM AXIAL DE TIPO FLAIR- FUENTE: [11]*

Por último, se encuentra las imágenes de tipo T1gd o T1-c. Como su nombre indica son imágenes de tipo T1 a las que se les mejora el contraste mediante sustancias basadas en gadolinio, las cuales se inyectan a los pacientes.



*ILUSTRACIÓN 7 – IMAGEN IRM AXIAL DE TIPO T1C O T1GD. FUENTE: [33]*

En relación con las IRM, aunque parezca una técnica milagrosa, no está exenta de fallos y aspectos perjudiciales. Hay que entender que el cuerpo humano está compuesto de diferentes órganos y tejidos. Por lo tanto, para un caso de visualización puede resultar a veces tedioso y complicado. Muchas veces otros tejidos y órganos bloquean la visión de aquello que se quiere capturar en la IRM. Es por ello por lo que diversas secuencias IRM han de ser empleadas para conseguir una imagen general de calidad y detallada. En segundo lugar, existen riesgos relacionadas con las imágenes de contraste. Las sustancias basadas en el gadolinio son muy tóxicas y por ende han de ser tratadas previamente para que posteriormente el organismo pueda expulsarlo.

### 3.3 ITK-SNAP: Herramienta de segmentación y visionado de IRM

Dentro del campo de análisis de imágenes médicas, uno de los principales retos es la segmentación. La elección de una herramienta cómoda y correcta para trabajar con las imágenes es crucial. En este caso se ha decantado por la aplicación ITK-SNAP.

ITK-Snap [12] es una herramienta multipropósito para la visualización de imágenes, capacidad de segmentación manual y semiautomática mediante métodos de contorneado activo. Creada alrededor de los 2000, fue concebida como una aplicación que fuese intuitiva y sencilla para usuarios no tan expertos con los ordenadores y con una visión centralizada al 100% en la segmentación de imágenes.

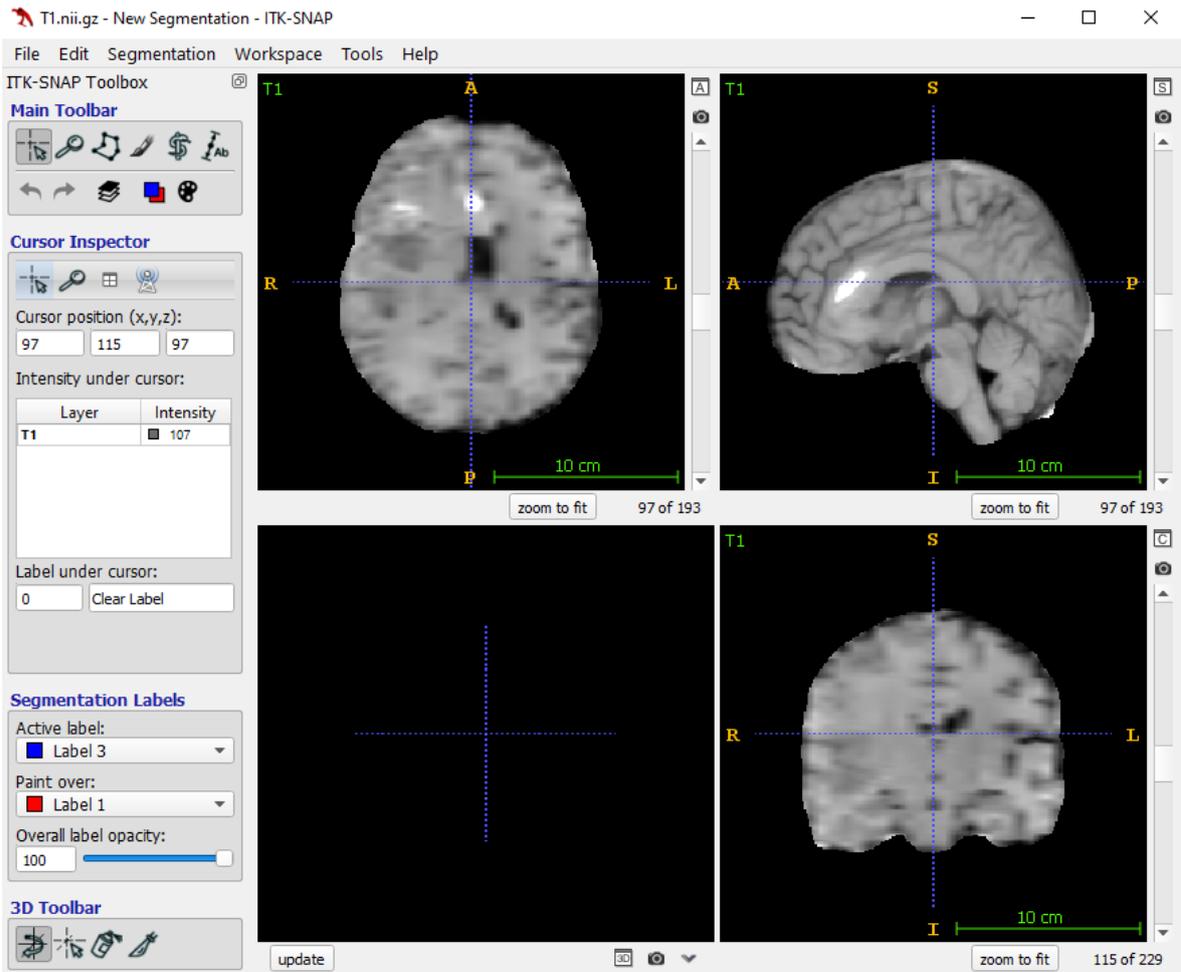
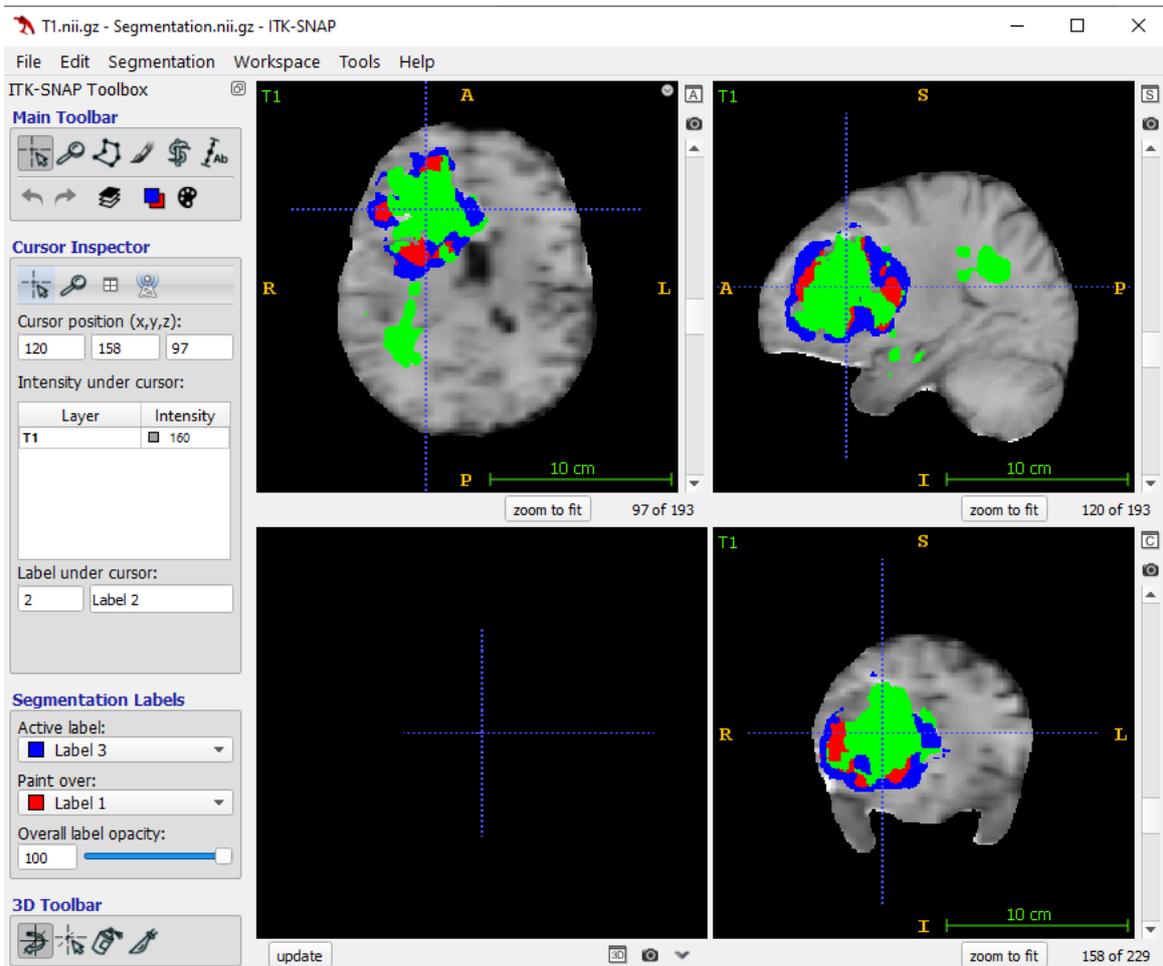


ILUSTRACIÓN 8 – ITK-SNAP (VERSIÓN 3.6) CON UNA IRM DE TIPO T1. FUENTE: ELABORACIÓN PROPIA

Desde su creación, ITK-SNAP se ha posicionado como una de las herramientas más usadas por la comunidad médica. Esto se debe a que, aparte de las funcionalidades mencionadas anteriormente, ITK-SNAP presenta también las siguientes características:

- Soporte para distintos tipos de imágenes, entre ellas 3D (como por ejemplo el formato NifTI).
- Soporte para color, multicanales e imágenes de dependencia temporal.
- Cursor vinculado para una navegación 3D amena.



*ILUSTRACIÓN 9 – ITK-SNAP (VERSIÓN 3.6) CON UNA IRM DE TIPO T1 CON SU CORRESPONDIENTE MÁSCARA DE SEGMENTACIÓN, DIVIDA EN DISTINTOS COLORES SEGÚN EL TIPO DE TEJIDO. FUENTE: ELABORACIÓN PROPIA.*

Para el desarrollo de este trabajo, el uso de ITK-SNAP ha sido fundamental en los primeros momentos. Ha servido primer contacto para el visionado de las IRM, entendimiento del formato NifTI y sobre todo para comprender en primera instancia el reto que supone la correcta segmentación de estas imágenes.



# Metodología

Para este cuarto capítulo se procede a explicar los modelos para segmentación de imágenes que posteriormente se utilizan, haciendo hincapié en su arquitectura. Se procede en primer lugar a dar una primera base sobre qué son las redes neuronales artificiales, explicando modelos más sencillos, para pasar posteriormente a comentar las redes neuronales convolucionales. Posteriormente se detalla la arquitectura *UNet*, sus principios y detalles y por último se describen modelos que nacen como variantes de esta arquitectura.

## 4.1 Redes neuronales artificiales

En el ámbito de la inteligencia artificial existen infinitudes de redes neuronales capaces de resolver problemas que aparecen en distintos ámbitos de la sociedad. En su definición, las redes neuronales artificiales son modelos computacionales que permiten de una manera u otra recrear el comportamiento racional del humano.

La primera red básica que se puede pensar es la neurona artificial. Como indica su nombre, esta funciona similarmente a una neurona del sistema nervioso. El modus operandi consiste en: recibe información, la procesa y posteriormente la transmite. La neurona sienta la base para la creación de posteriores modelos más complejos y desarrollados.

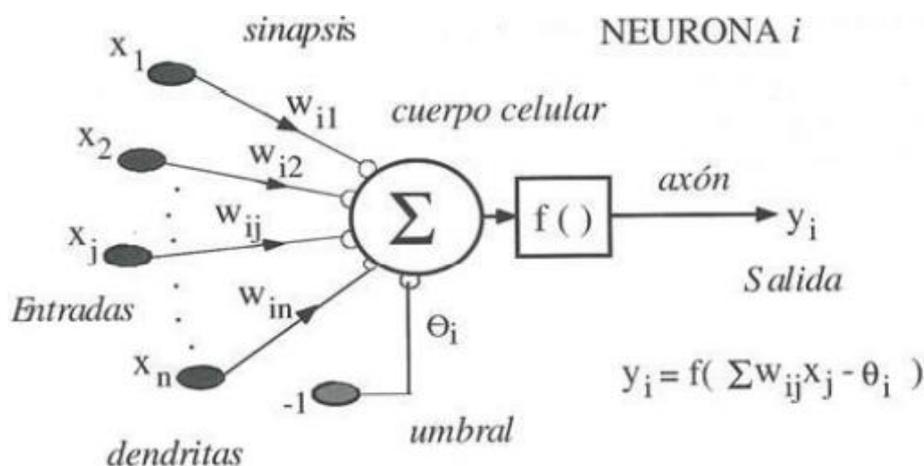


ILUSTRACIÓN 10 – MODELO DE UNA NEURONA ARTIFICIAL. FUENTE: [13]

Como se puede ver en la ilustración 10, el modelo se define de la siguiente forma:

- La neurona cuenta con un total de  $m$  entradas.
- Cada entrada tiene asociada un peso  $w_i$ , tal que  $i > 0$ .
- La neurona cuenta con un peso umbral  $\theta_i$ .
- Una función de activación  $f$ .
- Y finalmente, una salida  $y_i$ .

Al final, la neurona realiza el siguiente cálculo matemático:

$$y_i = \sum_{j=1}^n w_{ij}x_j - \theta_i$$

Por norma general, la función de activación  $f()$  se considera determinista y asimismo monótona y continua. Este modelo, denominado también Perceptrón simple, se idea en 1957. Aun así, en el año 1969 se demuestra matemáticamente que este modelo solo sirve para problemas linealmente separables.

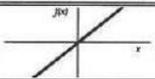
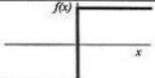
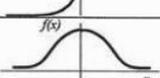
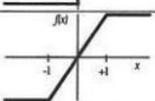
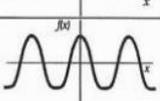
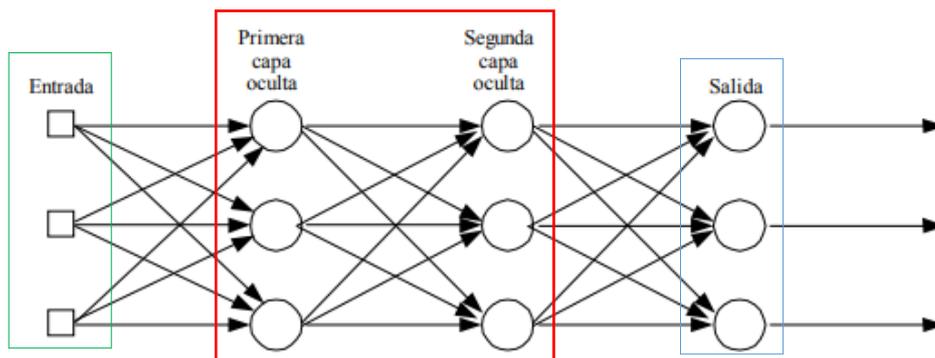
	Función	Rango	Gráfica		Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$		Sigmoidea	$y = \frac{1}{1+e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$		Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } +l \leq x \leq -l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$		Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	

ILUSTRACIÓN 11 – FUNCIONES DE ACTIVACIÓN MÁS COMUNES. FUENTE: [13]

Como solución al impedimento que presenta el modelo del Perceptrón simple se propone un nuevo modelo, el Perceptrón multicapa. Este modelo no es más que una sucesión de neuronas conectadas entre sí y distribuidas entre distintas capas. Entre las susodichas capas, existe una distinción que cabe mencionar:

- Capa de entrada: Forma la primera parte del modelo. Aquí las neuronas reciben los datos desde las entradas, pero sin realizar ninguna operación.

- Capas de salida: Son aquel conjunto de neuronas cuya salida representa la salida formal del modelo.
- Capas ocultas: comprendidas entre la capa de entrada y de salida, las neuronas apostadas reciben datos de aquellas de la capa anterior, realizan los cálculos indicados y propagan la salida a la siguiente capa de neuronas. Aquí es donde se realiza la gran parte de los cálculos del Perceptrón multicapa.



*ILUSTRACIÓN 12 – PERCEPTRÓN MULTICAPA CON DOS CAPAS OCULTAS. DELIMITADA EN VERDE, LA CAPA DE ENTRADA; EN ROJO, LAS OCULTAS Y EN VERDE, LA DE SALIDA. FUENTE: [15]*

Este modelo con una configuración correcta y funciones de activación convenientes es capaz de ser aplicado a problemas no linealmente separables. Sin embargo, la desventaja que presentaban en aquel momento era que debían ser entrenados manualmente. En otras palabras, no existía un algoritmo que entrenase al modelo automáticamente. No es hasta el año 1986 que Hinton, Rumelhart y Williams idean el algoritmo de retropropagación o propagación hasta atrás [16]. Esto permitió la posibilidad de automatizar el proceso de entrenamiento de estas redes.

## 4.2 Redes neuronales convolucionales

Las redes neuronales convolucionales representan un tipo de redes neuronales artificiales la cual imita el procesamiento de las neuronas de la corteza primaria del cerebro [18] con la finalidad de detectar diferentes características de los datos de entrada. Los principios de las redes neuronales convolucionales se sustentan en *Neocognitron* [19], ideada por el japonés Kunihito Fukushima, siendo mejorado posteriormente por Yann LeCun [20] para introducir un método que entrenase correctamente mediante el algoritmo de propagación hacia atrás.

La estructura principal de este tipo de modelos se configura a base de diferentes capas ocultas que siguen una jerarquía. Esto implica que las primeras capas son capaces de reconocer patrones o características simples, mientras que cuanto más se profundiza las capas más se especializan, llegando a reconocer patrones más complejos como la forma de un órgano o la silueta de un tipo de tejido cancerígeno, en el caso de imágenes médicas.

Un concepto también importante a la hora de tratar este tipo de redes es la convolución. La convolución se trata de un conjunto de operaciones matriciales entre la matriz de entrada (en nuestro caso de una imagen) y una segunda matriz denominada *kernel* o filtro. Durante este proceso, el filtro se mueve de izquierda a derecha y de arriba hacia abajo determinando cuantas posiciones debe moverse mediante un parámetro denominado *stride*. Como resultado, se obtiene una nueva matriz convolucionada. Después de este proceso se le aplica una función de activación, normalmente siendo *ReLU* (función definida como  $F(x) = \max(0, x)$ ).

Aunque en un principio las redes neuronales convolucionales fueron diseñadas para imágenes en 2 dimensiones, los avances han permitido que se puedan aplicar tanto 1 como 3 dimensiones. En la ilustración 13 se observa para un ejemplo en 2 dimensiones.

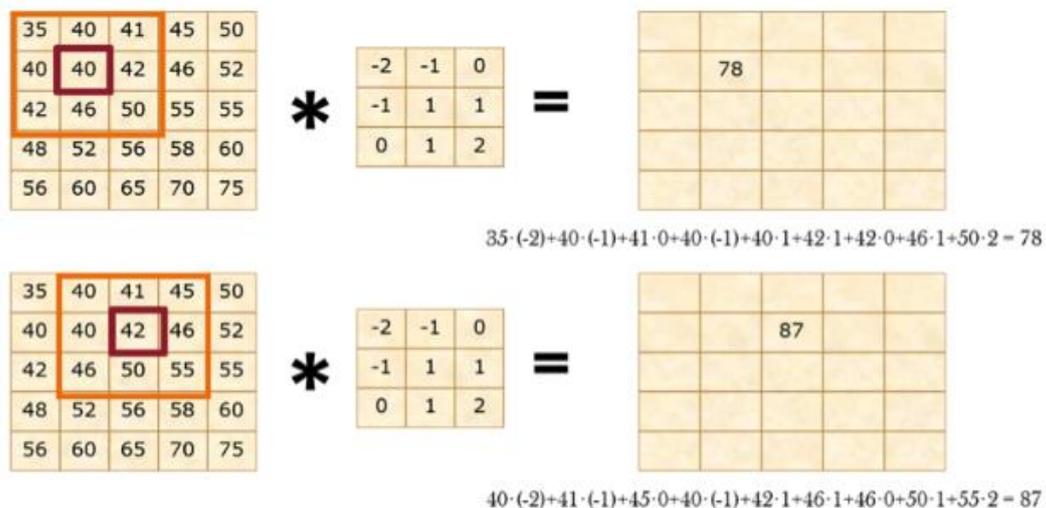


ILUSTRACIÓN 13 – CONVOLUCIÓN DE UN MATRIZ DE 5X5 CON UN FILTRO DE 3X3 CON UN STRIDE DE 1. FUENTE: [21]

La finalidad del proceso de convolución es permitir descubrir características ocultas y, aunque parezca contradictorio, representar composiciones completas mediante representaciones más simples. Esto se debe a que las redes convolucionales profundas a medida que aumenta la profundidad de la imagen aumenta la carga semántica de esta. Es decir, tal y como se explicó que estas redes seguían una jerarquía de capas, esto

también se traduce que, a mayor número de capas ocultas, mayor será la capacidad de construir patrones más complejos mediante representaciones más sencillas.

Por último, tras realizar el proceso de convolución se aplica un componente de submuestreo o *sub-sampling*, que en definitiva implica reducir el número de datos conforme se avanza por las distintas capas de la red. Aunque se pierdan datos durante este proceso, se asegura que las características más relevantes prevalezcan.

Aunque hay distintos tipos de submuestreo, las dos variantes más usadas son *Max-Pooling* y *Average-Pooling*. El primero discrimina a partir el valor más alto de la ventana de tamaño X y la segunda calcula la media de los datos dentro de la ventana.

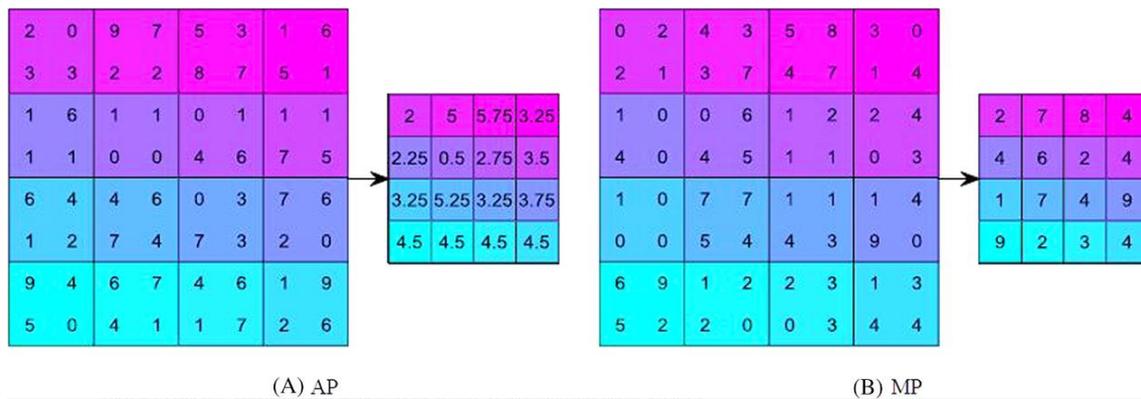


ILUSTRACIÓN 14 – EJEMPLO DE DOS TIPOS DE SUBMUESTREO. A LA DERECHA AVERAGE-POOLING Y A LA IZQUIERDA MAX-POOLING. FUENTE: [23]

### 4.3 Redes convolucionales UNet

En su definición, *UNet* es una técnica de segmentación de imágenes creada con el mismo propósito. La arquitectura *UNet* se basa en redes neuronales y su estructura básica consta de dos caminos bien diferenciados, adoptando una forma de “u” que es lo que le da el nombre. La primera es denominada contracción y se asemeja a una red neuronal convolucional. Cada bloque de este camino se caracteriza por dos sucesivas convoluciones de 3x3 a las que se les aplica una activación de tipo *ReLU* y una capa de submuestreo *Max-Pooling*. Este proceso es repetido varias veces. La novedad de esta arquitectura viene en el segundo camino expansivo. En cada paso de este camino se realiza una decodificación la matriz de características mediante una convolución traspuesta de 2x2. Tras esto, la matriz de características se recorta y se propaga a otra matriz de características. Después de ello se realiza otra doble convolución de 2x2 y una

reactivación *ReLU*. Al final del modelo, se realiza una última convolución de 1x1 para reducir la matriz de características al número de canales especificados y finalmente se produce la imagen segmentada.

Aunque parezca ilógico, el proceso de recortado y propagación es esencial dado que los píxeles de características que conforman los bordes poseen la menor cantidad de información posible y por ende deben ser descartados.

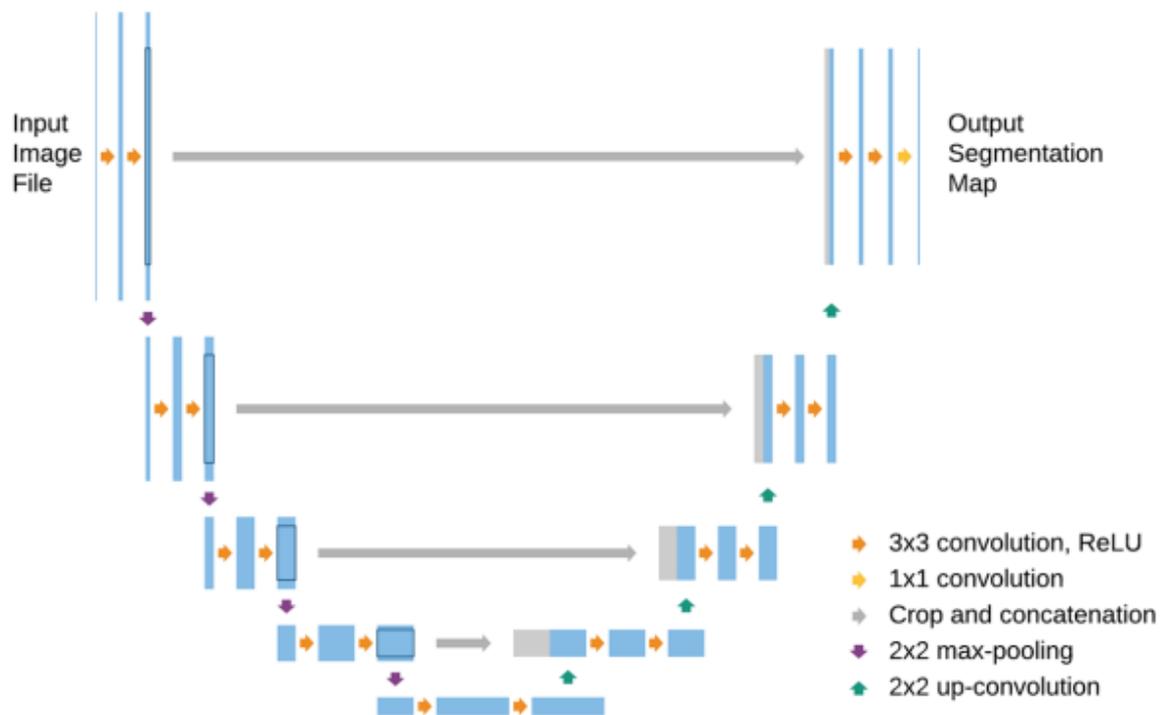


ILUSTRACIÓN 15 – ARQUITECTURA DEL MODELO UNET. EN AZUL LAS MATRICES DE CARACTERÍSTICAS. LAS FLECHAS DE COLOR REPRESENTAN LAS DISTINTAS OPERACIONES Y LAS GRISES EL PROCESO DE RECORTADO. FUENTE: [24]

Las arquitecturas *UNet* y sus derivadas son excelentes aproximaciones al problema de la segmentación de imágenes médicas por los problemas de escasez de datos y el desbalanceo de clases. Asimismo, también presenta una mayor facilidad para trabajar con imágenes en 3 dimensiones (como ocurre en este trabajo). *UNet* cuenta con el modelo *3D-UNet*, que básicamente sustituye todas las operaciones de convolución, convolución traspuesta y *Max-Pooling* a sus respectivas operaciones en 3D.

## 4.4 Redes ya existentes

Como se puede entender, concebir desde 0 una red neuronal convolucional no es una tarea fácil, pues no existe un número correcto de capas ocultas, el tipo de estas o las conexiones que debe haber. Por esto en muchos casos se recurre a la implementación de modelos ya existentes con resultados precisos y concluyentes probados en distintos conjuntos de datos. Este último subapartado pretende mostrar alguno de estos modelos, sus particulares y arquitecturas.

### 4.4.1 X-Net

X-Net es un diseño de modelo basado en arquitecturas de redes neuronales convolucionales propuesta en 2019 [25]. El nombre del modelo viene por el uso que se le había propuesto, segmentar imágenes médicas de rayos X. La arquitectura se basa en la tupla codificador-decodificador, común en las arquitecturas de segmentación de imágenes.

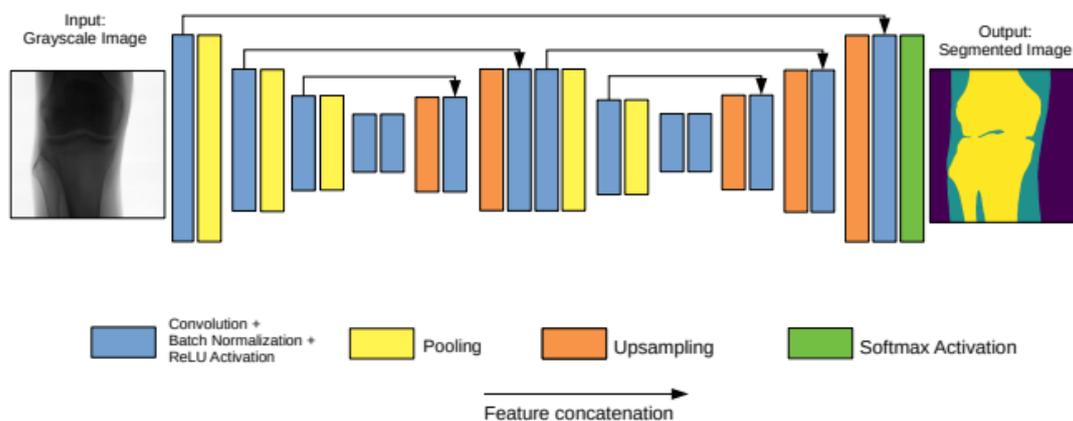


ILUSTRACIÓN 16 – ARQUITECTURA DEL MODELO X-NET. INCLUYE DE IMAGEN DE ENTRADA Y POSTERIOR IMAGEN DE SEGMENTACIÓN DE SALIDA. FUENTE: [25]

El codificador se compone de distintas capas convolucionales para extracción de características y *Max-Pooling* para reducción de tamaño de la imagen. El decodificador por otro lado realiza convoluciones invertidas para crear la imagen con la segmentación, ajustándose a las dimensiones de la imagen de entrada.

El uso de distintas capas de convolución traspuesta con consecuentes capas de convolución posibilita la variación de distintos grados de reconstrucción de características de grano fino. Asimismo, las concatenaciones de características de dimensiones iguales permiten evitar la pérdida de detalles de grano fino.

El modelo, entrenado para un conjunto de 150 imágenes obtenidas mediante Rayos X para la segmentación de tejido óseo y blanco, obtuvo un F1-Score del 0.87 y 0.90 respectivamente.

#### **4.4.2 SegNet**

En 2015, Vijay Badrinarayanan, Alex Kendall, y Roberto Cipolla proponen un novedoso modelo basado en redes neuronales convolucionales, *SegNet* [26]. La motivación para crear este modelo surge a partir de la baja resolución que producen el *Max-Pooling* y el submuestreo en los mapas de características. *SegNet* se proponía como una arquitectura rápida y precisa para segmentación precisa de píxeles (*pixel-wise*). Este modelo serviría, por ejemplo, para aplicaciones que requieran el procesado de la actividad de las calles. Estas requieren la capacidad de modelar la apariencia, forma de distintos elementos (coches, peatones, calles) mientras que a su vez entiende la relación espacial entre las distintas clases.

La arquitectura de *SegNet* presenta una red de codificación y una de decodificación. La primera se compone de 13 capas de convolución. A diferencia de otros modelos, en esta se elige omitir la conexión total de capas a favor de mantener una mayor resolución del mapa de características a la salida de la capa más profunda del codificador, mientras que a su vez reduce el número de parámetros del modelo (de 124M a 14.7M). Al igual que el codificador, el decodificador cuenta con 13 capas de convolución traspuesta. La salida del decodificador finalmente se pasa por una capa de clasificador multiclase *softmax*.

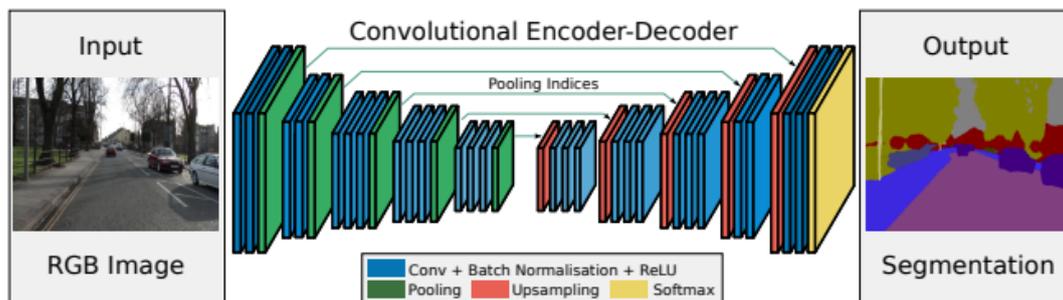


ILUSTRACIÓN 17 – ARQUITECTURA DEL MODELO SEGNET. INCLUYE EJEMPLO DE IMAGEN DE ENTRADA Y POSTERIOR SEGMENTACIÓN DE SALIDA. FUENTE: [26]

El modelo, probado con un conjunto de imágenes de calles, obtuvo excelentes resultados mejorando incluso de aquellos propuestos por modelos de redes neuronales convolucionales similares.

#### 4.4.3 Multires-UNet

*Multires-UNet* [27] nace como un nuevo modelo como propuestas de implementaciones que podrían aplicarse al modelo básico de *UNet* de la mano de Nabil Ibtehaz y M. Sohel Rahman en el año 2019.

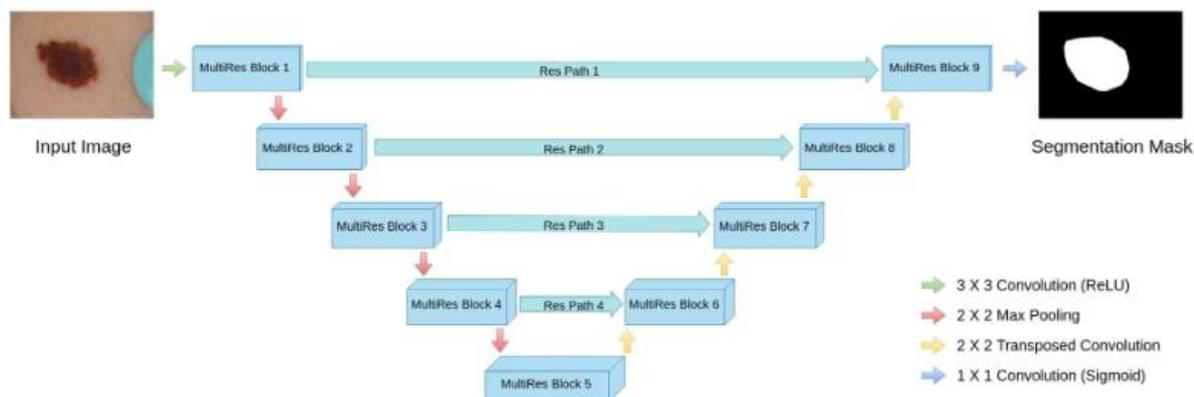


ILUSTRACIÓN 18 – ARQUITECTURA MULTIRES-UNET. FUENTE: [27]

El primer cambio que proponen es la introducción de los bloques *Multires* en sustitución de la secuencia de dos capas de convolución de 2x2. Estos nuevos bloques tienen

asignados cada uno un parámetro  $W$  que indica los números de filtros de las capas de convolución de cada bloque, de tal manera que:

$$W = \alpha * U$$

- $U$  correspondiendo al número de filtros en cada capa del modelo *UNet*.
- $\alpha$  como un parámetro escalar.

Aparte de la introducción de los nuevos bloques, otra característica que añaden como modificación es el camino de recortado, proponiendo unos nuevos denominados recortados *Res*.

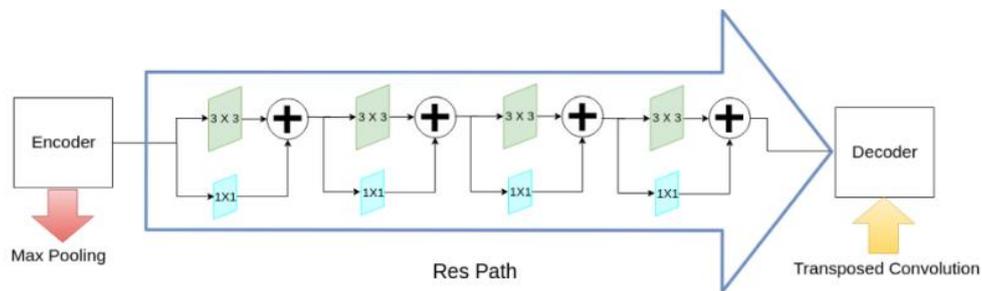


ILUSTRACIÓN 19 – VISUALIZACIÓN DEL CAMINO DE RECORTADO RES. FUENTE: [27]

A diferencia del modelo *UNet*, en vez de propagar el mapa de características del codificador directamente con el del decodificador, se propone pasar antes por una serie de capas de convolución. Esto, junto a otras operaciones no lineales permiten reducir la diferencia semántica entre las características del codificador y decodificador. Asimismo, conforme más profunda es la capa, se reducen el número de capas convoluciones, al ir disminuyendo progresivamente la diferencia semántica entre el codificador y decodificador.

Este modelo atacó el problema que implicaban las imágenes médicas, donde en muchos casos las segmentaciones a realizar son de escalar diferentes e irregulares.

#### 4.4.4 R2U-Net

*R2U-Net* surge como una variante del modelo base de *UNet* usando redes recurrentes convolucionales con residual. Ideada en 2018 [28], esta recoge las ventajas de 3 modelos diferentes, *UNet*, RNCC [29] y modelos profundos residuales [30] para mejorar las tasas de entrenamiento y contexto entre capas.

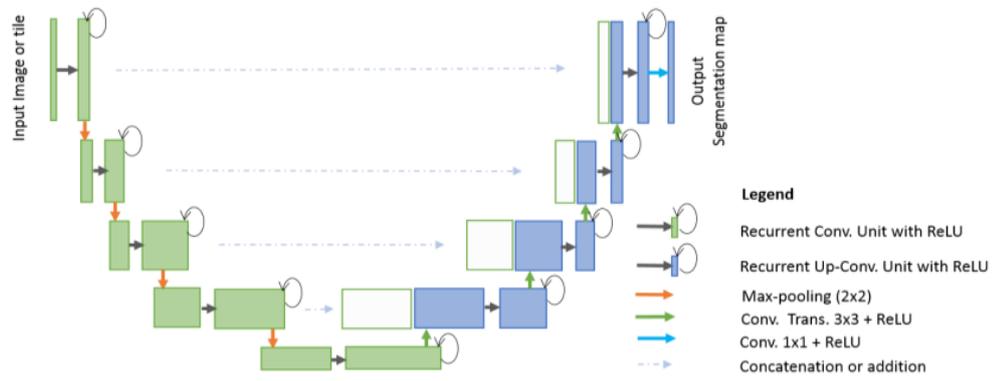


ILUSTRACIÓN 20 – ARQUITECTURA R2U-NET. FUENTE: [28]

Este modelo aporta mejoras de rendimiento en diferentes pruebas de rendimiento en tareas de reconocimiento, clasificación y segmentación de datos.



# Marco experimental

El quinto capítulo se trata sobre los datos empleados por los modelos. Se describen y posteriormente se procede a discernir los canales de las IRM usados para entrenar los distintos modelos. Luego se detalla las técnicas de aumento de datos que se han aplicado a los datos y el porqué de ellas. De seguido se explican los hiperparámetros usados para los modelos. Más adelante se detallan las métricas que sirven en el capítulo siguiente para evaluar nuestro modelo. Por último, se comentan las tecnologías utilizadas para la implementación de los modelos y el equipo informático utilizado.

## 5.1 Descripción de los datos

En la realización de este trabajo se han empleado un total de dos conjuntos de datos médicos. El primero corresponde al “Ivy Glioblastoma Atlas Project (IvyGAP)”, [31]. Este banco de datos consiste en imágenes IRM de tumores cerebrales para un total de 39 pacientes, incluyendo precirugía y postcirugía. El segundo corresponde al “The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM)” [32], conteniendo imágenes de un total de 262 personas.

Sin embargo, el gran tamaño de estos dos *datasets*, (130.4 GB y 73.5 GB) imposibilita la implementación de los modelos por problemas de carencia de un hardware competente. Por ello, se ha optado por utilizar subconjuntos de estos dos bancos de datos mencionados anteriormente, en específico dos subconjuntos usados por ONCOhabitats en dos distintos estudios: *ONCOhabitats results for Ivy Glioblastoma Atlas Project (Ivy Gap): Segmentation and Hemodynamic Tissue Signature* y *ONCOhabitats results for The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM): Segmentation and Hemodynamic Tissue Signature* [33] [34]. Por comodidad y sencillez, renombraré estos dos conjuntos de datos a ONCO\_Ivy y ONCO\_TCGA.

En estos bancos de datos, cada paciente contiene los siguientes tipos de IRM: T1, T2, *FLAIR* y T1-gd. Además, los tejidos tumorales o aquellos que se encuentran entre el tejido se dividen en 3 grupos: *edema* (acumulación de líquido en espacios extracelulares), *enhancing tumor* (núcleo del tumor) y *necrosis* (tejido muerto). Estas se pueden definir en cada una de las IRM definidas anteriormente.

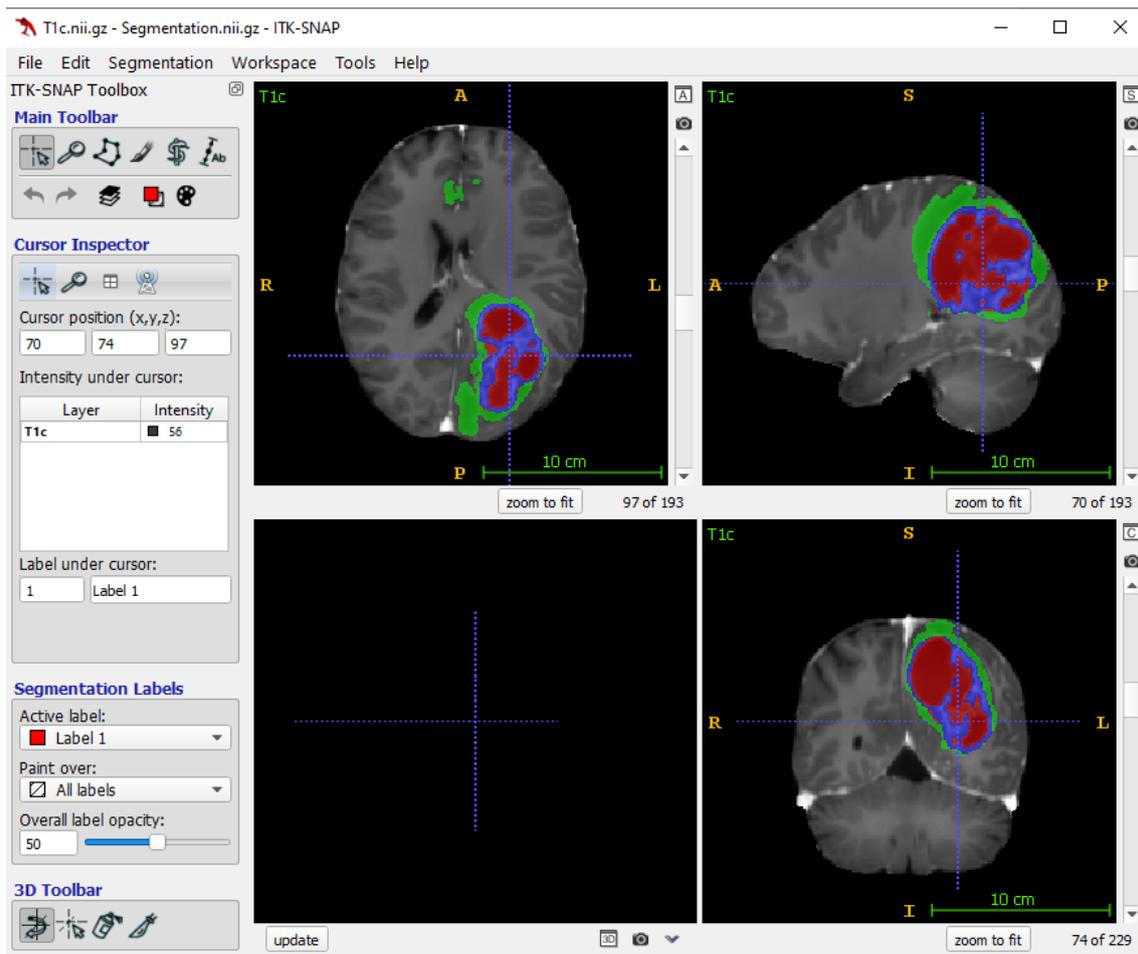


ILUSTRACIÓN 21 – IMAGEN IRM T1-gd CON MÁSCARA DE SEGMENTACIÓN CORRESPONDIENTE AL DATASET ONCO\_IVY. EN VERDE, TEJIDO CORRESPONDIENTE A EDEMA; EN AZUL A ENHANCING TUMOR Y EN ROJO A NECROSIS. FUENTE: ELABORACIÓN PROPIA.

El tejido correspondiente a *edema* puede ser observado por las imágenes de tipo T2 y *FLAIR*, aunque la segunda es más usada para observar la extensión o para distinguirla de otro tipo de estructuras líquidas.

T1 y T1-gd por otro lado son usadas para observar los tejidos correspondientes a *enhancing tumor* y *necrosis*. *Enhancing tumor* se puede captar variando el umbral de T1-gd mientras que el tejido correspondiente a *necrosis* es visible mediante el mismo canal T1-gd.

Por ende, los tres principales canales a elegir serán T2, *FLAIR* y T1-gd. T1 se excluye debido a que ya tenemos la información de T1-gd.

Por último y para finalizar este apartado, estas imágenes han sido sometidas a un preprocesado. En primer lugar, se ha corregido las inhomogeneidades de campo de polarización magnética, ruido o artefactos de pico. Por otro lado, se ha realizado la normalización de la intensidad y

extracción cerebral para generar una resonancia magnética multiparamétrica consistente de alta calidad del cerebro.

## 5.2 Preprocesado, técnicas de aumento y regularización de datos

Antes de acoplar las imágenes y alimentar a los modelos, se ha necesitado realizar un preprocesado y aplicar técnicas de aumento de datos para paliar, entre algunos problemas, el desbalanceo de clases (es decir, hay más muestras de algunas clases que de otras). Para nuestro caso se ha realizado para los datos ya existentes. Sin embargo, hay que ser cuidadosos al aplicar estas técnicas de aumento de datos. Un reciente estudio determinó que las técnicas aplicadas dependen de la clase y, por lo tanto, no todas deberían ser aplicadas [35]. Como resultado a este estudio, se ha procedido a realizar las siguientes técnicas a los datos.

De primeras en el apartado anterior se menciona que T1 se excluye, partiendo que la base que ya se cuenta con la información que aporta T1-gd. Con esto, y aquellos cambios realizados previamente, se ha realizado las dos siguientes técnicas:

### 5.2.1 Normalización

La normalización de los datos es una de las técnicas más comunes en el preprocesado de datos [43]. En muchos casos se piensa que es innecesario y tedioso. No obstante, muchas veces la agrupación de datos a usar sin normalizar presenta muestras cuyos rangos numéricos pueden variar en cuanto a otras. Esto afecta considerablemente a la velocidad de convergencia del modelo, pues el optimizador corrige el factor de aprendizaje para una característica correctamente pero muy equivocadamente para el resto de estas.

Para este trabajo se ha optado por la normalización *Z-Score*. Esta normaliza cada valor del conjunto de datos tal que la media de todos los valores se iguala a 0 y la desviación estándar a 1. Esta sigue la fórmula que se expresa a continuación:

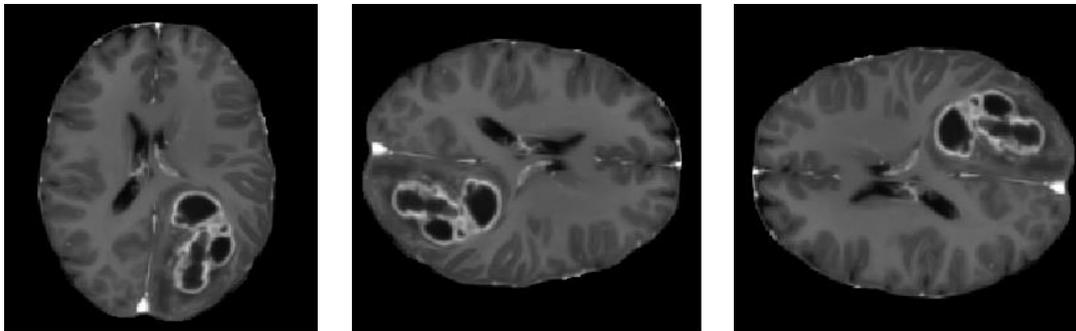
$$y = \frac{(x - \mu)}{\sigma}$$

- $y$  corresponde al nuevo valor.
- $x$  corresponde al valor original.
- $\mu$  simboliza la media de los datos.
- $\sigma$  representa la desviación estándar de los datos.

Cabe destacar que esta normalización se ha realizado excluyendo a los *vóxeles* (píxeles volumétricos) que no incluyen al cerebro.

### 5.2.2 RandomRotation

Esta técnica es una de las más simples de aplicar a los datos. En su práctica definición consiste en rotar la imagen cierta cantidad de grados aleatoriamente. Para nuestro caso de estudio, se permitirá rotaciones en un intervalo  $[-90,90]$  con transiciones de 15 grados, es decir:  $0^\circ$ ,  $15^\circ$ ,  $30^\circ$ , etc.



*ILUSTRACIÓN 22 EJEMPLO DE RANDOMROTATION, PARA  $-90^\circ$  Y  $90^\circ$  RESPECTIVAMENTE. FUENTE: ELABORACIÓN PROPIA.*

## 5.3 Hiperparámetros

Durante la fase de entrenamiento, es importante especificar las configuraciones del algoritmo. Sin embargo y aunque parezca raro, no existe una configuración óptima establecida a la que se deba ajustar los parámetros. La decisión de qué hiperparámetros usar y que valores darle es cuestión única del propio programador. Tras esta aclaración, se procede a comentar los hiperparámetros empleados.

### 5.3.1 Función de pérdida

La función de pérdida es aquella encargada de comparar y evaluar la predicción obtenida por el modelo y el valor real (etiquetas) de aquellas muestras empleadas durante la fase de

entrenamiento. Al generar la función el gradiente correspondiente, es esencial escoger la más adecuada para que el modelo presente una correcta convergencia.

Para el caso de estudio, se ha optado por la función conocida como *dice loss*, concreto una variación de esta propuesta en 2016 [37]. Esta función es ampliamente utilizada en modelos de segmentación donde el conjunto de datos presenta un serio desbalanceo de clases. La fórmula es la siguiente:

$$DL = \frac{1}{N} \sum_i \left[ 1 - \frac{2p_{i1}y_{i1} + Y}{p_{i1}^2 + y_{i1} + Y} \right]$$

- $p_i$  denota la probabilidad de pertenencia a una clase.
- $y_i$  denota la etiqueta a la que pertenece la clase.
- $Y$  implica un factor de suavizado, normalmente de 1.

### 5.3.2 Optimizador del planificador de aprendizaje

A la hora de que un modelo realmente aprenda, es necesario ir ajustando los diferentes parámetros que lo compone. Sin embargo, realizar esta tarea manualmente puede suponer un suplicio, pues la cantidad de parámetros que puede tener cada modelo puede ser del orden de miles de millones.

Para ello se utiliza un optimizador. Este se encarga de modificar los parámetros o pesos de la red en función del factor de aprendizaje o *learning rate*. El *learning rate* indica cuan de grande es el cambio que realizan los parámetros.

En este caso concreto se ha procedido a usar el optimizador *AdamW* [38] con un factor de aprendizaje de  $10^{-3}$ . Este es un optimizador con unos resultados más estables y robustos en comparación con su antecesor, el optimizador *Adam* [39].

## 5.4 Métricas de evaluación

A la hora de evaluar modelos de segmentación de imágenes, existen distintas métricas para definir el rendimiento (*pixel accuracy*, *Jaccard index*, etc.). No obstante, tras varios estudios y proposiciones se ha optado por implementar una única métrica cuantitativa, *Dice Coefficient* o *F1-Score*. Esta métrica calcula la similitud de dos conjuntos A y B, siguiendo la siguiente fórmula:

$$Dice = \frac{2 * |A \cap B|}{|A| + |B|}$$

Esta métrica es sencilla y precisa para dar un resultado concreto del resultado de un modelo en la tarea de segmentación. Para este estudio, *Dice Coefficient* se ha calculado sobre el total de las tres clases del conjunto de datos (*edema*, *enhancing tumor* y *necrosis*) y posteriormente se ha calculado una media aritmética con los tres resultados.

## 5.5 Tecnologías utilizadas

### 5.5.1 Equipo, sistema operativo, entorno de desarrollo y lenguaje de programación

Para la realización de este trabajo se ha utilizado un ordenador HP All-in-One 24-df00xxx con las siguientes especificaciones:

- Procesador Intel(R) Core (TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz
- 8 GB de RAM
- Sistema operativo de 64 bits, procesador basado en x64.

Asimismo, el sistema operativo empleado ha sido *Windows 10 Pro, versión 21H2*. El entorno de desarrollo utilizado ha sido *Visual Studio Code*. Por último, el lenguaje principal de programación ha sido Python, concretamente la versión 3.9, por comodidad, accesibilidad y conocimientos de estas. Además de poseer las librerías necesarias para la implementación de los modelos.

### 5.5.2 Librerías utilizadas

#### **Numpy**

Numpy es la librería por excelencia de Python para la manipulación de estructuras numéricas. El tratamiento de matrices y vectores se realiza eficientemente gracias a la función *ndarray* que implementa esta librería. Muchas otras librerías utilizadas tienen dependencia de esta.

## **Scikit-learn**

Scikit-learn [\[40\]](#) es otra librería de Python que hace uso de Numpy. Esta librería ofrece un amplio abanico de distintos algoritmos relacionados con la clasificación, regresión, etc. Con ella se pueden implementar nuevos algoritmos y como usarlos con nuestros datos.

## **Pytorch**

Pytorch [\[41\]](#) es una librería basada en Torch dedicada al aprendizaje automático. Se caracteriza por la implementación de algoritmos basados en la visión artificial o el procesamiento de lenguajes naturales. También ofrece posibilidades de escalado y optimización de modelos.

## **TorchIO**

TorchIO [\[42\]](#) es una librería para la carga eficiente de datos, preprocesado y aumentación de estos y para la división de imágenes médicas. Esta librería ha sido útil en los primeros pasos del trabajo para, como se ha comentado al principio, la carga eficiente y la aplicación de distintas técnicas de preprocesado a nuestros datos.



## CAPÍTULO 6

# Resultados

Este último apartado sirve para visualizar el resultado y mostrar el desempeño de los distintos modelos para la segmentación de los dos conjuntos de imágenes seleccionados. La estructura de este penúltimo capítulo se divide en dos secciones: uno mostrando los resultados cuantitativos en base a la métrica implementada. Y en segundo lugar se opta por un acercamiento más visual de los resultados. Es decir, se procede pues a comparar la segmentación predicha por el modelo con la original.

### 6.1 Resultados cuantitativos

#### 6.1.1 Dataset ONCO\_Ivy

La siguiente tabla presenta los resultados de rendimiento de los modelos para el *dataset* ONCO\_Ivy. *Dice Coefficient* se ha calculado como un porcentaje.

	Dice Coefficient			
	Edema (%)	Enhancing core (%)	Necrosis (%)	Media (%)
<i>3D-Net</i>	80.09	79.56	81.68	80.44
<i>X-Net</i>	85.59	88.97	87.12	87.22
<i>SegNet</i>	90.59	89.67	89.78	90.01
<i>Multires-UNet</i>	90.09	89.21	88.56	89.28
<b><i>R2U-Net</i></b>	<b>91.92</b>	<b>93.23</b>	<b>90.07</b>	<b>91.74</b>
<i>Media por tejido</i>	87.65	88.12	87.56	

TABLA 1 – TABLA COMPARATIVA DE RESULTADOS PARA DATASET ONCO\_IVY. FUENTE: ELABORACIÓN PROPIA.

A la vista de los resultados, el peor rendimiento lo obtiene el modelo 3D-UNet, con una media *Dice Coefficient* del 80.44%. Por otro lado, el resto de las modelos presentan resultados positivos y similares entre ellos. Sin embargo y para este conjunto de datos, el rendimiento en la tarea de segmentación ha sido R2U-Net, con una tasa del 91.74% de media entre todos los tejidos.

Otro aspecto que nos ofrece la tabla es la media del coeficiente entre los tres tejidos respecto a los 5 modelos. Como se observa, la media oscila entre el 87% y 88%. Esto nos da a entender que el problema del desbalanceo de clases ha sido paliado mediante las técnicas de preprocesado utilizadas.

### 6.1.2 Dataset ONCO\_TCGA

La siguiente tabla los resultados de rendimiento de los modelos para el *dataset* ONCO\_TGCA. Al igual que en el anterior apartado, *Dice Coefficient* se presenta como porcentaje.

	<b>Dice Coefficient</b>			
	<b>Edema (%)</b>	<b>Enhancing core (%)</b>	<b>Necrosis (%)</b>	<b>Media (%)</b>
<i>3D-Net</i>	81.56	78.99	82.56	81.03
<i>X-Net</i>	86.59	87.12	88.26	87.32
<i>SegNet</i>	90.56	88.26	89.62	89.48
<i>Multires-UNet</i>	90.23	90.26	90.34	90.27
<b><i>R2U-Net</i></b>	<b>92.17</b>	<b>93.55</b>	<b>91.98</b>	<b>92.5</b>
<i>Media por tejido</i>	88.22	87.63	88.55	

TABLA 2 – TABLA COMPARATIVA DE RESULTADOS PARA DATASET ONCO\_TCGA. FUENTE: ELABORACIÓN PROPIA.

Analizando la tabla observamos una vez más que el mejor resultado es obtenido por el modelo R2U-Net (92.5%) mientras que el peor desempeño es para el modelo 3D-UNet (81.3%). Asimismo, el resto de los modelos presentan un resultado favorable y cercano al 90%. Por último, al igual que en la última tabla, se observa que la media de *Dice Coefficient* para cada tejido oscila entre el 88% y 87%, demostrando una vez más que las técnicas de preprocesado implementadas han ayudado con el problema del desbalanceo de clases.

## 6.2 Resultados cualitativos

El objetivo principal es crear un modelo que segmente las imágenes IRM de gliomas en sus respectivos tejidos. En contraposición a los resultados cuantitativos, también es interesante mostrar una aproximación visual del desempeño del modelo. Al fin y al cabo, lo que le interesa al experto no es una tabla de resultados, sino una imagen con la segmentación. Por ello, en este apartado se procede a comparar una máscara de segmentación predicha por el mejor modelo de cada conjunto de datos (en este caso el mejor modelo en ambos ha sido el R2U-Net) y

compararlo con la segmentación original. Los colores indican el tipo de tejido tumoral: en verde para *edema*, en rojo para *necrosis* y en azul para *enhancing tumor*.

### 6.2.1 Dataset ONCO\_Ivy

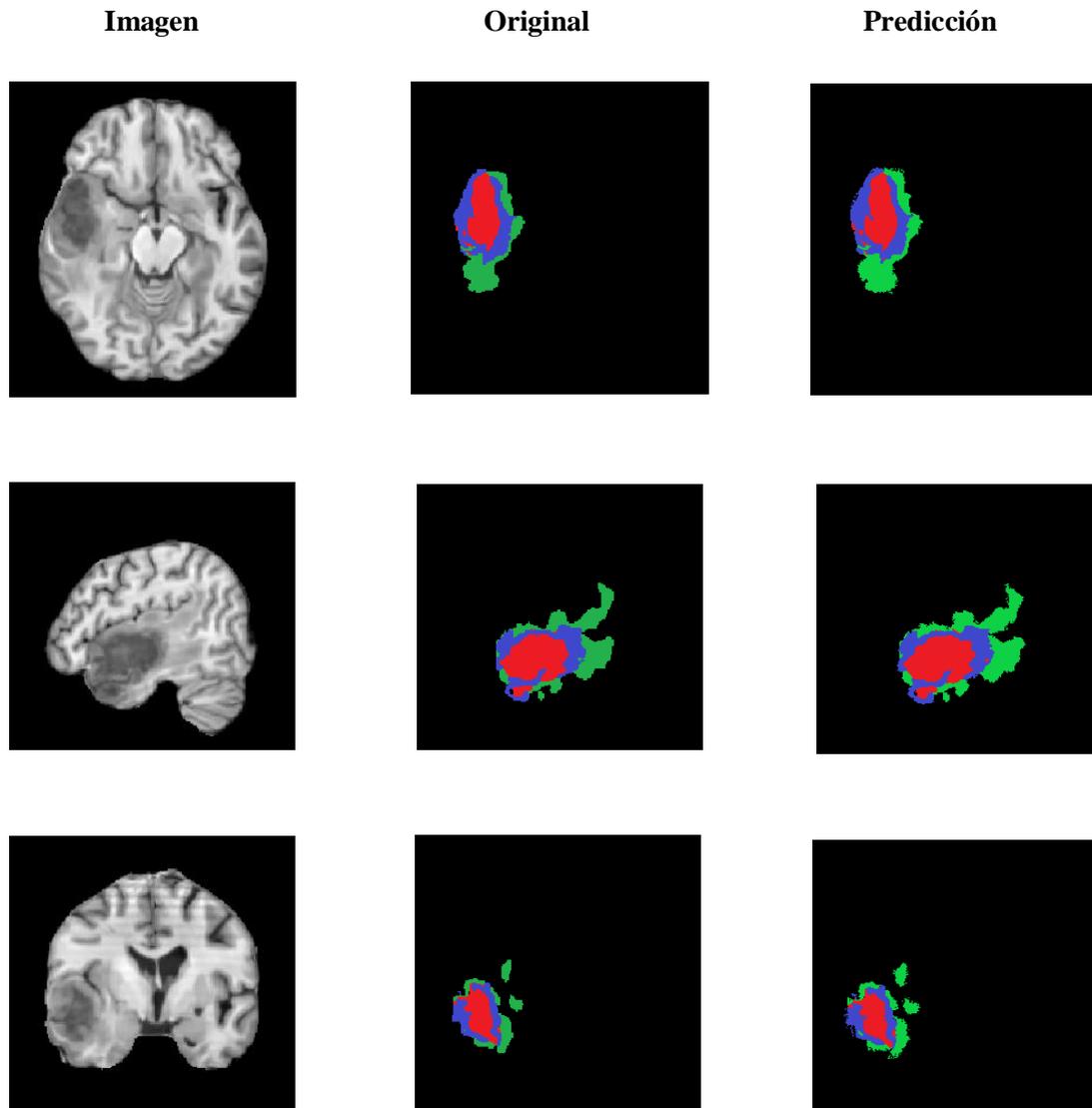


ILUSTRACIÓN 23 – RESULTADOS COMPARATIVOS ENTRE MÁSCARA ORIGINAL Y PREDICHA POR MODELO R2U-NET PARA ONCO\_IVY. FUENTE: ELABORACIÓN PROPIA

### 6.2.2 Dataset ONCO\_TCGA

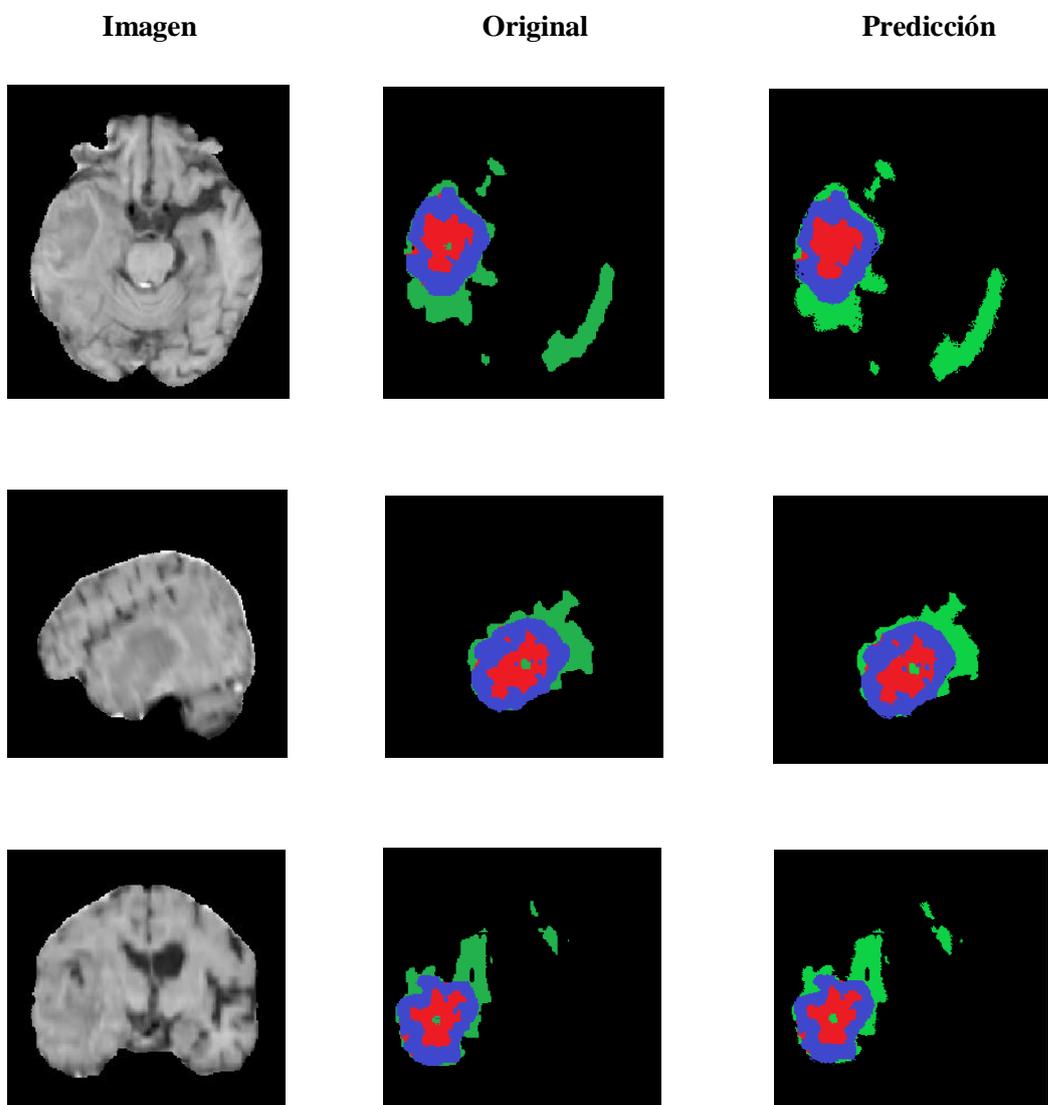


ILUSTRACIÓN 24 – RESULTADOS COMPARATIVOS ENTRE MÁSCARA ORIGINAL Y PREDICHA POR EL MODELO R2U-NET PARA ONCO\_TCGA. FUENTE: ELABORACIÓN PROPIA.



# Conclusiones

La segmentación de tumores cerebrales manualmente sobre imágenes IRM se considera una tarea costosa y que requiera la precisión y conocimientos de personal experto. Sin embargo, se ha demostrado y actualmente se siguen desarrollando métodos tecnológicos de vanguardia en el campo del aprendizaje automático que permiten aminorar la carga laboral de los médicos y que incluso sirven de apoyo para la monitorización y seguimiento en el tratamiento de este tipo de tumor cerebral.

En este trabajo ha partido de un conjunto de imágenes IRM de pacientes de edad avanzada a las que se han eliminado los nombres y sexo para anonimizarlos. Posteriormente se ha decidido tras un estudio premeditado los canales indicados de las imágenes IRM para discernir los distintos tejidos del tumor. Seguidamente se han realizado varios preprocesados sobre las imágenes como técnicas de aumentación de datos. Con esto queda demostrado que se ha cumplido el primer de los subobjetivos marcados al inicio del trabajo.

Tras ello, se han contemplado modelos basados en la arquitectura *UNet* (considerada la más relevante y de renombre en cuanto estado del arte) para la segmentación de imágenes médicas, llegando a escoger un total de 5: *3D-UNet*, *R2U-Net*, *SegNet*, *X-Net* y *Multires-UNET*. Después, se ha procedido a entrenarlos con el conjunto de datos previamente procesados. Con ello se entiende como cumplido el segundo subobjetivo.

Por último, se ha comprobado la eficacia y rendimiento mediante la implementación de distintas métricas adecuadas al problema de segmentación. A posteriori, con los resultados obtenidos se ha escogido el mejor modelo para comparar las máscaras de segmentación reales y las obtenidas por las predicciones. Esto aporta una comparación más visual que métricas basadas en porcentajes y realista dado que podría ser perfectamente un caso de uso en el ámbito profesional. Con esto, se cumplimenta el tercer y último subobjetivo y, por ende, el objetivo final que se había marcado al principio para este trabajo.

Como colofón a este apartado conclusivo, la realización de este trabajo de fin de grado ha supuesto un gran reto a nivel material, intelectual, pero sobre todo a nivel personal.

Materialmente, al disponer de únicamente de un equipo de sobremesa de gama básica, el tiempo de procesado y de entrenamiento de los modelos en un problema a resolver. En segundo lugar, intelectualmente ha sido un reto en calidad de aplicar los conocimientos aprendidos a lo largo de

4 años a un trabajo de esta índole y entender nuevas formas de datos, modelos y técnicas que en la carrera no llegan a verse o profundizarse.

## 7.1 Futuro trabajo

Tras la realización de este trabajo, se plantean diferentes propuestas de ampliación:

- Optimización de diversos hiperparámetros a partir de modelos “preentrenados” y hacer comparativas. Por ejemplo, mediante el uso de la librería de Python Optuna.
- Implementar mejoras de hardware y entrenar los modelos mediante los *datasets* principales. Posibilidad de métricas de rendimiento temporal.
- Ampliar el espectro de segmentación a “semisupervisada” y no supervisada.

## 7.2 Relación con los estudios cursados

A lo largo de la realización de este trabajo ha sido necesario poner en práctica múltiples conocimientos aprendidos al largo de la carrera. De la misma manera ha servido para pulir y masterizar la programación a niveles de eficiencia, precisión y rapidez. Aun así, cabe destacar ciertas asignaturas que han sido clave para este proceso.

En primer lugar, hay que destacar las asignaturas de Introducción a la Informática y Programación (IIP) y Programación (PRG), dado que aportan los conocimientos básicos e intermedios de programación para crear algoritmos con la mayor eficiencia posible.

En segundo lugar, la asignatura de Estadística (EST) ha sido fundamental, pues los conocimientos y las bases estadísticas enseñadas se encuentran en casi todos los ámbitos del aprendizaje automático.

En tercer lugar, Estructuras de Datos y algoritmos (EDA) y Lenguajes, Tecnologías y Paradigmas de la programación (LTP) ha sido fundamental para la creación de un código eficiente y preciso.

En referente a las asignaturas de la rama de Computación, Sistemas Inteligentes (SIN) y Aprendizaje Automático (APR) han aportado el conocimiento y las bases sobre el mundo del aprendizaje automático y reconocimiento de formas e imágenes.

Por último, dentro de las asignaturas de 4º año opcionales se destaca Bioinformática (BIO), que ha sido útil al aportar los conocimientos básicos sobre la bioinformática y el potencial que tiene el aprendizaje automático aplicado al sector de la salud.



# Bibliografía

- [1] GARCÍA, Diana Nancy Martínez, et al. Avances de la inteligencia artificial en salud. *Dominio de las Ciencias*, 2019, vol. 5, no 3, p. 603-613.
- [2] Stein, Rob. (May 21, 2008). Kennedy's Cancer is Highly Lethal. *The Washington Post*.
- [3] WORLD HEALTH ORGANIZATION, et al. Informe Mundial sobre el Cáncer, 2014. *Lyon, France: IARC*, 2014.
- [4] ARVELO, Francisco; POUPON, Marie-France. Aspectos moleculares y celulares de la metástasis cancerosa. *Acta Cient Venez*, 2001, vol. 52, no 4, p. 304-12.
- [5] LOUIS, David N., et al. The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica*, 2007, vol. 114, no 2, p. 97-109.
- [6] BARKER, Fred G., et al. Age and the risk of anaplasia in magnetic resonance-nonenhancing supratentorial cerebral tumors. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 1997, vol. 80, no 5, p. 936-941.
- [7] CONTRERAS, Luis Enrique. Epidemiología de tumores cerebrales. *Revista Médica Clínica Las Condes*, 2017, vol. 28, no 3, p. 332-338.
- [8] ORELLANA, Andrés. Formación de Imágenes por Resonancia Magnética. *Facultad de Ingeniería, Universidad Politécnica Salesiana, Cuenca-Ecuador*, 2016.
- [9] DEONI, Sean CL; PETERS, Terry M.; RUTT, Brian K. High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 2005, vol. 53, no 1, p. 237-241.
- [10] GANZETTI, Marco; WENDEROTH, Nicole; MANTINI, Dante. Mapping pathological changes in brain structure by combining T1-and T2-weighted MR imaging data. *Neuroradiology*, 2015, vol. 57, no 9, p. 917-928.
- [11] ZEINELDIN, Ramy A., et al. DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. *International journal of computer assisted radiology and surgery*, 2020, vol. 15, no 6, p. 909-920.
- [12] YUSHKEVICH, Paul A.; GAO, Yang; GERIG, Guido. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. En *2016 38th annual*

*international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2016. p. 3342-3345.

[13] NACELLE, Andres; MIZRAJI, E. Redes neuronales artificiales. *Núcleo de ingeniería biomédica–Universidad de la Republica Uruguay*, 2009.

[14] TABARES, Héctor; BRANCH, John; VALENCIA, Jaime. Generación dinámica de la topología de una red neuronal artificial del tipo perceptron multicapa. *Revista Facultad de Ingeniería Universidad de Antioquia*, 2006, no 38, p. 146-162.

[15] PONCE, Hector Uriel Rodriguez. Perceptrón multicapa para reconocimiento de objetos sobre planos. 2004.

[16] RUMELHART, David E.; HINTON, Geoffrey E.; WILLIAMS, Ronald J. *Learning internal representations by error propagation*. California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[17] HERNÁNDEZ, Jairo; RODRÍGUEZ, José. Algoritmos de Retropropagación con restricciones para la estimación de parámetros de curvas de titulación. *Ciencia e Ingeniería*, 2018, vol. 39, no 1, p. 13-26.

[18] MUÑOZ, Sandra Milena Medrano. Fundamentos de campo visual. *Ciencia y tecnología para la salud visual y ocular*, 2007, no 8, p. 85-92.

[19] FUKUSHIMA, Kunihiko; MIYAKE, Sei. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. En *Competition and cooperation in neural nets*. Springer, Berlin, Heidelberg, 1982. p. 267-285.

[20] LECUN, Yann, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, vol. 86, no 11, p. 2278-2324.

[21] PALOMARES, Fernando Giménez; MONSORIU, Juan A.; ALEMANY, Elena. Aplicación de la convolución de matrices al filtrado de imágenes. *Modelling in Science Education and Learning*, 2016, vol. 9, no 1, p. 97-108.

[22] RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. En *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015. p. 234-241.

[23] LI, Zhi, et al. Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. *International Journal of Imaging Systems and Technology*, 2019, vol. 29, no 4, p. 577-583.

- [24] SIDDIQUE, Nahian, et al. U-net and its variants for medical image segmentation: A review of theory and applications. *Ieee Access*, 2021, vol. 9, p. 82031-82057.
- [25] BULLOCK, Joseph; CUESTA-LÁZARO, Carolina; QUERA-BOFARULL, Arnau. XNet: A convolutional neural network (CNN) implementation for medical X-ray image segmentation suitable for small datasets. En *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*. SPIE, 2019. p. 453-463.
- [26] BADRINARAYANAN, Vijay; KENDALL, Alex; CIPOLLA, Roberto. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2017, vol. 39, no 12, p. 2481-2495.
- [27] IBTEHAZ, Nabil; RAHMAN, M. Sohel. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural networks*, 2020, vol. 121, p. 74-87.
- [28] ALOM, Md Zahangir, et al. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.
- [29] Liang, Ming, and Xiaolin Hu. "Recurrent convolutional neural network for object recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
- [30] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- [31] SHAH, N., et al. Data from Ivy GAP [Data set]. The Cancer Imaging Archive. 2016.
- [32] SCARPACE, L., et al. Radiology Data from The Cancer Genome Atlas Glioblastoma Multiforme [TCGA-GBM] collection [Data set]. The Cancer Imaging Archive. 2016.
- [33] JUAN-ALBARRACÍN, Javier, et al. Glioblastoma: vascular habitats detected at preoperative dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging predict survival. *Radiology*, 2018, vol. 287, no 3, p. 944-954.
- [34] DEL MAR ÁLVAREZ-TORRES, María, et al. Robust association between vascular habitats and patient prognosis in glioblastoma: An international multicenter study. *Journal of Magnetic Resonance Imaging*, 2020, vol. 51, no 5, p. 1478-1486.
- [35] BALESTRIERO, Randall; BOTTOU, Leon; LECUN, Yann. The effects of regularization and data augmentation are class dependent. *arXiv preprint arXiv:2204.03632*, 2022.
- [36] PATRO, S.; SAHU, Kishore Kumar. Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*, 2015.

- [37] KODYM, Oldřich; ŠPANĚL, Michal; HEROUT, Adam. Segmentation of head and neck organs at risk using cnn with batch dice loss. En *German conference on pattern recognition*. Springer, Cham, 2018. p. 105-114.
- [38] LOSHCHILOV, Ilya; HUTTER, Frank. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.
- [39] KINGMA, Diederik P.; BA, Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [40] PEDREGOSA, Fabian, et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 2011, vol. 12, p. 2825-2830.
- [41] PASZKE, Adam, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 2019, vol. 32.
- [42] PÉREZ-GARCÍA, Fernando; SPARKS, Rachel; OURSELIN, Sébastien. TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine*, 2021, vol. 208, p. 106236.
- [43] SINGH, Dalwinder; SINGH, Birmohan. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 2020, vol. 97, p. 105524.
- [44] GORDILLO, Nelly; MONTSENY, Eduard; SOBREVILLA, Pilar. State of the art survey on MRI brain tumor segmentation. *Magnetic resonance imaging*, 2013, vol. 31, no 8, p. 1426-1438.
- [45] OTSU, Nobuyuki. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 1979, vol. 9, no 1, p. 62-66.
- [46] KASS, Michael; WITKIN, Andrew; TERZOPOULOS, Demetri. Snakes: Active contour models. *International journal of computer vision*, 1988, vol. 1, no 4, p. 321-331.
- [47] SALMAN, Yasser M. Modified technique for volumetric brain tumor measurements. *Journal of Biomedical Science and Engineering*, 2009, vol. 2, no 1, p. 16.
- [48] KHALID, Noor Elaiza Abdul; IBRAHIM, Shafaf; HANIFF, P. N. M. M. MRI Brain Abnormalities Segmentation using K-Nearest Neighbors(k-NN). *International Journal on Computer Science and Engineering*, 2011, vol. 3, no 2, p. 980-990.
- [49] KHARRAT, Ahmed, et al. A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine. *Leonardo journal of sciences*, 2010, vol. 17, no 1, p. 71-82.

[50] KAVITHA, A. R.; CHELLAMUTHU, C.; RUPA, Kavin. An efficient approach for brain tumour detection based on modified region growing and neural network in MRI images. En *2012 international conference on Computing, Electronics and Electrical Technologies (ICCEET)*. IEEE, 2012. p. 1087-1095.





## ANEXO

### OBJETIVOS DE DESARROLLO SOSTENIBLE

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No Procede
ODS 1. <b>Fin de la pobreza.</b>				X
ODS 2. <b>Hambre cero.</b>				X
ODS 3. <b>Salud y bienestar.</b>	X			
ODS 4. <b>Educación de calidad.</b>				X
ODS 5. <b>Igualdad de género.</b>				X
ODS 6. <b>Agua limpia y saneamiento.</b>				X
ODS 7. <b>Energía asequible y no contaminante.</b>				X
ODS 8. <b>Trabajo decente y crecimiento económico.</b>		X		
ODS 9. <b>Industria, innovación e infraestructuras.</b>		X		
ODS 10. <b>Reducción de las desigualdades.</b>				X
ODS 11. <b>Ciudades y comunidades sostenibles.</b>				X
ODS 12. <b>Producción y consumo responsables.</b>				X
ODS 13. <b>Acción por el clima.</b>				X
ODS 14. <b>Vida submarina.</b>				X
ODS 15. <b>Vida de ecosistemas terrestres.</b>				X
ODS 16. <b>Paz, justicia e instituciones sólidas.</b>				X
ODS 17. <b>Alianzas para lograr objetivos.</b>				X



Reflexión sobre la relación del TFG/TFM con los ODS y con el/los ODS más relacionados.

El trabajo de final de grado presentado se titula “Diseño de una herramienta de clasificación semántica de imágenes de gliomas de alto grado mediante redes convolucionales profundas”. Este trabajo conforma un estudio de distintos modelos neuronales convolucionales para la segmentación de imágenes médicas. Por lo tanto, guarda cierta relación con distintos puntos de los Objetivos de Desarrollo Sostenible (ODS).

Como preludio a los posteriores párrafos, los puntos con los que guarda relación este trabajo con los ODS son tres en concreto: “ODS 3 Salud y bienestar”, “ODS 8 Trabajo decente y crecimiento económico” y “ODS 9 Industria, innovación e infraestructuras”. A continuación, se desarrolla la justificación de la relación de cada ODS.

La conexión entre este proyecto y el ODS 3 es más que evidente, y uno de los de mayor grado de relación. En el tratamiento de pacientes con gliomas o glioblastomas, obtener información clara, concisa y precisa a través de imágenes y la posterior segmentación del tumor es crucial. El uso de modelos para la obtención rápidas y precisas de estas segmentaciones ayudan a los expertos a mejorar y proponer mejores tratamientos para cada caso de uso. Esto se traduciría en los mejores casos el aumento de la probabilidad de recuperación de los pacientes y, en el peor, la posibilidad de aumentar la esperanza de vida.

En segundo lugar, el ODS 8 Trabajo decente y crecimiento económico se relaciona con el trabajo puesto que la propuesta de modelos que segmenten las imágenes IRM está también pensado para ayudar a los profesionales del sector de la salud. En general, el proceso de segmentación de distintas IRM a mano es un proceso largo, muy exhaustivo y que requiere la participación de distintos expertos. Mediante el uso de estos modelos como apoyo, o posiblemente en un futuro como herramienta principal, pueden llegar a aliviar esa carga laboral y permitir que ese tiempo sea empleado eficientemente en otros aspectos del tratamiento de los pacientes.

Por otro lado, el ODS 9 Industria, innovación e infraestructura presenta un grado medio de relación con el trabajo presentado. El campo de la inteligencia artificial está en constante cambio y crecimiento. Aunque este proyecto haya presentado como modelos aquellos en el estado del arte en cuanto a segmentación de imágenes médicas, no se descarta que en un futuro se desarrollen otros modelos que dejen obsoletos a los propuestos actualmente. No obstante, la utilidad de este proyecto también reside en la posibilidad de ser usado como punto de apoyo para nuevos proyectos que innoven en nuevos modelos de segmentación.



Como se ha podido observar en los párrafos anteriores, este trabajo presenta cierta relación con diversos puntos de los Objetivos de Desarrollo Sostenibles, en un caso concreto una relación directa. Esto se traduce que aquellos proyectos, tanto pasados como futuro y cercanos también presentarán un grado de relación media con los ODS, destacando su alto grado de relación con “ODS 3 Salud y bienestar”. Por ende, se concluye que el trabajo presenta un grado de relación medio con los ODS.

