

**Cuantificación automática de la carga metabólica tumoral en  
imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA  
en pacientes con carcinoma de próstata**

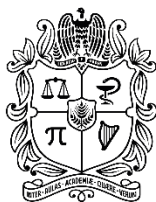


UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

**Camilo Ernesto Sarmiento Torres**

Universidad Nacional de Colombia sede Bogotá  
Facultad de Ciencias  
Departamento de Física  
Maestría en Física Médica  
Bogotá D.C.  
2023

**Cuantificación automática de la carga metabólica tumoral en imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA en pacientes con carcinoma de próstata**



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

Trabajo final presentado como requisito parcial para optar al título de Magister en Física Médica

Presentado por:

**Camilo Ernesto Sarmiento Torres**

MSc. en Ingeniería Física

**Director:**

Francisco Albeiro Gómez Jaramillo

PhD (Dr. Ing) en Ciencias de la Computación

Universidad Nacional de Colombia sede Bogotá

**Codirector:**

Nathaly Barbosa Parada

MSc. en Física Médica

**Línea de Investigación:**

Medicina Nuclear

Universidad Nacional de Colombia sede Bogotá

Facultad de Ciencias

Departamento de Física

Maestría en Física Médica

Bogotá D.C.

2023

## Nota de aval

Aprobado por el comité de director y codirector en cumplimiento a los requisitos exigidos por la Universidad Nacional de Colombia para la presentación del Trabajo Final de la Maestría en Física Médica.

Francisco Gómez J

---

PhD. Francisco Albeiro Gómez Jaramillo  
Director



---

MSc, Nathaly Barbosa Parada  
Codirector

## Agradecimientos

El autor expresa su más sincero y profundo agradecimiento a todas las personas que contribuyeron al desarrollo de este trabajo. En especial, quisiera agradecer a mi familia, amigos y compañeros por su apoyo incondicional a lo largo de este proceso. Un agradecimiento especial al Dr. Francisco Albeiro Gómez, por su invaluable dirección y orientación, que fueron fundamentales para el éxito de este proyecto. A la profesora Nathaly Barbosa, mi tutora, le agradezco enormemente su dedicación y valiosos aportes en mi formación, así como por ayudarme a encontrar el enfoque adecuado para el desarrollo de este trabajo. También quiero expresar mi gratitud a la Dra. Mónica Naranjo por sus enseñanzas y valiosos aportes en el campo de la medicina, los cuales fueron fundamentales para el desarrollo de esta investigación. Agradezco de igual manera a la profesora Lorena Sandoval, quien me brindó una sólida formación en el campo de la Medicina Nuclear, lo que resultó crucial para el éxito de este trabajo. Además, quiero extender mi reconocimiento al Instituto Nacional de Cancerología, que me brindó la oportunidad de formarme y permitió el desarrollo de este trabajo. Por último, mi agradecimiento a la Universidad Nacional de Colombia y todo su talento humano, quienes contribuyeron significativamente a mi formación profesional y enriquecieron este proyecto.

¡Muchas gracias a todos por su inestimable apoyo!

## Resumen

**Título:** Cuantificación automática de la carga metabólica tumoral en imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA en pacientes con carcinoma de próstata

**Contexto:** El cáncer de próstata es el segundo tipo de cáncer más frecuente en hombres a nivel mundial, y su incidencia varía considerablemente según las regiones geográficas. En Latinoamérica, se estima que es el tercer tipo de cáncer más común y el más letal entre los hombres. Para un tratamiento efectivo, es crucial determinar con precisión la extensión de la enfermedad, y las imágenes PET/CT con radiotrazadores específicos del antígeno de membrana de próstata (PSMA) han demostrado ser útiles para este propósito. Estas imágenes también permiten medir la carga metabólica tumoral, un indicador importante para evaluar la respuesta al tratamiento, ya que una mayor carga metabólica se relaciona con un grado de malignidad más elevado y un posible avance de la enfermedad. No obstante, las imágenes PET/CT presentan desafíos debido a su baja resolución y alto ruido, lo que dificulta la delimitación precisa de la carga metabólica tumoral y requiere una considerable cantidad de tiempo. Los métodos comunes para esta delimitación son la interpretación visual y el uso de umbrales en software semiautomáticos, pero ambos presentan limitaciones y variabilidad dependiendo del observador o del umbral seleccionado. En el caso del cáncer de próstata metastásico, la delimitación manual es aún más compleja debido a la presencia de múltiples metástasis, lo que puede afectar la determinación de los tratamientos. Para afrontar estas dificultades, en este trabajo se han explorado enfoques basados en aprendizaje de máquina, que ofrecen una mayor precisión y se han revelado como una herramienta valiosa para mejorar la delimitación y cuantificación del MTB, brindando un método validado, reproducible y preciso para los centros oncológicos.

**Objetivo:** El objeto de este trabajo fue cuantificar de forma automática la carga metabólica tumoral a través de algoritmos de aprendizaje de máquina en imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA en pacientes diagnosticados con carcinoma de próstata.

**Metodología:** Este trabajo adopto la metodología CRISP-DM (Cross Industry Standard Process of Data Mining) pero se realizaron adaptaciones donde fue necesario. Se recolectaron 100 exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA en pacientes diagnosticados con carcinoma de próstata los cuales fueron anotados y estandarizados para después entrenar y evaluar una red neuronal convolucional con la arquitectura Attention U-NET en la tarea de la segmentación semántica de la carga metabólica tumoral.

**Resultados:** Los resultados principales incluyen la creación de un modelo que puede cuantificar de manera automática la carga metabólica tumoral en exámenes PET/CT de

$^{68}\text{Ga}$ -PSMA en pacientes diagnosticados con carcinoma de próstata en la región de la pelvis. Este modelo destaca las captaciones en la próstata y los ganglios linfáticos.

## **Palabras clave**

Cáncer de próstata;  $^{68}\text{Ga}$ -PSMA; Carga metabólica tumoral; Redes neuronales convolucionales; PET/CT.

## Abstract

**Title:** Automatic quantification of metabolic tumor burden on  $^{68}\text{Ga}$ -PSMA positron emission tomography images in patients with prostate cancer.

**Context:** Prostate cancer is the second most frequent type of cancer in men worldwide, and its incidence varies considerably according to geographical regions. In Latin America, it is estimated that it is the third most common type of cancer and the most lethal among men. For effective treatment, it is crucial to accurately determine the extent of the disease, and PET/CT imaging with prostate membrane antigen (PSMA)-specific radiotracers has proven useful for this purpose. These images also make it possible to measure the metabolic tumor burden, an important indicator to assess the response to treatment, since a higher metabolic burden is related to a higher degree of malignancy and a possible progression of the disease. However, PET/CT imaging presents challenges due to its low resolution and high noise, making precise delineation of metabolic tumor burden difficult and requiring a considerable amount of time. The common methods for this delimitation are visual interpretation and the use of thresholds in semi-automatic software, but both present limitations and variability depending on the observer or the selected threshold. In the case of metastatic prostate cancer, manual delimitation is even more complex due to the presence of multiple metastases, which can affect the determination of treatments. To face these difficulties, in this work we have explored approaches based on machine learning, which offer greater precision and have been revealed as a valuable tool to improve the delimitation and quantification of MTB, providing a validated, reproducible and precise method for cancer centers.

**Objective:** The aim of this work was to automatically quantify the tumor metabolic burden through machine learning algorithms in  $^{68}\text{Ga}$ -PSMA positron emission tomography images in patients diagnosed with prostate carcinoma.

**Methodology:** This work adopted the CRISP-DM (Cross Industry Standard Process of Data Mining) methodology, but adaptations were made where necessary. One hundred  $^{68}\text{Ga}$ -PSMA PET/CT scans were collected from patients diagnosed with prostate carcinoma, which were annotated and standardized to later train and evaluate a convolutional neural network with the Attention U-NET architecture in the task of semantic segmentation of metabolic tumor burden.

**Results:** The main results include the creation of a model that can automatically quantify the metabolic tumor burden on  $^{68}\text{Ga}$ -PSMA PET/CT scans in patients diagnosed with prostate carcinoma in the pelvic region. This model highlights uptakes in the prostate and lymph nodes.

## Key-words

Prostate cancer;  $^{68}\text{Ga}$ -PSMA; Metabolic tumor burden; Convolutional neural networks; PET/CT.



# Tabla de contenido

Capítulo 1.....	1
Justificación y planteamiento del problema .....	1
1.1 Instrucción .....	1
1.2 Justificación .....	2
1.3 Antecedentes.....	4
1.4 Planteamiento del problema.....	6
1.5 Pregunta de investigación .....	7
1.6 Objetivos .....	7
1.6.1 General.....	7
1.6.2 Específicos .....	7
1.7 Contribuciones .....	8
Capítulo 2.....	9
Carcinoma de próstata y aprendizaje profundo.....	9
2.1 Carcinoma de próstata.....	9
2.1.1 Estudios a través de imágenes médicas .....	10
2.2 Carga metabólica tumoral.....	10
2.3 Medicina nuclear .....	12
2.4 Radiofármaco $^{68}\text{Ga}$ -PSMA.....	13
2.5 Tomografía por emisión de positrones (PET/CT) .....	15
2.6 Método de reconstrucción iterativo.....	16
2.6.1 Parámetros en la reconstrucción iterativa.....	17
2.7 Aprendizaje de máquina .....	18
2.7.1 Aprendizaje supervisado .....	19
2.8 Redes neuronales convolucionales.....	19
2.8.1 Arquitectura general de las CNN .....	20
2.9 Transferencia de aprendizaje y ajuste fino.....	24
Capítulo 3.....	26
Materiales y métodos. ....	26
3.1 Descripción de la metodología .....	26

3.1.1	Comprensión del problema .....	26
3.1.2	Comprensión de los datos .....	26
3.1.3	Preparación de los datos .....	27
3.1.4	Modelado .....	27
3.1.5	Evaluación .....	27
3.1.6	Despliegue .....	27
3.2	Datos .....	27
3.2.1	Prescripción de la actividad a administrar .....	28
3.2.2	Exploración anatómica axial .....	29
3.2.3	Exploración cuerpo completo .....	29
3.2.4	Protocolo de adquisición y parámetros técnicos .....	30
3.2.5	Procesamiento de las imágenes .....	31
3.2.6	Recolección de los datos .....	31
3.3	Preparación de los datos .....	32
3.3.1	Estandarización .....	32
3.3.2	Anotación de los datos .....	32
3.3.2	Binarización .....	34
3.4	Modelado .....	34
3.4.1	U-NET .....	34
3.4.2	U-NET++ .....	36
3.4.3	Attention U-NET .....	37
3.4.4	Columna vertebral de las arquitecturas .....	40
3.4.5	Parámetros utilizados en las arquitecturas .....	41
3.5	Evaluación .....	42
3.5.1	Validación cruzada anidada en k particiones .....	42
3.5.2	Índices .....	44
3.6	Despliegue .....	49
Capítulo 4.....		50
Resultados y discusión .....		50
4.1	Conjunto de datos .....	50
4.2	Modelado.....	52

4.2.1 Segmentación del examen PET/CT de <b>68Ga</b> -PSMA .....	52
4.2.2 Segmentación del examen PET/CT <b>68Ga</b> -PSMA a nivel de grupo .....	54
4.2.3 Determinación de la arquitectura .....	56
4.2.4 Evaluación de la capacidad de generalización .....	57
4.2.5 Comparación con método de umbralización .....	59
4.3 Discusión de los resultados .....	60
4.4 Despliegue del componente inteligente .....	62
Capítulo 5.....	65
Conclusiones y trabajos futuros .....	65
5.1 Conclusiones .....	65
5.2 Trabajos futuros .....	66
Referencias .....	68
Anexos .....	75
Anexo I. Columnas vertebrales .....	75
Anexo II. Resultados de la selección de la arquitectura.....	76
Anexo III. Evaluación de la capacidad de generalización. ....	81

## Lista de Figuras

Figura 2.1 Ejemplo de pacientes con carga tumoral baja, media y alta en imágenes PET de $^{68}\text{Ga}$ -PSMA [35].	11
Figura 2.2 Biodistribución del $^{68}\text{Ga}$ -PSMA-11 [39].	14
Figura 2.3 Ejemplo de una adquisición con diferentes reconstrucciones iterativas [37].	17
Figura 2.4 Detección de objetos vs segmentación semántica [46].	20
Figura 2.5 Analogía de la operación de convolución [48].	21
Figura 2.6 Arquitectura general de una red neuronal convolucional [50].	22
Figura 2.7 Ejemplos de capas convolucionales y mapeo espacial [41].	22
Figura 2.8 Mapeo espacial de la matriz [41].	23
Figura 12.9 Operaciones de agrupación [41].	23
Figura 3.1 Metodología CRISP-DM [30].	26
Figura 3.2 Equipo biomédico Siemens Biograph™ TruePoint PET/CT 40.	28
Figura 3.3 Paciente decúbito supino brazos arriba entrando cráneo.	29
Figura 3.4 Paciente en posición decúbito supino con los brazos abajo entrando pies primero.	30
Figura 3.5 Reconstrucción del examen.	31
Figura 3.6 Ejemplo del proceso de segmentación de examen PET/CT en 3D Slicer.	33
Figura 3.7 Ejemplo de un corte de interés y la correspondiente máscara segmentada por el médico, la cual está asociada a una captación del radiofármaco.	33
Figura 3.8 Ejemplo de arquitectura U-NET [54].	35
Figura 3.9 Arquitectura de U-NET++ [55].	37
Figura 3.10 Ejemplo de implementación de la atención suave aditiva [57].	39
Figura 3.11 Ejemplo de la implementación de una columna vertebral [58].	40
Figura 3.12 Validación cruzada anidada en k particiones.	44
Figura 3.13 Bucles del proceso de evaluación.	44
Figura 3.14 Coeficiente Sørensen-Dice [62].	48
Figura 3.15 Arquitectura del componente inteligente.	49
Figura 4.1 Distribución de imágenes por región anatómica.	51
Figura 4.2 Ejemplo de lesiones en la región de la pelvis. Las flechas verdes indican las lesiones en el corte axial, a la izquierda en un ganglio y a la derecha en la próstata.	51

## Lista de tablas

Tabla 3.1 Actividad del radiofármaco. ....	29
Tabla 3.2 Duración de la cama en función del peso del paciente. ....	30
Tabla 3.3 Protocolos de adquisición de exámenes PET/CT de 68Ga-PSMA. ....	30
Tabla 3.4 Matriz de confusión. ....	45
Tabla 4.1 Demografía de los pacientes sometidos al examen PET/CT. ....	50
Tabla A.1 Columnas vertebrales compatibles con keras-unet-collection ....	75
Tabla A.2 Coeficientes de Dice para el bucle interno ....	76
Tabla A.3 Función de pérdida de Jaccard para el bucle interno. ....	77
Tabla A.4 Exactitud para el bucle interno. ....	78
Tabla A.5 Precisión para el bucle interno. ....	78
Tabla A.6 Recall para el bucle interno. ....	79
Tabla A.7 Área bajo la curva ROC (AUC) para el bucle interno. ....	80
Tabla A.8 Coeficiente de Dice para la evaluación de la capacidad de generalización.....	81
Tabla A.9 Función de pérdida de Jaccard para la evaluación de la capacidad de generalización. ....	82
Tabla A.10 Exactitud para la evaluación de la capacidad de generalización. ....	83
Tabla A.11 Precisión para la evaluación de la capacidad de generalización.....	83
Tabla A.12 Recall para la evaluación de la capacidad de generalización. ....	84
Tabla A.13 AUC para la evaluación de la capacidad de generalización. ....	85

# Capítulo 1.

## Justificación y planteamiento del problema

### 1.1 Introducción

El presente trabajo explora la cuantificación automática de la carga metabólica tumoral través de la implementación de redes neuronales convolucionales en exámenes de tomografía por emisiones positrones (PET/CT) de  $^{68}\text{Ga}$ -PSMA debido a que convencionalmente esta tarea requiere gran cantidad de tiempo por parte del médico nuclear. Las dos técnicas convencionales utilizadas comúnmente son la cuantificación manual y el empleo de técnicas de umbralización. Sin embargo, ambas presentan limitaciones que pueden mitigarse mediante la implementación de redes neuronales convolucionales.

La cuantificación manual, aunque precisa, es un proceso laborioso y dependiente de la interpretación subjetiva del médico. Por otro lado, las técnicas de umbralización pueden subestimar o sobreestimar la carga metabólica tumoral debido a la variabilidad en la selección de umbrales.

La aplicación de redes neuronales convolucionales emerge como una solución prometedora para superar estas limitaciones. Estas redes pueden aprender patrones complejos y características relevantes de manera automática a partir de los datos de imágenes PET/CT, permitiendo una cuantificación más precisa y eficiente de la carga metabólica tumoral. Su capacidad para identificar y analizar patrones sutiles en las imágenes puede ofrecer resultados más consistentes y objetivos, reduciendo la subjetividad inherente a otros métodos.

Al implementar estas redes neuronales convolucionales, se busca no solo agilizar el proceso de cuantificación, sino también mejorar la precisión y fiabilidad de la evaluación de la carga metabólica tumoral, lo que potencialmente podría tener un impacto significativo en la práctica clínica y el seguimiento de los pacientes con cáncer.

A lo largo de este capítulo se exploran el contexto de investigación, los antecedentes, los objetivos planteados y las contribuciones desarrolladas. Estos aspectos permiten introducir al lector en el tema. Sin embargo, se recomienda explorar el Capítulo 2 previamente para comprender los conceptos necesarios que son fundamentales para abordar con precisión el problema en cuestión, ya que se requieren conocimientos

específicos en este campo. El Capítulo 3 se enfoca en el método propuesto, detallando cómo se desarrolló la solución planteada. A continuación, en el Capítulo 4, se presentan los resultados obtenidos, y finalmente, en el Capítulo 5, se exponen las conclusiones y se plantean las líneas de investigación futuras.

## 1.2 Justificación

El cáncer de próstata es la segunda neoplasia maligna más frecuente en los hombres en el mundo, con un total de 1 276 106 casos nuevos y 358 989 muertes en el año 2018, lo cual corresponde al 3,8 % de las muertes por cáncer en hombres [1]. La tasa de incidencia de esta forma de cáncer varía según las regiones y las poblaciones. En particular, la tasa estandarizada por edad es más alta en Oceanía (79,1 por cada 100 000 hombres) y América del norte (73,7 por cada 100 000 hombres). Por el contrario, regiones como África y Asia presentan tasas de incidencias inferiores con 26,6 y 11,5 por cada 100 000 hombres, respectivamente [1].

En Latinoamérica la información respecto a la incidencia del cáncer de próstata es escasa. En particular, a partir de 2011 tan solo el 21% de los países de la región tenían un registro detallado de este tipo de cáncer. No obstante, se estima que el cáncer de próstata es el tercer tipo de tumor más frecuente (después de pulmón y mama), y el más frecuente entre los hombres, con la tasa de mortalidad más alta entre todos los cánceres. Al respecto, la tasa de incidencia es de 27,3 casos por 100 000 hombres, con una tasa de mortalidad de 11,3 muertes por 100 000 hombres en las últimas 3 décadas [2]. En Colombia, este tipo de cáncer es el de mayor incidencia en hombres. Datos recientes sugieren la existencia de 8,872 casos nuevos anualmente, para una tasa ajustada de 46,5 afectados por 100 000 habitantes (cálculo hipotético para la comparación entre grupos) [3].

La determinación precisa de la extensión del cáncer de próstata es fundamental para el diseño de terapias personalizadas. En particular, la determinación temprana y precisa de la carga metabólica de la enfermedad está asociada a una mayor probabilidad de supervivencia del paciente [4]. Las aproximaciones más comunes para la determinación de la extensión de la enfermedad incluyen cuantificaciones sobre tomografía computarizada (Computed Tomography, CT), la imagen por resonancia magnética (Magnetic Resonance Imaging, MRI) y la gammagrafía ósea. No obstante, la imagen de CT y MRI no consideran el metabolismo de la enfermedad y la gammagrafía ósea tienen una menor sensibilidad para la detección de lesiones metastásicas respecto de la imagen, PET/CT (Positron Emission Tomography and Computed Tomography) [5], [6]. Con el fin de abordar estas limitaciones, recientemente se propuso el uso del antígeno de membrana de próstata (PSMA, por sus siglas en inglés), como un marcador más preciso de la actividad metabólica asociada al cáncer de próstata, utilizado para la toma de imágenes PET. Esta aproximación ha mostrado resultados prometedores en tareas asociadas con la planificación del tratamiento, particularmente, en la localización de tumores no solamente en la próstata sino también en

toda la pelvis, y en otras partes del cuerpo, en los casos en que los tumores han migrado [7].

Es importante anotar que la variabilidad en la detección para los dos radiotrazadores principales basados en PSMA ( $^{68}\text{Ga}$ -PSMA y el  $^{18}\text{F}$ -PSMA) es baja. Comúnmente estos se diferencian en la identificación de lesiones presuntamente inespecíficas o de bajo grado de relevancia clínica en relación con marcaciones malignas. Usualmente, la elección del radiotrazador depende de la disponibilidad local [8], [9].

La evidencia más reciente sugiere que la medición de los volúmenes metabólicos del tumor en imágenes PET tiene un gran impacto en la determinación precisa de la carga metabólica tumoral (MTB) [7]. La cual, a su vez, se correlaciona con la probabilidad de supervivencia global del paciente. El MTB cuantifica la cantidad de cáncer en todo el cuerpo, y permite estimar cantidades como el volumen metabólico del tumor (MTV) y la glucólisis total de la lesión (TLG). El MTV y el TGL corresponden al tamaño de la lesión y la expresión de PSMA intralesional, respectivamente, y permiten predecir la curación metastásica [10]. Por esta razón, la planificación del tratamiento basada en la extensión del MTB ha cobrado importancia recientemente en la terapia del cáncer de próstata porque permite predecir el resultado y evaluar la respuesta al tratamiento en terapias radiometabólicas [11]. En este sentido, varios ensayos clínicos han sugerido la utilización del MTB para la estratificación de los pacientes en enfermedad de bajo o alto volumen [10].

Por ejemplo, en el cáncer de próstata hormonosensible metastásico la delimitación del MTB ha permitido diferenciar a los pacientes con un alto volumen de enfermedad respecto de los pacientes con un bajo volumen de la enfermedad, para determinar los tratamientos o terapias que se realizan [12]. Para un alto volumen de enfermedad en [13] se planteó un ensayo clínico aleatorizado de quimioterapia hormonal versus ablación de andrógenos para enfermedad extensa en cáncer de próstata - Chemohormonal Therapy Versus Androgen Ablation Randomized Trial for Extensive Disease in Prostate Cancer (CHAARTED), reportando altos beneficios del tratamiento quimiohormonal. El ensayo clínico STAMPEDE (Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer - radioterapia al tumor primario para el cáncer de próstata metastásico recién diagnosticado: un ensayo aleatorizado controlado de fase 3) [14] demostró una mejor supervivencia en pacientes con cáncer de próstata metastásico LVD que son sometidos a radioterapia del tumor primario. Como se evidencia, en ambas pruebas la carga tumoral es importante para la predicción de la respuesta del tratamiento y habitualmente son evaluados con imágenes convencionales como CT, sin embargo, la PET/CT utilizando radiofármacos de antígeno específico de próstata (PSMA) poseen una precisión superior para la localización del cáncer de próstata en el diagnóstico inicial, su recurrencia bioquímica y el nivel de avance de la enfermedad.

Además de las terapias mencionadas anteriormente, las terapias con  $^{177}\text{Lu}$ -PSMA son una opción terapéutica prometedora para los pacientes con cáncer de próstata en etapa terminal. En donde la respuesta temprana que se presenta al tratamiento (reducción



del volumen) en la primera reclasificación (reestadificación) después de dos ciclos de terapia se correlaciona con una alta eficacia del tratamiento y una supervivencia general (SG) prolongada. Por lo cual, es relevante determinar y cuantificar la reducción de la carga tumoral ya que puede ser un parámetro significativo y valioso para pronosticar la supervivencia general del paciente y la respuesta al tratamiento [15], [16].

### **1.3 Antecedentes**

Recientemente, se han explorado diferentes modelos basados en inteligencia artificial (IA) para una variedad de tareas relacionadas con la planeación en radioterapia [17]. En particular, algunos de estos trabajos se concentran en las imágenes moleculares de PET/CT. Es importante anotar que la construcción de métodos automatizados sobre este tipo de datos constituye un reto significativo, dada la alta variabilidad, baja resolución espacial y los altos niveles de ruido, presentes en este tipo de imagen médica. Diferentes modelos automatizados basados en IA se han explorado en tareas como: la estratificación del riesgo del paciente, la estimación del volumen tumoral bruto (gross tumor volume) (GTV), la estimación del MTV, el cálculo del volumen objetivo clínico (clinical target volume) (CTV), la estimación del volumen objetivo planificado (planning target volume) (PTV) y delineación de órganos en riesgo (organs at risk) (OAR), entre otras.

Los abordajes más comunes a estos problemas incluyen, la umbralización de los niveles de intensidad, el crecimiento de regiones, la construcción de clasificadores, el uso de algoritmos de agrupamiento y/o detección de bordes, entre otros [18]. Sin embargo, durante los últimos años se ha evidenciado una tendencia creciente en el uso de métodos basados en las técnicas de aprendizaje de máquina, y en particular aprendizaje profundo, por su alta precisión y posibilidad de mejorar la eficiencia del flujo de trabajo clínico [18].

Durante los últimos años se han realizado trabajos basados en aprendizaje de máquina para la segmentación del tumor, lesiones y la carga metabólica tumoral [18]. Sin embargo, es importante tener en cuenta que la mayoría de estos trabajos se centran en la imagen de Fluorodesoxiglucosa (FDG), que no es específica para el cáncer de próstata y puede acumularse en varios tipos de tumores. En particular, este radiotrazador es utilizado para evaluar la actividad metabólica global de los tumores independientemente de su origen. En contraste, existe una cantidad limitada de trabajos basados en PSMA, y comúnmente se centran en el volumen biológico el cual es diferente de la carga y volumen metabólico, no obstante, estos trabajos han mostrado la factibilidad del uso de estas aproximaciones para la segmentación de las imágenes.

En este contexto, uno de los trabajos más importantes es la delimitación del volumen tumoral con técnicas basadas en aprendizaje de máquina que fue propuesto por Evans et al [19] Este trabajo evaluó la viabilidad y el impacto del uso de un nuevo método avanzado de auto segmentación de PET en la planificación del tratamiento de radioterapia (RT) de cabeza y cuello. Esta segmentación fue realizada utilizando un árbol de decisiones,

el cual fue validado previamente con datos preclínicos. El algoritmo entrenado se aplicó a datos  $^{18}\text{F}$ -FDG-PET/CT de 20 pacientes, y los resultados se compararon con una delimitación realizada por expertos utilizando resonancia magnética nuclear y tomografía computarizada. Los resultados sugieren que el volumen tumoral fue menor en el 70% de los casos y de esta forma se cambió directamente el plan de tratamiento de radioterapia [19].

En otro trabajo Z. Zhong et al investigó el uso y la eficiencia del aprendizaje profundo para la segmentación tumoral simultánea en imágenes de cáncer de pulmón de células no pequeñas [20]. Este trabajó utilizó redes completamente convolucionales (DFCN) para la representación, y su nivel de desempeño fue cuantificado utilizando el algoritmo de estimación del nivel de rendimiento y verdad simultáneo (STAPLE) utilizando las delineaciones de 3 médicos expertos [20]. Los resultados sugieren que la similitud promedio entre CT y PET para el algoritmo propuesto fue de 0,86 y 0,83, respectivamente.

Por su parte, Guo et al [21] propusieron un método para la segmentación automática de tumores pulmonares en imágenes PET/CT. Para la segmentación utilizaron un modelo difuso de campo aleatorio de Markov (MRF). Las segmentaciones realizadas por el método fueron comparadas con segmentaciones manuales realizadas por un oncólogo radioterápico experto. Los resultados de segmentación obtenidos con los dos métodos fueron similares con un coeficiente de similitud de Dice (DSC) de 0,85.

Por otra parte, en el caso específico de cáncer de próstata se han realizado algunos trabajos respecto a la segmentación de la carga metabólica tumoral y la predicción del tratamiento. Por ejemplo [22], introdujo y validó un paquete de software semiautomático denominado qPSMA para la evaluación de la carga tumoral en todo el cuerpo en pacientes con cáncer de próstata utilizando  $^{68}\text{Ga}$ -PSMA en imágenes PET/CT. El método propuesto se basa en la utilización de una máscara para la captación normal junto con un umbral SUV de 3 y un umbral basado en el hígado para segmentar las lesiones óseas y las partes blandas. Por su parte, F. Barbato et al [22], Implementaron el software qPSMA en un ensayo que buscaba cuantificar la carga tumoral a través de las captaciones de  $^{68}\text{Ga}$ -PSMA en imágenes PET/CT en 105 pacientes con cáncer de próstata hormonosensible metastásico (mHSPC). En particular, este trabajo estableció un valor umbral para diferenciar la carga tumoral y permitir propuestas alternativas al tratamiento del paciente. Los ensayos clínicos basados en esta aproximación obtuvieron un área bajo la curva roc (AUC) de 0,86 [12].

Los trabajos descritos y otros diversos autores [23]–[26] han determinado varios enfoques en los que se evidencia la viabilidad de delimitar de forma automática el MTB y los MTV. Sin embargo, los métodos utilizados todavía requieren ajustes para mejorar la precisión y el coeficiente de similitud. Por ejemplo, estos métodos tienden a subestimar la captación en una lesión, lo cual puede resultar en una identificación y evaluación erróneas de la respuesta al tratamiento. En este sentido, resulta conveniente explorar alternativas más precisas que además no dependan del umbral de captación. Adicionalmente, se requiere que estas aproximaciones sean evaluadas en poblaciones con un mayor número

de pacientes para elevar la fiabilidad de los experimentos. Finalmente, es importante anotar que la gran mayoría de estas herramientas no son de código abierto, ni están disponibles públicamente, lo que dificulta su validación externa.

## 1.4 Planteamiento del problema

En la práctica clínica se utilizan las imágenes producidas por el PET/CT como una herramienta clave para determinar el estadio y la carga metabólica tumoral del cáncer de próstata. Sin embargo, estas imágenes tienen baja resolución y altos niveles de ruido, lo que dificulta la delimitación precisa y exacta del MTB y el MTV llevando a consumir gran cantidad de tiempo [27].

Los métodos usuales para la delimitación de estas variables son interpretación visual y el uso de umbrales en software semiautomáticos. En la interpretación visual, los bordes de la lesión o tumor se estiman como resultado de una revisión visual exhaustiva por parte del médico nuclear. Este método es susceptible a sesgos de interpretación, y puede ser influenciado por el nivel de experiencia de quien inspecciona la imagen, además de las diversas variables que afectan la interpretación como la capacidad visual, estrés, nivel de fatiga, monotonía de la tarea, las cargas de trabajo, entre otras, que conllevan las labores de intervención humana. Por ser una tarea manual tiene una alta variabilidad dependiendo del observador.

Alternativamente, es posible utilizar un valor de umbral límite para decidir qué región es de interés en una imagen [26], [27]. En principio, este umbral permite distinguir el MTB que serán incluido. Además, este umbral es seleccionado a partir de la concentración del material radiactivo. Sin embargo, la selección de este umbral depende del individuo. En donde al cambiar el umbral cambia el tamaño del MTB seleccionado lo que permite que existan imprecisión ya que será altamente dependiente del umbral seleccionado.

En el cáncer de próstata metastásico la tarea de delimitación del volumen del tumor tiene una dificultad elevada ya que estadios altos de la enfermedad comúnmente poseen gran cantidad de metástasis, lo que agrava la cuantificación de la carga de la enfermedad. Por esta razón, habitualmente esta cuantificación que se realiza de forma manual consume gran cantidad de tiempo y puede representar una limitante en la determinación de los tratamientos del paciente.

Debido a estas problemáticas, durante los últimos años se han explorado aproximaciones basadas en aprendizaje de máquina que han demostrado tener mejor precisión y exactitud en la delimitación del volumen y carga metabólica del tumor. Estas aproximaciones pueden constituirse en una herramienta que permita disminuir las brechas de las tareas en las que hay intervención e interpretación humana. Por este motivo, en este trabajo se propone la implementación de algoritmos de aprendizaje de máquina para optimizar el proceso de delimitación del MTB y MTV ya que responde a la necesidad de los

centros oncológicos de tener un método validado, reproducible y de alta precisión para la su delimitación y cuantificación en pacientes con cáncer de próstata metastásico.

## 1.5 Pregunta de investigación

En este contexto y en relación a los enfoques mencionados anteriormente, se plantea la siguiente pregunta de investigación: **¿Cómo cuantificar de forma automática la carga metabólica tumoral en pacientes con diagnóstico de carcinoma de próstata a través de imágenes PET/CT de  $^{68}\text{Ga}$ -PSMA?**

Para responder a esta pregunta de investigación se propone la siguiente hipótesis: **La implementación de algoritmos de aprendizaje de máquina permite cuantificar la carga metabólica tumoral de forma automática en imágenes PET/CT de  $^{68}\text{Ga}$ -PSMA.**

## 1.6 Objetivos

Con el fin de abordar la hipótesis propuesta se proponen los siguientes objetivos de investigación:

### 1.6.1 General

Cuantificar de forma automática la carga metabólica tumoral a través de algoritmos de aprendizaje de máquina en imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA en pacientes diagnosticados con carcinoma de próstata.

### 1.6.2 Específicos

Para abordar el objetivo general propuesto se proponen los siguientes objetivos específicos:

- Construir un conjunto de datos de imágenes de tomografía por emisión de positrones de  $^{68}\text{Ga}$ -PSMA de pacientes con cáncer de próstata que incluya la delimitación de la carga metabólica del tumor.
- Implementar un algoritmo de aprendizaje de máquina para la delimitación de la carga metabólica del tumor.
- Evaluar el rendimiento del algoritmo en la delimitación de la carga metabólica del tumor.

## 1.7 Contribuciones

Las contribuciones realizadas durante el desarrollo de este trabajo son las siguientes:

- La principal contribución de este trabajo es un modelo para la detección automática de captaciones tumorales y metástasis asociadas al carcinoma de próstata en exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA que permite la cuantificación de la carga metabólica tumoral. Para ello el modelo incluye captaciones en la pelvis de próstata y ganglios. Esta cuantificación permite al médico nuclear establecer futuros tratamientos y la estadificación volumétrica de la enfermedad.
- Un conjunto de datos de 100 exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA de pacientes que han sido sometidos a la prueba, los cuales contienen anotaciones según la presencia o ausencia de las captaciones anómalas asociadas al carcinoma de próstata.
- La contribución dirigida a la aprobación social del conocimiento se realizó una ponencia de socialización de resultados en el evento II Congreso Colombiano de Física Médica, entre el 3 y 4 de noviembre de 2023.
- Se realizará el sometimiento de los resultados más importantes del trabajo a una revista indexada. Categoría Q1.
- Se realizará el respectivo Registro Software ante la Dirección Nacional de Derechos de Autor.

## **Capítulo 2.**

### **Carcinoma de próstata y aprendizaje profundo**

#### **2.1 Carcinoma de próstata**

El cuerpo humano está compuesto por trillones de células que durante toda la vida crecen normalmente y se dividen según sea necesario. Cuando las células son anormales o envejecen, por lo general mueren. El cáncer comienza cuando algo sale mal en este proceso y las células siguen produciendo nuevas células y las viejas o anormales no mueren cuando deberían. A medida que las células cancerosas crecen sin control, pueden desplazar a las células normales lo que dificulta que el cuerpo funcione correctamente [28].

El cáncer de próstata comienza cuando las células de la glándula prostática crecen sin control. La próstata está debajo de la vejiga y enfrente del recto. Casi todos los cánceres de próstata son adenocarcinomas y se desarrollan a partir de las células de las glándulas (las células que producen el líquido prostático que se agrega al semen). Existen otros tipos de cáncer que pueden comenzar en la próstata, sin embargo, son poco comunes: carcinomas de células pequeñas, tumores neuroendocrinos (que no sean carcinomas de células pequeñas), carcinomas de células de transición y sarcomas [28].

El cáncer de próstata puede no causar signos o síntomas en sus primeras etapas mientras que el cáncer de próstata avanzado puede causar signos y síntomas como: dificultad para orinar, disminución de la fuerza en el chorro de orina, sangre en la orina, sangre en el semen, dolor óseo, pérdida peso e impotencia [29].

Por lo general, los cánceres de próstata en etapas iniciales no causan síntomas, en cambio los cánceres más avanzados se pueden detectar debido a los síntomas que causan. La mayoría de los cánceres de próstata se detectan primero al hacer la prueba de sangre para detectar el antígeno prostático específico (Prostate-Specific Antigen, PSA) o mediante el tacto rectal. El diagnóstico definitivo de cáncer de próstata se puede llevar a cabo únicamente mediante una biopsia de la próstata [30].

### 2.1.1 Estudios a través de imágenes médicas

Los estudios por imágenes utilizan ondas sonoras, rayos X, imagen por resonancia magnética, gammagrafía ósea y tomografía por emisión de positrones del interior del cuerpo. Se puede emplear uno o más estudios por imágenes para:

- Buscar cáncer en la próstata.
- Ayudar al médico a observar la próstata durante ciertos procedimientos (como una biopsia de próstata o ciertos tipos de tratamiento para el cáncer de próstata).
- Buscar la propagación del cáncer de próstata a otras partes del cuerpo.

El paciente que es sometido a una exploración PET/CT de  $^{68}\text{Ga}$ -PSMA que incluye las siguientes indicaciones del estudio:

- Localización del tejido tumoral en cáncer de próstata recurrente.
- Estadificación primaria en la enfermedad de alto riesgo antes de la cirugía, procedimientos o planificación de radiación externa.
- Estadificación antes y durante la radioterapia dirigida.
- Biopsia dirigida después de una biopsia negativa previa en pacientes con alta sospecha de cáncer de próstata.
- Monitorización del tratamiento sistémico en cáncer de próstata metastásico.

## 2.2 Carga metabólica tumoral

La carga tumoral metabólica, también conocida como "metabolic tumor burden" (MTB) en inglés, se refiere a la cantidad y actividad metabólica de las células tumorales en el cuerpo. Es una medida que evalúa la cantidad y el grado de actividad de los tumores en un individuo [31].

Cuando se forma un tumor, las células cancerosas experimentan cambios metabólicos significativos. En comparación con las células normales, las células cancerosas a menudo tienen una tasa metabólica más alta y dependen de diferentes vías metabólicas para obtener energía y nutrientes necesarios para su crecimiento y proliferación. Estos cambios metabólicos pueden incluir un aumento en la captación de glucosa, una mayor producción de lactato y una mayor utilización de diferentes metabolitos.

La carga tumoral metabólica se puede evaluar mediante técnicas de imagen molecular, como la tomografía por emisión de positrones (PET) con un trazador metabólico como la glucosa marcada con fluorodeoxiglucosa (FDG), PSMA, entre otras. Esta técnica

permite detectar y cuantificar la captación de glucosa en los tejidos tumorales, lo que proporciona información sobre la actividad metabólica de los tumores (Figura 1.1) [32].

Además, la carga tumoral metabólica es una medida importante en la evaluación de la respuesta al tratamiento del cáncer. Una mayor carga metabólica tumoral generalmente se correlaciona con un mayor grado de malignidad y puede indicar una enfermedad más avanzada. Asimismo, la disminución de la carga tumoral metabólica después del tratamiento puede ser un indicador de respuesta favorable [33]. El MTB ha demostrado ser un predictor pronóstico independiente de la supervivencia en pacientes con cáncer de pulmón y otros tipos de cáncer. Además, el análisis de las diferencias metabólicas entre los tejidos cancerosos y no cancerosos puede proporcionar información valiosa sobre la biología del tumor, que está estrechamente relacionada con el pronóstico y el diagnóstico de la enfermedad [34].

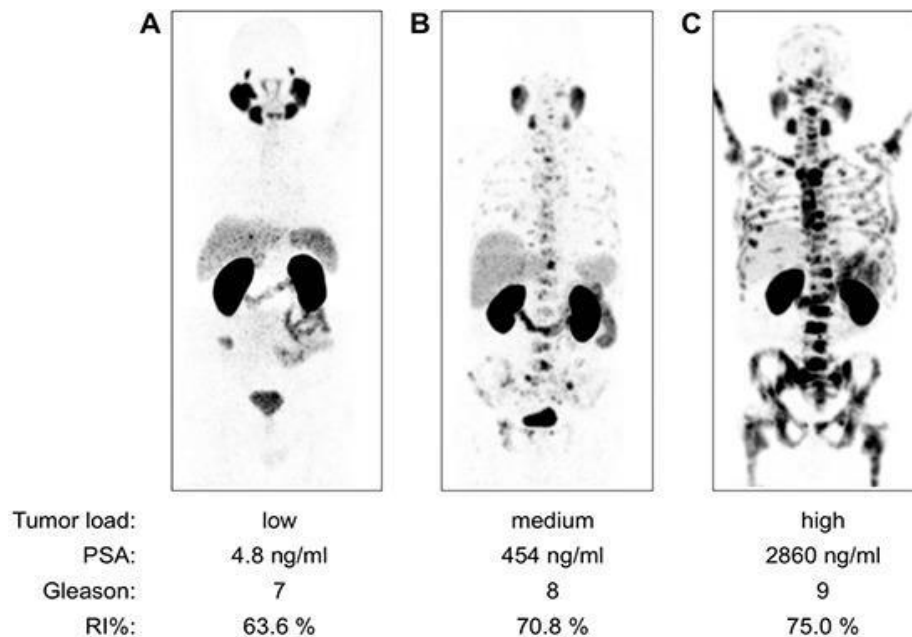


Figura 2.1 Ejemplo de pacientes con carga tumoral baja, media y alta en imágenes PET de  $^{68}\text{Ga}$ -PSMA [35].

Como se observa en la Figura 2.1, en el caso del paciente (A), clasificado visualmente con una carga tumoral baja, la imagen PET/CT reveló la presencia de dos metástasis óseas positivas para PSMA en la pelvis. Se observa una captación fisiológicamente elevada en las glándulas lagrimales, glándulas salivales (parótidas, submandibulares y sublinguales) y en los riñones. En cuanto al paciente (B), clasificado visualmente con una carga tumoral media, se identificaron metástasis óseas dispersas con una acumulación del trazador principalmente leve a moderada, así como múltiples metástasis de ganglios linfáticos positivas para PSMA en la pelvis, abdomen y tórax. Por otro lado, en el paciente (C),



clasificado visualmente con una carga tumoral alta, se encontró una intensa acumulación de  $^{68}\text{Ga}$ -PSMA-11 en las metástasis óseas dispersas. Además, se observó una pequeña recurrencia local positiva para PSMA y pequeños ganglios linfáticos mediastínicos [36].

## 2.3 Medicina nuclear

La ciencia y la práctica clínica de la medicina nuclear implican la administración de trazas de compuestos marcados con radiactividad (radiofármacos) que se utilizan para proporcionar información de diagnóstico en una amplia variedad de estados patológicos. Aunque los radiofármacos también tienen algunos usos terapéuticos, con principios físicos subyacentes similares, este trabajo se centra en los usos diagnósticos de los radionucleidos en la medicina moderna [37].

En su forma más básica, un estudio de medicina nuclear consiste en inyectar un compuesto, que está marcado con un radionúclido emisor de rayos gamma o emisor de positrones en el cuerpo. El compuesto radiomarcado se denomina radiofármaco, o más comúnmente, trazador o radiotrazador. Cuando el radionúclido se desintegra, se emiten rayos gamma o fotones de alta energía. La energía de estos rayos gamma o fotones es tal que un número significativo puede salir del cuerpo sin dispersarse ni atenuarse. Una “cámara” externa de rayos gamma sensible a la posición puede detectar los rayos gamma o los fotones y formar una imagen de la distribución del radionúclido y, por lo tanto, del compuesto (incluidos los productos radiomarcados de las reacciones de ese compuesto) al que estaba adherido [37].

Hay dos tipos de imágenes de medicina nuclear: imágenes de fotón único [que incluye tomografía computarizada por emisión de fotón único (SPECT)] e imágenes de positrones [tomografía por emisión de positrones (PET)]. Las imágenes de positrones utilizan radionúclidos que se descomponen por emisión de positrones. El positrón emitido tiene una vida útil muy corta y, tras la aniquilación con un electrón, produce simultáneamente dos fotones de alta energía que posteriormente son detectados por una cámara de imágenes. Las imágenes tomográficas se forman recopilando datos desde muchos ángulos alrededor del paciente, lo que da como resultado imágenes PET [37].

La ventaja de la medicina nuclear radica en su capacidad para proporcionar medidas muy sensibles de una amplia gama de procesos biológicos en el cuerpo. Otras modalidades de imágenes médicas, como la resonancia magnética nuclear (RMN), las imágenes de rayos X y la tomografía computarizada (TC) de rayos X, proporcionan imágenes anatómicas sobresalientes, pero tienen una capacidad limitada para proporcionar información biológica. Por ejemplo, los métodos de resonancia magnética generalmente tienen un límite inferior de detección en el rango de concentración milimolar ( $\approx 6 \times 10^7$  moléculas

por ml de tejido), mientras que los estudios de medicina nuclear detectan rutinariamente sustancias radiomarcadas en el rango nanomolar ( $\approx 6 \times 10^{11}$  moléculas por ml de tejido) o rango picomolar ( $\approx 6 \times 10^8$  moléculas por ml de tejido). Esta ventaja de sensibilidad, junto con la selección cada vez mayor de compuestos radiomarcados, permite que los estudios de medicina nuclear se centren en los procesos biológicos muy específicos que subyacen a la enfermedad [37].

## 2.4 Radiofármaco $^{68}\text{Ga}$ -PSMA

El antígeno de membrana específico de la próstata (PSMA) se utiliza cada vez más como una diana molecular en imágenes clínicas y terapia con radionucleidos para el cáncer de próstata. Esta proteína se encuentra normalmente en la superficie de las células de la próstata y es más abundante en las células cancerosas de la próstata, lo que ayuda a detectar células cancerosas, especialmente cuando hay sospecha de recurrencia del cáncer o metástasis a otras partes del cuerpo [38].

Debido a su similitud con el Fe(III), los complejos de Ga(III) están emergiendo como una alternativa interesante a los agentes anticancerígenos. Desde el punto de vista del diagnóstico, las versiones de Ga(III) que emiten positrones se pueden utilizar para la obtención de imágenes de tumores. El  $^{68}\text{Ga}$  se obtiene a partir de un generador de  $^{68}\text{Ge}/^{68}\text{Ga}$  ( $^{68}\text{Ge}$ ,  $t_{1/2} = 270,8$  días), lo que lo hace independiente de un ciclotrón in situ. Por lo tanto, los agentes de PET basados en Ga-68 poseen un potencial comercial significativo y sirven como una alternativa conveniente a los isótopos basados en ciclotrones para la tomografía por emisión de positrones (PET), como el  $^{18}\text{F}$  o el  $^{124}\text{I}$  [38].

El  $^{68}\text{Ga}$  tiene una alta fracción emisora de positrones (89% de su decaimiento total). La energía máxima de positrones de  $^{68}\text{Ga}$  (energía máxima = 1,92 MeV, media = 0,89 MeV) es mayor que la de  $^{18}\text{F}$  (máxima = 0,63 MeV, media = 0,25 MeV). Sin embargo, un estudio de resolución espacial utilizando el análisis de Monte Carlo reveló que, bajo el supuesto de una resolución espacial de 3 mm para la mayoría de los detectores PET, los valores del ancho a la mitad de la altura (FWHM) del  $^{18}\text{F}$  y  $^{68}\text{Ga}$  son indistinguibles en los tejidos blandos (3,01 mm frente a 3,09 mm). Ese hallazgo implica que con la resolución espacial estándar de 5-7 mm para los escáneres clínicos actuales, la calidad de la imagen que utiliza radiotrazadores basados en Ga-68 probablemente no se distinga de los basados en F-18, lo que estimula el interés en el desarrollo de compuestos marcados con  $^{68}\text{Ga}$  para imágenes médicas. Con una vida media física de 68 min, el  $^{68}\text{Ga}$  también se adapta muy bien a la farmacocinética de muchos péptidos utilizados para imágenes. Además, el  $^{68}\text{Ga}$  se introduce en las biomoléculas a través de quelantes macrocíclicos, lo que permite la posible formulación de

kits y una amplia disponibilidad de los agentes de formación de imágenes correspondientes [38].

Las imágenes PET/CT con ligandos de  $^{68}\text{Ga}$ -PSMA son un método altamente sensible para la detección de tumores primarios, así como de lesiones localmente recurrentes o metastásicas después del tratamiento del cáncer de próstata primario. En consecuencia, varios ligandos de PSMA se utilizan en la práctica clínica. Se ha demostrado recientemente que la imagenología y la terapia (I&T) con  $^{68}\text{Ga}$ -PSMA localizan el cáncer de próstata primario y proporcionan altas tasas de detección en el cáncer de próstata recurrente. PSMA I&T se puede etiquetar con  $^{68}\text{Ga}$  o  $^{177}\text{Lu}$  con fines teragnósticos [10].

Por otra parte, es importante anotar que existen captaciones fisiológicas que son benignas relacionadas con el metabolismo del radiofármaco como se observa en la Figura 1.2, como en la corteza renal, duodeno, íleon, parótida, submandibular y glándulas lagrimales, hígado, riñones y vejiga [39].



Figura 2.2 Biodistribución del  $^{68}\text{Ga}$ —PSMA-11 [39].

En la Figura 2.2 se ilustra un examen de PET/CT que muestra el patrón de distribución de  $^{68}\text{Ga}$ -PSMA-11 en un paciente de 64 años que tenía un nivel de PSA creciente de 1,06 ng/ml después de la cirugía (prostatectomía radical, Gleason Score = 8) y fue derivado para reestadificación [39].

## 2.5 Tomografía por emisión de positrones (PET/CT)

La Tomografía por Emisión de Positrones, es un examen diagnóstico que se utiliza para obtener imágenes funcionales de todo el cuerpo. Se realiza a través de la administración intravenosa de moléculas orgánicas similares a las naturales, marcadas con sustancias radioactivas. La PET/TC combina la PET con la tomografía computarizada (TC) para proporcionar imágenes bidimensionales detalladas que muestran la anatomía y la funcionalidad [40].

La imagen PET se basa en la detección de los positrones, que son partículas subatómicas emitidas por ciertos elementos radiactivos. Estos elementos radiactivos se conocen como radionúclidos y se combinan con compuestos químicos para formar los trazadores radiactivos utilizados en la PET. Los trazadores se inyectan en el cuerpo del paciente y se distribuyen en los tejidos en función de su actividad metabólica[37].

Cuando los positrones emitidos por los radionúclidos se encuentran con electrones en el cuerpo, se produce una aniquilación mutua, generando dos fotones gamma. Estos fotones gamma son detectados por una cámara especial, que está compuesta por una matriz de cristales detectores y fotomultiplicadores. La información recopilada por el detector PET se utiliza para reconstruir imágenes tridimensionales del cuerpo [37].

La clave de la PET radica en los trazadores radiactivos. Estos trazadores se diseñan para dirigirse a procesos metabólicos específicos en el cuerpo, como la captación de glucosa en las células cancerosas o el flujo sanguíneo cerebral. El trazador más comúnmente utilizado en la PET es la glucosa marcada con fluorodeoxiglucosa (FDG), que se acumula en las células con alta actividad metabólica, como las células cancerosas [37].

Las imágenes de PET muestran la distribución del trazador radiactivo en el cuerpo, lo que permite detectar y visualizar áreas de alta actividad metabólica. Esto es especialmente útil en la detección y seguimiento de enfermedades como el cáncer, donde la actividad metabólica anormal puede indicar la presencia de tumores, su extensión y la respuesta al tratamiento [37].

Además de la PET convencional, también se utiliza la PET/CT, que combina la tomografía por emisión de positrones con la tomografía computarizada (CT). La PET/CT fusiona las imágenes metabólicas de la PET con las imágenes anatómicas de la CT, lo que proporciona una localización precisa de las áreas con actividad metabólica anormal [37].

## 2.6 Método de reconstrucción iterativo

En esta sección se analiza el método de reconstrucción iterativa, ampliamente empleado en los exámenes PET/CT y el elegido como algoritmo de reconstrucción de las imágenes.

El método de reconstrucción iterativa en PET/CT es una técnica avanzada utilizada para mejorar la calidad de las imágenes obtenidas en los exámenes. Este método se basa en la aplicación de algoritmos iterativos que mejoran la precisión y reducen los artefactos en las imágenes [37]. El proceso de reconstrucción iterativa consta de los siguientes pasos:

- **Adquisición de datos:** Durante el examen PET/CT, el paciente es inyectado con una sustancia radiactiva, que emite positrones que son detectados por el sistema de PET. Al mismo tiempo, se realiza una tomografía computarizada (CT) para obtener imágenes anatómicas detalladas.
- **Sinograma:** Los datos capturados durante el escaneo se almacenan en un sinograma, que es una representación matemática de la información recopilada por los detectores de PET.
- **Proceso iterativo:** El proceso de reconstrucción iterativa comienza con una estimación inicial de la distribución de radioactividad en el cuerpo del paciente y se repite iterativamente varias veces hasta que se obtenga una imagen final más precisa.
- **División en subconjuntos:** Los datos capturados durante el escaneo se dividen en varios subconjuntos, lo que significa que solo se utiliza una parte de los datos en cada iteración del proceso de reconstrucción.
- **Actualización de la imagen:** En cada iteración, el algoritmo compara la imagen reconstruida con el sinograma adquirido. Luego, ajusta los valores de la imagen para minimizar la diferencia entre la proyección simulada y la proyección real.
- **Regularización:** Para mejorar la estabilidad y la calidad de la reconstrucción, se aplica una técnica de regularización que penaliza soluciones no suaves o con ruido excesivo. Esto evita que el algoritmo amplifique el ruido y mejora la calidad general de la imagen.

En la Figura 2.3, se aprecia el proceso iterativo y la progresiva evolución de la reconstrucción. A medida que avanzamos en las iteraciones, es notable cómo la representación del objeto se aproxima cada vez más a la realidad.

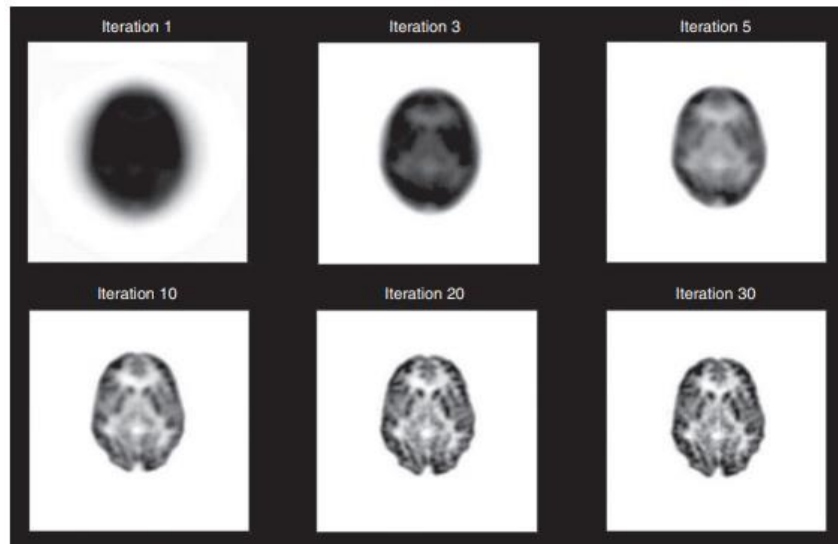


Figura 2.3 Ejemplo de una adquisición con diferentes reconstrucciones iterativas [37].

### 2.6.1 Parámetros en la reconstrucción iterativa

Los parámetros específicos para la reconstrucción iterativa por subconjuntos pueden variar dependiendo del software y el equipo utilizado en el sistema de PET. Sin embargo, a continuación, se presentan algunos de los parámetros comunes que se pueden encontrar en esta técnica:

- **Número de subconjuntos:** Este parámetro indica en cuántos grupos se dividirán los datos adquiridos durante el escaneo PET. Un mayor número de subconjuntos podría mejorar la calidad de la reconstrucción, pero también aumentaría el tiempo de procesamiento. Se debe encontrar un equilibrio entre la calidad de la imagen y la eficiencia computacional.
- **Número de iteraciones por subconjunto:** Indica cuántas veces se repetirá el proceso de actualización de la imagen en cada subconjunto. Un mayor número de iteraciones podría mejorar la precisión de la reconstrucción, pero también aumentaría el tiempo de procesamiento. Es importante encontrar un valor adecuado que logre un buen balance entre precisión y eficiencia.

- **Algoritmo de reconstrucción iterativa:** Existen diferentes algoritmos de reconstrucción iterativa que se pueden aplicar a cada subconjunto. Algunos ejemplos incluyen el algoritmo MLEM (Máxima Verosimilitud Esperada), OSEM (Estimación de Máxima Verosimilitud Ordenada), y AW-OSEM (Estimación de Máxima Verosimilitud Ordenada con Pesos Anatómicos). Cada algoritmo puede tener sus propias configuraciones y ajustes específicos.
- **Tamaño del píxel:** La resolución espacial de la imagen final dependerá del tamaño del píxel utilizado en la reconstrucción. Un tamaño de píxel más pequeño mejorará la resolución espacial, pero también aumentará el tiempo de procesamiento y el tamaño del archivo de imagen.
- **Parámetros de regularización:** Estos parámetros controlan la cantidad de suavizado o regularización aplicada en el proceso iterativo. La regularización es una técnica utilizada para mejorar la estabilidad y la calidad de la reconstrucción, penalizando soluciones no suaves o con ruido excesivo. Ajustar estos parámetros puede afectar la nitidez y los detalles en la imagen final.
- **Kernel de filtrado:** Es un filtro que se aplica en cada iteración para reducir el ruido en la imagen. La elección del kernel de filtrado puede afectar la calidad y los detalles en la imagen reconstruida.

Es importante tener en cuenta que la optimización de estos parámetros depende del objetivo específico de la reconstrucción y de las características del sistema de PET utilizado. La selección adecuada de los parámetros es crucial para obtener imágenes de alta calidad y una cuantificación precisa de la actividad metabólica en los estudios de PET/CT.

## 2.7 Aprendizaje de máquina

El aprendizaje automático es la ciencia (y el arte) de programar computadoras para que puedan aprender de los datos [41].

Una definición un poco más general sería:

“El aprendizaje automático es el campo de estudio que brinda a las computadoras la capacidad de aprender sin ser programadas explícitamente”. —Arthur Samuel, 1959

Y una más orientado a la ingeniería:

“Se dice que un programa de computadora aprende de la experiencia E con respecto a alguna tarea T y alguna medida de desempeño P, si su desempeño en T, medido por P, mejora con la experiencia E”. —Tom Mitchell, 1997

En resumen, es el subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que las computadoras aprendan. Se dice que un agente aprende cuando su desempeño mejora con la experiencia y mediante el uso de datos; es decir, cuando la habilidad no estaba presente en su genotipo o rasgos de nacimiento [42]. "En el aprendizaje de máquinas un computador observa datos, construye un modelo basado en esos datos y utiliza ese modelo a la vez como una hipótesis acerca del mundo y una pieza de software que puede resolver problemas" [43]. Por lo tanto, el aprendizaje automático es ideal para:

- Problemas para los cuales las soluciones existentes requieren muchos ajustes o largas listas de reglas: un algoritmo de aprendizaje automático a menudo puede simplificar el código y funcionar mejor que el enfoque tradicional.
- Problemas complejos para los que el uso de un enfoque tradicional no ofrece una buena solución: las mejores técnicas de Machine Learning quizás puedan encontrar una solución.
- Entornos fluctuantes: un sistema de Machine Learning puede adaptarse a nuevos datos.
- Obtener información sobre problemas complejos y grandes cantidades de datos.

Hay diferentes tipos de algoritmos de aprendizaje de máquina, sin embargo, en este trabajo se implementa el supervisado.

### **2.7.1 Aprendizaje supervisado**

El aprendizaje supervisado consiste en crear una función capaz de predecir el valor correspondiente a cualquier objeto de entrada válida después de haber examinado una serie de ejemplos, los datos de entrenamiento. Para ello, tiene que generalizar a partir de los datos presentados a las situaciones no examinadas previamente. Esta técnica de aprendizaje solo se puede aplicar en datos que son etiquetados [44].

## **2.8 Redes neuronales convolucionales**

Las redes neuronales convolucionales (Convolutional Neural Network, CNN) surgieron del estudio de la corteza visual del cerebro y se han utilizado en el reconocimiento de imágenes desde la década de 1980. En los últimos años, gracias al aumento del poder



computacional, la cantidad de datos de entrenamiento disponibles y diferentes metodologías para entrenar redes profundas (Deep learning), las CNNs han logrado un rendimiento sobrehumano en algunas tareas visuales complejas. Impulsan servicios de búsqueda de imágenes, automóviles autónomos, sistemas automáticos de clasificación de videos y entre otras [45].

Además, las CNN se utilizan principalmente en diferentes tareas visuales como la detección de objetos que consiste en clasificar varios objetos en una imagen y colocar cuadros delimitadores a su alrededor y la segmentación semántica que consiste en clasificar cada píxel de una imagen según la clase del objeto que representa, este trabajo se centra en la segmentación semántica (ver Figura 2.4) [45].

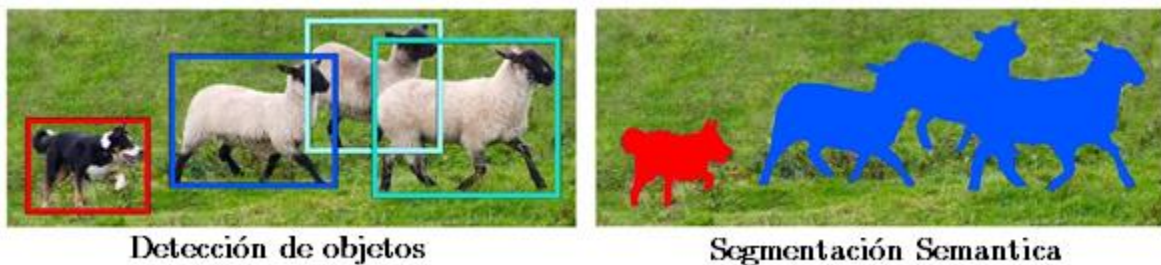


Figura 2.4 Detección de objetos vs segmentación semántica [46].

Las redes neuronales convolucionales son un tipo especializado de redes neuronales artificiales que utilizan una operación matemática llamada convolución en lugar de la multiplicación general de matrices en al menos una de sus capas que están diseñados específicamente para procesar datos de píxeles [47].

### 2.8.1 Arquitectura general de las CNN

Una CNN consta de una capa de entrada, capas ocultas y una capa de salida (Figura 1.6). Donde, las capas intermedias se denominan ocultas porque sus entradas y salidas están enmascaradas por la función de activación y la convolución final. En una red neuronal convolucional, las capas ocultas incluyen capas que realizan convoluciones. Una convolución es un operador matemático que transforma dos funciones  $f$  y  $g$  en una tercera función que en cierto sentido representa la magnitud en la que se superponen  $f$  y una versión trasladada e invertida de  $g$  (Figura 2.5).

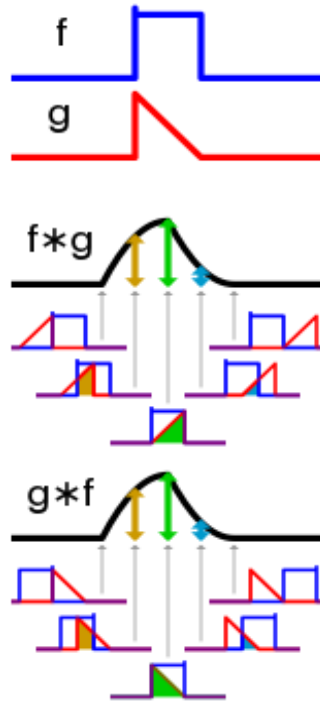


Figura 2.5 Analogía de la operación de convolución [48].

Además, las capas ocultas incluyen una capa que realiza un producto escalar del kernel de convolución con la matriz de entrada de la capa. Este producto suele ser el producto interno de Frobenius, y su función de activación es comúnmente ReLU. A medida que el kernel de convolución se desliza a lo largo de la matriz de entrada de la capa, la operación de convolución genera un mapa de características, que a su vez contribuye a la entrada de la siguiente capa. A esto le siguen otras capas, como las capas de agrupación, las capas totalmente conectadas y las capas de normalización.

En las capas convolucionales la entrada es un tensor con una forma: (número de imagen)  $\times$  (altura de la imagen)  $\times$  (ancho de la imagen)  $\times$  (canales de la imagen). Después de pasar por una capa convolucional, la imagen se abstrae en un mapa de características, también llamado mapa de activación, con forma: (número de entradas)  $\times$  (altura del mapa de características)  $\times$  (ancho del mapa de características)  $\times$  (canales del mapa de características). Las capas convolucionales convolucionan la entrada y pasan su resultado a la siguiente capa [49].

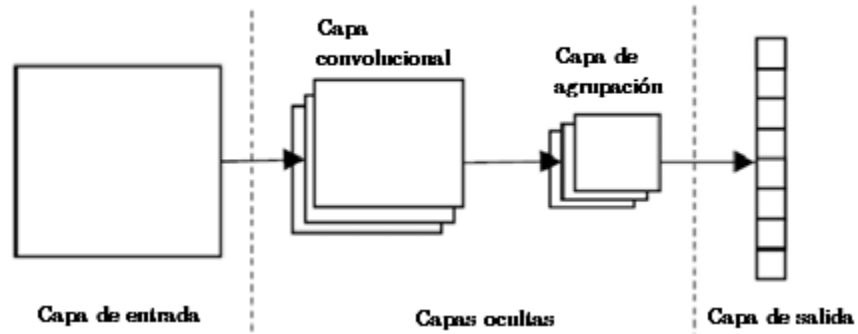


Figura 2.6 Arquitectura general de una red neuronal convolucional [50].

Otro factor fundamental son los estados de cada capa que se organizan de acuerdo con una estructura matricial espacial. Estas relaciones espaciales se heredan de una capa a la siguiente porque cada valor de característica se basa en una pequeña región espacial local en la capa anterior. Es importante mantener estas relaciones espaciales entre las celdas de la matriz, porque la operación de convolución y la transformación a la siguiente capa dependen de manera crítica de estas relaciones (Figura 2.7) [41].

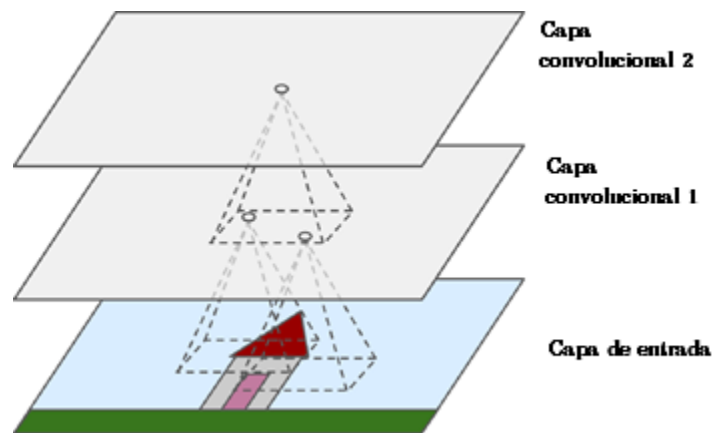


Figura 2.7 Ejemplos de capas convolucionales y mapeo espacial [41].

Una neurona ubicada en la fila  $i$ , columna  $j$  de una capa determinada se conecta a las salidas de las neuronas de la capa anterior ubicadas en las filas  $i$  a  $i + f_h - 1$ , columnas  $j$  a  $j + f_w - 1$ , donde  $f_h$  y  $f_w$  son la altura y el ancho del campo receptivo (ver Figura 1.8). Para que una capa tenga el mismo alto y ancho que la capa anterior, es común agregar ceros alrededor de las entradas, como se muestra en el diagrama [41].

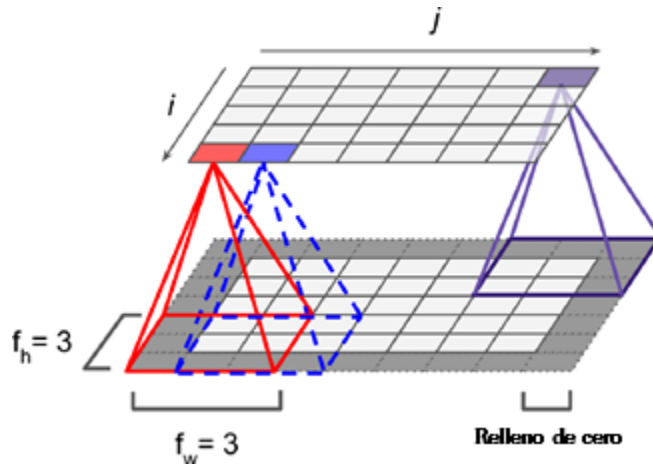


Figura 2.8 Mapeo espacial de la matriz [41].

Otra parte fundamental de la arquitectura son las capas de agrupación que tienen como objetivo submuestrear, es decir, reducir la imagen de entrada para reducir la carga computacional, el uso de memoria y la cantidad de parámetros (lo que limita el riesgo de sobreajuste). Las capas de agrupación todo lo que hacen es agregar las entradas usando una función de agregación como el máximo o la media (Figura 2.9) [41].

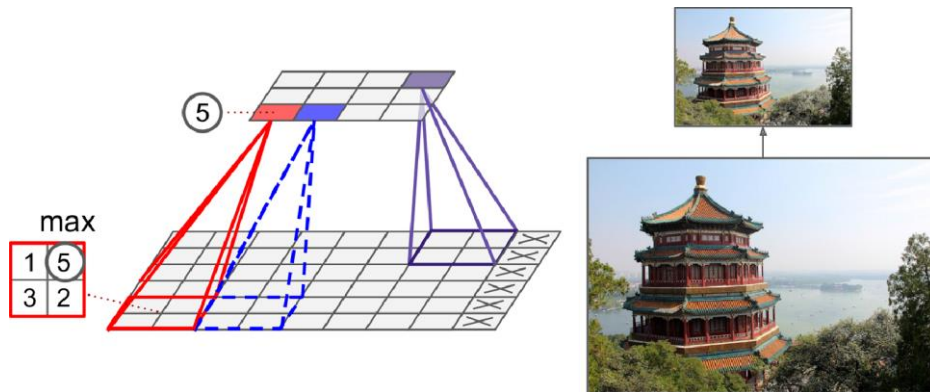


Figura 12.9 Operaciones de agrupación [41].

La figura 1.9 muestra una capa de agrupación máxima (max pooling layer), que es el tipo más común de capa de agrupación. En este ejemplo, usamos un núcleo de agrupación de 2 x 2, con un paso de 2 y sin relleno. Solo el valor de entrada máximo en cada campo receptivo pasa a la siguiente capa, mientras que las otras entradas se eliminan [41].

## 2.9 Transferencia de aprendizaje y ajuste fino

Transferir el aprendizaje (transfer learning en inglés) consiste en tomar características o conocimientos previos aprendidos en un problema y aprovecharlas en un problema o modelo nuevo similar. En aprendizaje profundo generalmente se realiza para tareas en las que su conjunto de datos tiene muy pocos datos para entrenar un modelo a gran escala desde cero [51]. En lugar de comenzar el entrenamiento de un modelo desde cero, el aprendizaje por transferencia aprovecha el conocimiento previamente adquirido por un modelo ya entrenado para realizar una tarea específica. El modelo pre-entrenado se puede considerar como una especie de "conocimiento general" que se puede aplicar a una nueva tarea para obtener un mejor rendimiento [51].

El proceso de transferencia de aprendizaje generalmente se divide en dos etapas:

1. Pre-entrenamiento: En esta etapa, un modelo se entrena en un conjunto de datos grande y diverso, generalmente utilizando una arquitectura de red neuronal profunda como una red neuronal convolucional (CNN) o una red neuronal recurrente (RNN). El objetivo del pre-entrenamiento es capturar características generales de los datos, lo que permite al modelo reconocer patrones comunes en los nuevos datos.
2. Afinamiento (fine-tuning): En esta etapa, se ajustan los parámetros del modelo pre-entrenado en un conjunto de datos más pequeño y específico relacionado con la tarea que se quiere realizar. Se utiliza una técnica llamada "congelación" para evitar que se modifiquen las capas de la red neuronal que ya han sido pre-entrenadas, y solo se ajustan las capas finales que están relacionadas con la nueva tarea.

El aprendizaje por transferencia es especialmente útil cuando el conjunto de datos de entrenamiento es limitado o costoso de obtener, ya que permite utilizar modelos previamente entrenados en tareas relacionadas para mejorar el rendimiento en una tarea nueva [51]. Además, el aprendizaje por transferencia ha demostrado ser una técnica efectiva para acelerar el proceso de entrenamiento de un modelo, ya que el modelo pre-entrenado puede utilizarse como una buena inicialización de los parámetros del nuevo modelo [51].

El procedimiento más frecuente del aprendizaje por transferencia en el contexto del aprendizaje profundo es el siguiente flujo de trabajo:

1. Tomar capas de un modelo previamente entrenado.
2. Congelar capas para evitar destruir la información que contienen durante futuras rondas de entrenamiento.

3. Agregar algunas capas nuevas y entrenables encima de las capas congeladas. Las cuales aprenderán a convertir las características antiguas en predicciones sobre un nuevo conjunto de datos.
4. Entrenar las nuevas capas en el conjunto de datos del nuevo problema.
5. Un último paso opcional es el ajuste fino, que consiste en descongelar todo el modelo que se obtuvo anteriormente (o parte de él) y volver a entrenarlo con los nuevos datos con una tasa de aprendizaje muy baja. Esto puede potencialmente lograr mejoras significativas, al adaptar de manera incremental las funciones preentrenadas a los nuevos datos.

## Capítulo 3.

### Materiales y métodos.

#### 3.1 Descripción de la metodología

La metodología adoptada para este trabajo es tipo CRISP-DM (Cross Industry Standard Process of Data Mining) [52], que es el método más utilizado en aprendizaje de máquina y consta de seis etapas (Figura 3.1): comprensión del problema, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Estas etapas se siguen de manera exhaustiva, pero se realizan adaptaciones donde sea necesario.



Figura 3.1 Metodología CRISP-DM [30].

##### 3.1.1 Comprensión del problema

Consiste en identificar los requerimientos y problemáticas asociadas a la delimitación y cuantificación de la carga metabólica del tumor para convertirlos en objetivos técnicos y elaborar un plan de trabajo. Plantear el problema desde el punto de vista del problema y desde el punto de vista del aprendizaje automático.

##### 3.1.2 Comprensión de los datos

En esta fase se realiza la recopilación inicial de datos, donde se construye el conjunto de datos a partir de imágenes PET/CT que contenga la delimitación de la MTB. Adicionalmente, se describen y exploran las imágenes para establecer relaciones, identificar patrones y características propias para finalmente realizar las anotaciones y descripción de los datos. Este proceso se llevó a cabo en el Instituto Nacional de Cancerología (INC) de forma retrospectiva con un total de 100 exámenes de imágenes PET/CT de  $^{68}\text{Ga}$ -PSMA diagnosticados con carcinoma de próstata que incluya la delimitación de la carga metabólica tumoral por parte del médico nuclear.

### **3.1.3 Preparación de los datos**

Posteriormente a la recolección de los datos, se procede a adaptarlos para el proceso de aprendizaje automático, donde se establecen los criterios de limpieza, estandarización y cambios de origen de formato.

### **3.1.4 Modelado**

En esta etapa se seleccionan y aplican diferentes algoritmos de visión por computadora basados en aprendizaje automático, donde se ajustan y se sintonizan sus hiperparámetros con el fin de determinar el modelo con desempeño más elevado en la segmentación del MTB. Además de establecer el algoritmo de cuantificación de la carga tumoral.

### **3.1.5 Evaluación**

En esta fase se evalúan los modelos por medio del conjunto de datos de validación con métricas de segmentación. Posteriormente, se evalúa el modelo seleccionado con nuevas imágenes PET/CT con el MTB delimitado contrastado con el diagnóstico y la delimitación del médico nuclear. Además, se evalúa el rendimiento del algoritmo de cuantificación del MTB.

### **3.1.6 Despliegue**

Construir la arquitectura final del componente de delimitación del MTB que incluya lectura de imágenes, estandarización, delimitación e interfaz de resultados.

## **3.2 Datos**

Los datos PET/CT para este trabajo se recolectaron de forma retrospectiva en un periodo de 6 meses comprendido entre octubre de 2021 hasta marzo del 2022 para



pacientes que se le indica la realización de un examen PET/CT de  $^{68}\text{Ga}$ -PSMA para diagnosticar cáncer de próstata en el Instituto Nacional de Cancerología E.S.E. En total se recolectaron 100 exámenes para una población total de 100 pacientes, todos ellos adultos mayores de 18 años.

La construcción del conjunto de datos se realizó de acuerdo a los estándares de bioética y con el aval del comité de ética del Instituto Nacional de Cancerología E.S.E, en cumplimiento del tratado de Helsinki. El protocolo de la institución establece que se realiza la exploración luego de obtener la firma y aceptación del consentimiento informado (Anexo I). La exploración se realizó utilizando el equipo biomédico Siemens Biograph™ TruePoint PET/CT 40 (Figura 3.2), el cual cuenta con autorización aprobada por el Servicio Geológico Colombiano y la Secretaría Distrital de Salud durante el periodo de los exámenes. El software utilizado para la adquisición fue desarrollado por el fabricante. Además, los exámenes fueron realizados por tecnólogos especializados en medicina nuclear, siguiendo el protocolo institucional descrito en las secciones 3.2.1, 3.2.4 y 3.2.5. Cada toma tuvo una duración máxima aproximada de 30 minutos.

A continuación, se proporciona una descripción detallada de la prescripción de la actividad a administrar y el protocolo de adquisición.



Figura 3.2 Equipo biomédico Siemens Biograph™ TruePoint PET/CT 40.

### **3.2.1 Prescripción de la actividad a administrar**

La actividad del radiofármaco para la exploración es descrita en la tabla 3.1 la cual requiere un tiempo de captación de 50 minutos para la biodistribución del mismo.

Tabla 3.1 Actividad del radiofármaco.

	Radiofármaco	Actividad recomendada administrada (mCi)	Actividad mínima (mCi)	Actividad máxima (mCi)	Tiempo de captación (min)
<b>Adulto</b>	$^{68}\text{Ga}$ -PSMA	3 – 4 40 a 200 kg	3	4	50

### 3.2.2 Exploración anatómica axial

El paciente es acostado en la camilla de exploración del PET/CT en posición decúbito supino con los brazos arriba sin objetos metálicos (Figura 3.3), la exploración se realiza desde el vértice craneal hasta la mitad de los muslos entrando el cráneo. En el caso que el paciente sea tratado posteriormente con radioterapia requiere el uso de una camilla plana.

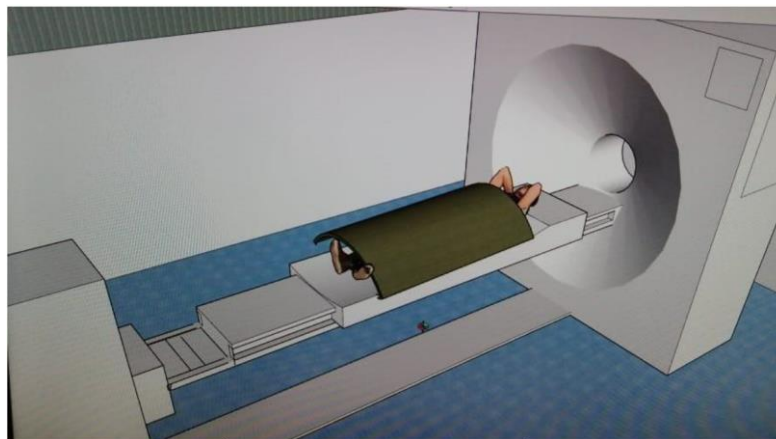


Figura 3.3 Paciente decúbito supino brazos arriba entrando cráneo.

### 3.2.3 Exploración cuerpo completo

Esta exploración es indicada para los pacientes de terapias con  $^{177}\text{Lu}$  y con metástasis óseas conocidas. Donde, el paciente se acostará en la camilla de exploración PET/CT en posición decúbito supino con los brazos abajo sin objetos metálicos (Figura 3.4), la exploración se realiza desde el vértice del cráneo hasta los pies, además entrando pies primero en dirección de exploración caudocraneal.

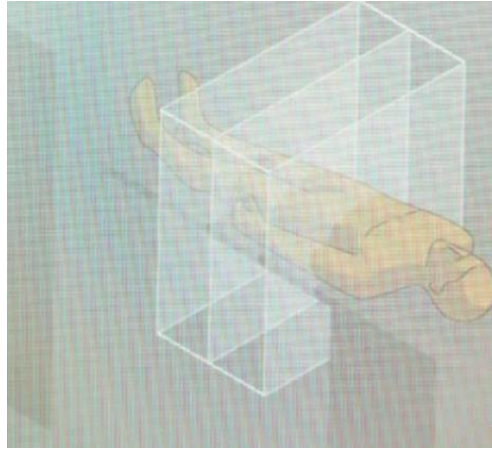


Figura 3.4 Paciente en posición decúbito supino con los brazos abajo entrando pies primero.

### 3.2.4 Protocolo de adquisición y parámetros técnicos

Los protocolos de adquisición y parámetros técnicos siguen los lineamientos establecidos en el SOP para tomografía por emisión de positrones con  $^{68}\text{Ga}$ -PSMA por parte del Instituto Nacional de Cancerología E.S.E. En este SOP, se registran inicialmente los datos del paciente que incluye nombre, número de historia clínica, peso, talla y número de acceso. Además, se distingue entre pacientes adultos y adultos obesos con un peso superior a 80 kg. La dosis inyectada neta, la hora de inyección y el fármaco utilizado se registran, y el tiempo de exploración por cama varía de 3 a 4 minutos dependiendo del peso corporal del paciente (Tabla 3.2).

Tabla 3.2 Duración de la cama en función del peso del paciente.

Peso (kg)	Duración por cama (min)
0 - 69	2
70 - 79	2,5
80 - 100	3
$\geq 100$	4

Inicialmente se adquiere el CT y posteriormente se adquiere la imagen PET bajo los parámetros de la tabla 3.3.

Tabla 3.3 Protocolos de adquisición de exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA.

Protocolo	Tomografía axial computarizada	Reconstrucción CT	Reconstrucción PET
PET WB	kVp: 120.	Espesor de corte: 3 mm.	Iteraciones: 7. Subconjuntos: 14.

Topograma: 1021 – 2048 mm (Según la longitud del paciente).	mA: 100-Con modulación. Espesor de corte: 3 mm. Pitch: 1,5.	Incremento: 2,5 mm. Filtro: B31f Ventana: Abdomen. FOV: 500 mm.	FWHM: 8. Matriz: 168 x 168. Corregido y no corregido por atenuación.
---	---	--	--

### 3.2.5 Procesamiento de las imágenes

Después de completar el estudio se realiza el procesamiento de las imágenes adquiridas las cuales son fusionadas con la condición de grosor de corte de cada imagen sin superar los 5.0 mm. Cada imagen es guardada en reconstrucción axial y almacenada en el sistema para archivo y comunicación de imágenes (Picture Archiving Communication System) (PACS) (Figura 3.5).

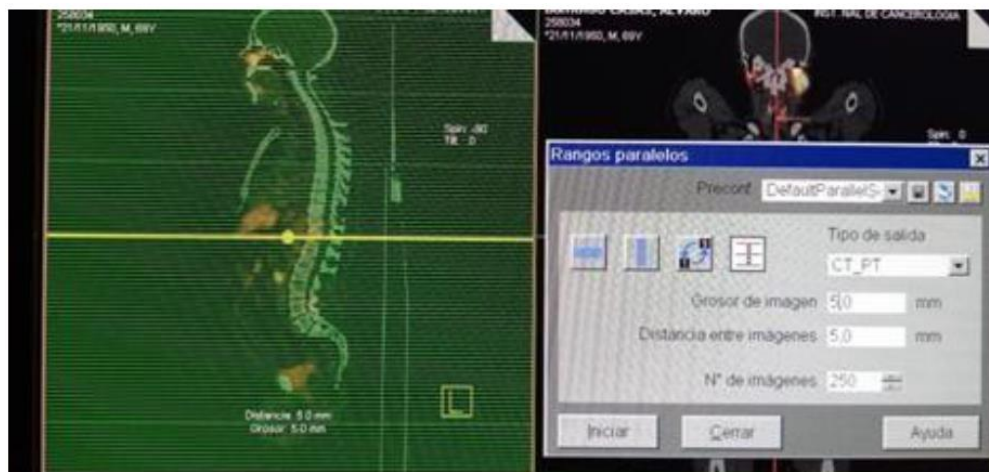


Figura 3.5 Reconstrucción del examen.

### 3.2.6. Recolección de los datos

Los datos se recopilaron manualmente utilizando el sistema PACS, en el cual las imágenes crudas se almacenaron y codificaron en carpetas individuales. Cada carpeta contiene las imágenes CT, PET y la fusión de ambas, en formato DICOM (Digital Imaging and Communication in Medicine). Adicionalmente se consideró el informe del médico nuclear que incluye la lectura e interpretación del examen, el cual se almacena en una carpeta separada.

Todos los exámenes se almacenan en la nube y están bajo la custodia del investigador, con restricciones de acceso para garantizar la seguridad y privacidad de los datos.

### **3.3 Preparación de los datos**

#### **3.3.1 Estandarización**

Cada uno de los exámenes, inicialmente almacenados en formato DICOM, se anonimizó y convirtió al formato TIFF. Esto se realizó para facilitar su procesamiento mediante lenguaje de programación y análisis utilizando el software de anotación 3D SLICER. En este software, el experto llevó a cabo la segmentación de las lesiones.

Además, se identificó que las imágenes fusionadas originalmente con una resolución de 512 x 512 contenían aproximadamente un 50% de píxeles con información nula en las periferias de la imagen. Por lo tanto, se realizó un proceso de corte para reducir su tamaño de ventana a 256 x 256 (ver Figura 3.5). Esta reducción garantiza que la información contenida en el centro de la imagen cubra toda la exploración, incluyendo la información anatómica y funcional captada. El corte se realizó de forma utilizando el software FIJI, que es un paquete de procesamiento de imágenes de código abierto.

#### **3.3.2 Anotación de los datos**

Los exámenes fueron interpretados por el médico nuclear con 5 años de experiencia y con más de 1700 estudios analizados. El médico examinó detalladamente corte por corte cada una de las imágenes y realizó diferentes anotaciones relacionadas con las captaciones anómalas del radiofármaco. Estas anotaciones incluyeron la segmentación y delimitación de las lesiones asociadas al carcinoma de próstata. Además, se identificaron captaciones normales relacionadas con la biodistribución del radiofármaco, sin embargo, estas no fueron anotadas.

Este proceso de anotación se llevó a cabo utilizando el software 3D Slicer, el cual es una herramienta gratuita y de código abierto utilizada para la visualización, procesamiento, segmentación, registro y análisis de imágenes médicas (ver Figura 3.6). Las anotaciones corresponden a máscaras binarias, donde la información de interés se muestra con una tonalidad clara correspondiente a un valor de píxel igual a 255, mientras que el resto de la información se muestra en tonalidad oscura con un valor de píxel igual a 0, indicando que no hay información relevante en esas áreas (ver Figura 3.7).

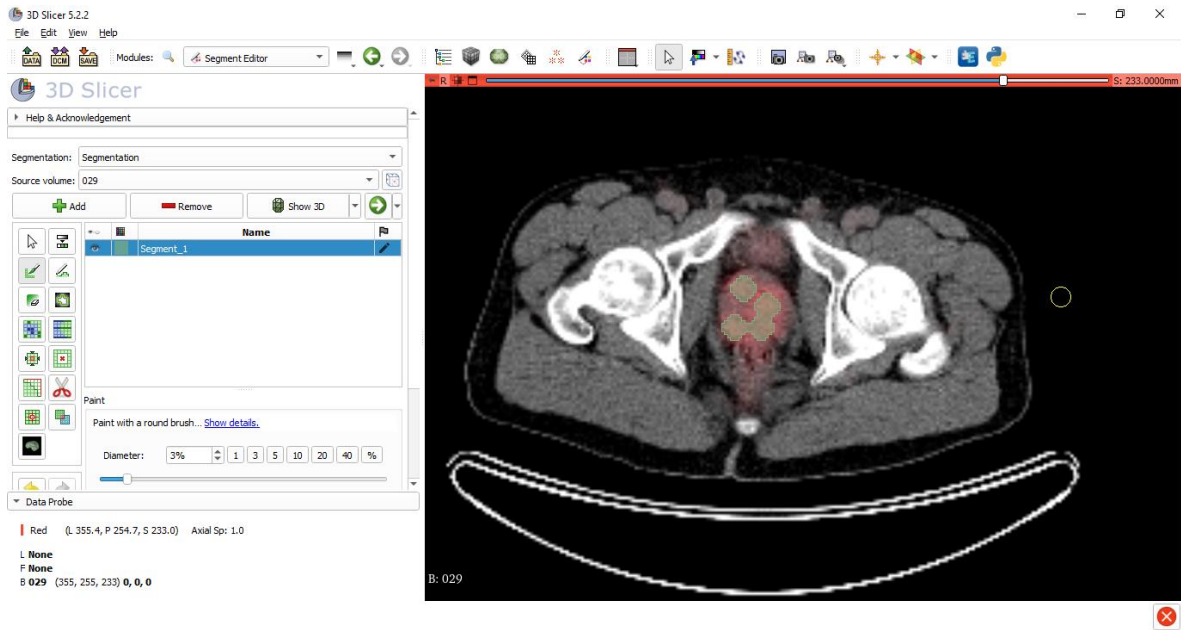


Figura 3.6 Ejemplo del proceso de segmentación de examen PET/CT en 3D Slicer.

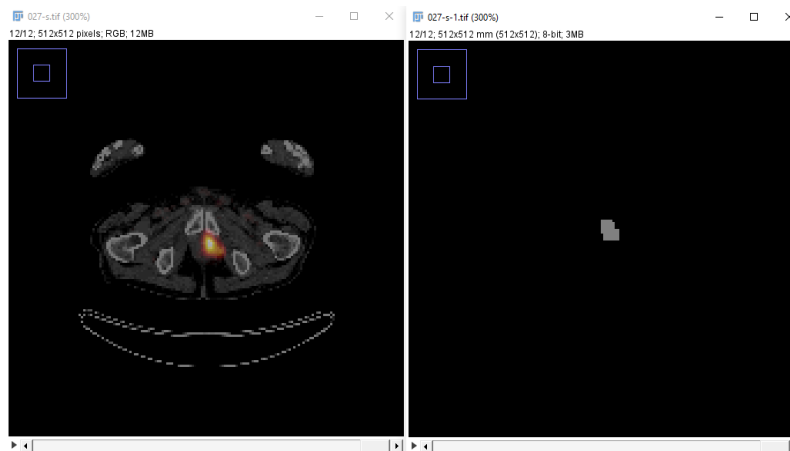


Figura 3.7 Ejemplo de un corte de interés y la correspondiente máscara segmentada por el médico, la cual está asociada a una captación del radiofármaco.

Las máscaras obtenidas corresponden a una serie de imágenes en formato NIFTI (Neuroimaging Informatics Technology Initiative), el cual es un formato estandarizado para Neur-o-imágenes. Esta extensión se utiliza debido a la compatibilidad del software, y posteriormente se convierte al formato TIFF.

### **3.3.2 Binarización**

Las imágenes se convirtieron a escala de grises para poder ser interpretadas, utilizando tonalidades desde 0 hasta 255. Por otro lado, las máscaras se normalizaron y se transformaron a una escala binaria, donde los valores van desde 0 hasta 1, en el módulo de preprocesamiento del algoritmo.

## **3.4 Modelado**

En esta etapa, llevamos a cabo el proceso de modelado del problema de segmentación semántica de la carga metabólica tumoral. Con este propósito, se exploran tres arquitecturas basadas en U-NET, en particular: U-NET, U-NET++ y Attention U-NET. U-NET es una red neuronal convolucional con una arquitectura codificador-decodificador y conexiones de salto, diseñada originalmente para la segmentación semántica de imágenes médicas [53]. Sin embargo, se ha modificado para abordar una amplia gama de problemas de aprendizaje.

A continuación, se realizará una descripción más detallada de las arquitecturas propuestas.

### **3.4.1 U-NET**

La arquitectura de la red se ilustra en la Figura 3.8. Esta arquitectura considera una ruta de contracción también conocida como codificador (lado izquierdo) y una ruta expansiva o decodificador (lado derecho).

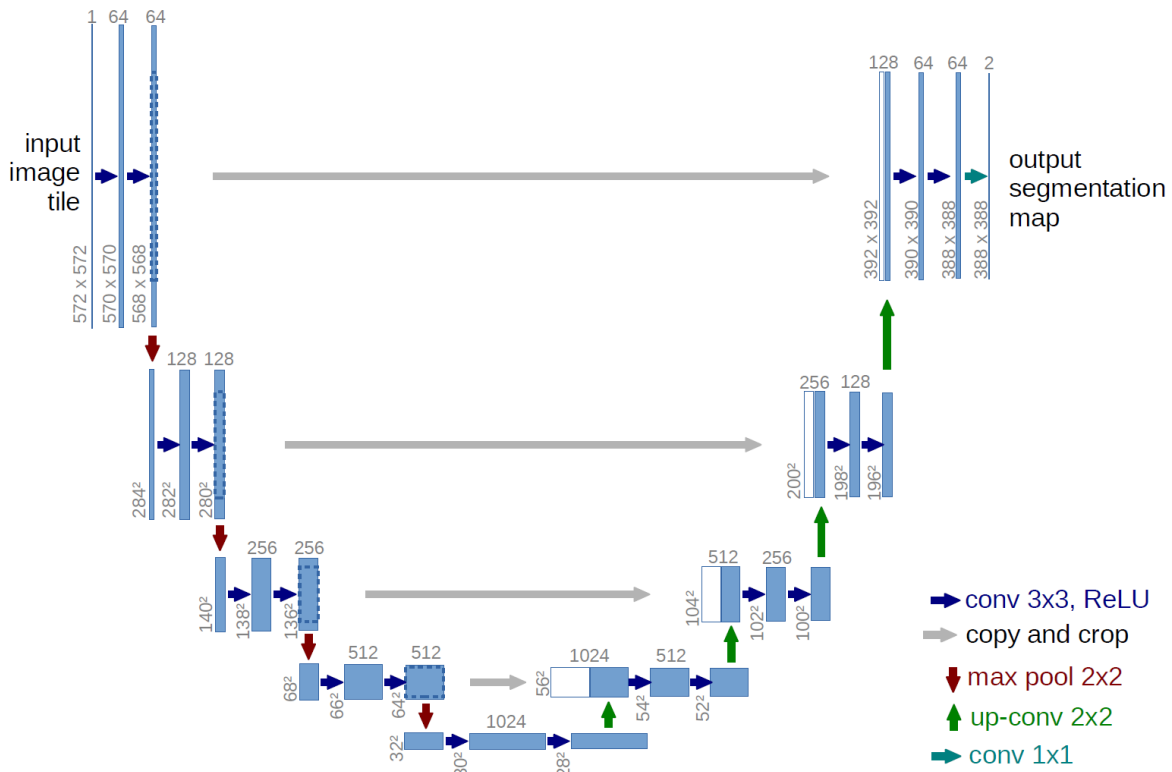


Figura 3.8 Ejemplo de arquitectura U-NET [54].

En la representación gráfica (Figura 3.8), la parte superior muestra el tamaño del filtro utilizado, mientras que en la parte inferior se indica el tamaño de la matriz correspondiente. Las flechas azules denotan operaciones de convolución de  $3 \times 3$  seguidas de una función de activación ReLU (unidad lineal rectificadora). Las flechas grises representan el recorte y copia de la imagen, es decir, el mapa de características. La flecha roja indica una operación de max pooling de  $2 \times 2$ , mientras que la flecha verde representa una operación de deconvolución de  $2 \times 2$ . Por último, la flecha de color agua marino simboliza una operación de convolución de  $1 \times 1$  [54].

La ruta de contracción sigue la arquitectura típica de una red neuronal convolucional, encargada de extraer características de la imagen de entrada. Este proceso se realiza mediante la sucesiva aplicación de capas convolucionales, que gradualmente disminuyen la resolución espacial de la entrada [54].

En cada etapa de la ruta de contracción, se emplean dos capas convolucionales de tamaño  $3 \times 3$ , seguidas de una función de activación ReLU. Posteriormente, se aplica una operación de agrupación máxima (Max pooling) de tamaño  $2 \times 2$  con un desplazamiento de 2, lo que resulta en una reducción de la resolución. Esta operación de agrupación máxima permite seleccionar el valor máximo dentro de una región vecina, ayudando a preservar las características más destacadas de la imagen mientras se reduce la dimensionalidad [54].



Cada paso en la ruta expansiva, que es la parte "ascendente" de la red o decodificador, presenta una estructura simétrica al codificador. Esta etapa se compone de capas convolucionales transpuestas, también conocidas como capas deconvolucionales, cuyo propósito es aumentar gradualmente la resolución espacial de las características [54].

El proceso comienza con un muestreo ascendente del mapa de funciones, seguido de una convolución ascendente de  $2 \times 2$ . Esta convolución reduce a la mitad el número de canales de funciones. Luego, se realiza una concatenación con el mapa de funciones correspondientemente recortado de la ruta de contracción. Posteriormente, se aplican dos convoluciones de  $3 \times 3$ , seguidas de una función de activación ReLU. Es importante realizar el recorte debido a la pérdida de píxeles de borde en cada convolución [54].

En la capa final de esta etapa, se emplea una convolución de  $1 \times 1$  para asignar cada vector de características de 64 componentes al número deseado de clases. Las capas deconvolucionales desempeñan un papel fundamental en la reconstrucción de la información detallada y de alta resolución a partir de las características extraídas en el codificador.

Una de las características distintivas de U-NET radica en sus conexiones de salto o conexiones residuales. Estas conexiones establecen un vínculo directo entre las capas del codificador y las capas del decodificador, lo que permite una transmisión de información más eficiente y preserva los detalles espaciales. Estas conexiones de salto desempeñan un papel fundamental al fusionar las características de niveles de resolución más bajos con las de niveles de resolución más altos, lo que facilita la reconstrucción de las salidas con una alta resolución [54].

### **3.4.2 U-NET++**

UNET++ es una mejora del modelo UNET original, diseñado para tareas de segmentación semántica en imágenes. UNET++ conserva la estructura básica de UNET, que consiste en un codificador y un decodificador conectados por pasos de "skip connection". Sin embargo, UNET++ introduce algunas mejoras significativas en términos de rendimiento y precisión (Figura 3.9) [55].

Una de las principales mejoras de UNET++ es la incorporación de conexiones de "niveles profundos". En UNET, las conexiones de "skip connection" se establecen sólo entre las capas correspondientes del codificador y del decodificador. Además de estas conexiones, se agregan conexiones entre los niveles profundos de la pirámide del codificador y del decodificador. Esto permite que la información de las características se transmita más detalladamente a través de múltiples escalas, mejorando la precisión de la segmentación [55].

Otra mejora clave en UNET++ es la inclusión de "bloques de agrupación de atención". Estos bloques capturan las relaciones de contexto a largo plazo en la imagen al enfocarse en las interacciones entre las características en diferentes niveles de la pirámide. Esto ayuda a capturar características relevantes y mejorar la precisión de la segmentación [55].

Además, UNET++ utiliza técnicas como la normalización de instancias y la activación ReLU modificada (ReLU+) para mejorar la estabilidad y la capacidad de generalización del modelo [55].

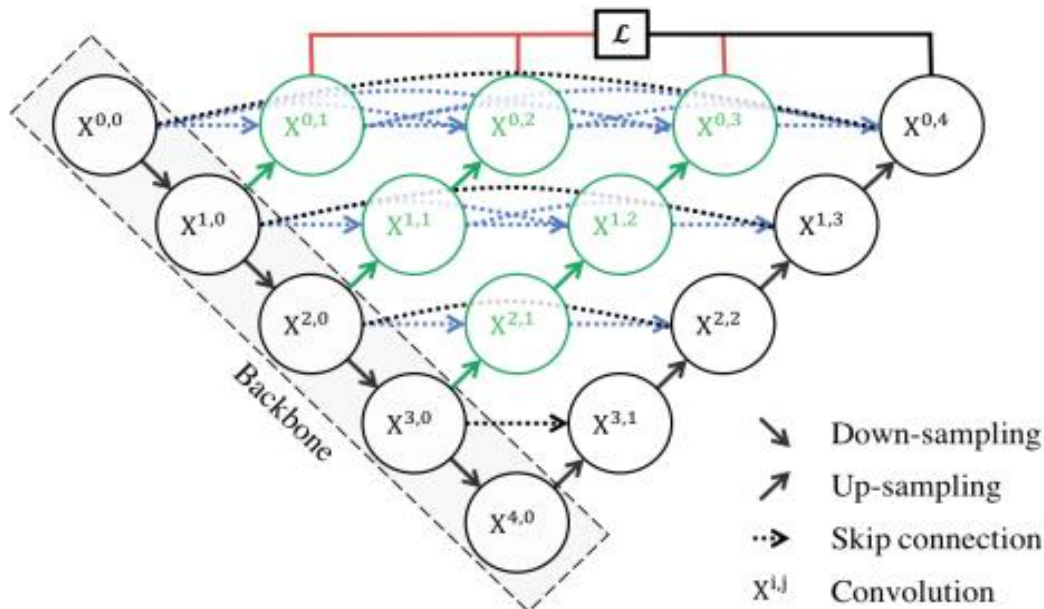


Figura 3.9 Arquitectura de U-Net++ [55].

En la Figura 3.9, el negro indica la U-Net original, el verde y el azul muestran bloques de convolución densos en las rutas de salto y el rojo indica supervisión profunda. Los componentes rojo, verde y azul distinguen a UNet++ de U-Net. La idea principal detrás de UNet++ es cerrar la brecha semántica entre los mapas de funciones del codificador y el decodificador antes de la fusión.

### 3.4.3 Attention U-NET

La arquitectura de Attention UNET es una variante del modelo UNET que incorpora bloques de atención para capturar relaciones de contexto a largo plazo y mejorar la precisión de la segmentación semántica en imágenes [56].

La atención, en el contexto de la segmentación de imágenes, es una forma de resaltar solo las activaciones relevantes durante el entrenamiento. Esto reduce los recursos

computacionales desperdiciados en activaciones irrelevantes, proporcionando a la red un mejor poder de generalización. Esencialmente, la red puede prestar "atención" a ciertas partes de la imagen [56].

Después de cada capa de convolución en el codificador, se agrega un bloque de atención. Este bloque captura las relaciones de contexto a largo plazo al enfocarse en las interacciones entre las características en diferentes niveles de la pirámide. El bloque de atención generalmente se compone de las siguientes etapas:

- **Codificación de características:** Se utiliza una capa convolucional para reducir la dimensionalidad de las características en el nivel actual del codificador. Esto se hace para que las operaciones posteriores sean computacionalmente eficientes.
- **Decodificación de características:** Se utilizan capas convolucionales y de muestreo espacial para aumentar la resolución y el tamaño del campo receptivo de las características codificadas, permitiendo una mejor comprensión del contexto.
- **Atención:** Aquí es donde se capturan las relaciones de contexto. Se realiza una operación de atención, que se basa en la comparación de características de diferentes niveles del codificador y del decodificador. Esto permite que la red se centre en regiones relevantes y resalte características importantes para la segmentación.
- **Fusión de características:** Las características atendidas se fusionan con las características del nivel actual del codificador mediante una operación de concatenación. Esto asegura que la información contextual relevante se transmita a través de la red.

Además de las conexiones de "skip connection", en Attention UNET también se agrega un bloque de atención después de cada capa de convolución en el decodificador, de manera similar a lo que se hace en el codificador. Esto ayuda a capturar relaciones de contexto adicionales y mejorar aún más la precisión de la segmentación [56].

En este caso, se implementó la atención suave aditiva como se observa en la figura.

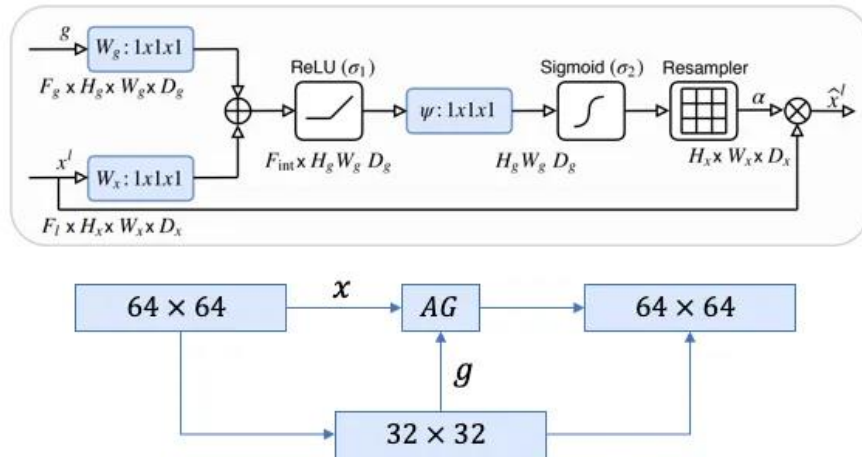


Figura 3.10 Ejemplo de implementación de la atención suave aditiva [57].

Primero, la puerta de atención (AG) en el modelo toma dos entradas: los vectores  $x$  y  $g$ . El vector  $g$  se obtiene de la capa más profunda de la red y tiene dimensiones más pequeñas pero una representación de características mejorada. En el ejemplo dado anteriormente, el vector  $x$  tendría dimensiones de  $64 \times 64 \times 64$  (filtros  $\times$  alto  $\times$  ancho), mientras que el vector  $g$  sería de  $32 \times 32 \times 32$  [57].

A continuación, el vector  $x$  se somete a una convolución escalonada, lo que reduce sus dimensiones a  $64 \times 32 \times 32$ . Por otro lado, el vector  $g$  pasa por una convolución  $1 \times 1$ , lo que lo transforma en un vector de dimensiones  $64 \times 32 \times 32$ . Posteriormente, los dos vectores se suman elemento a elemento. Este proceso resulta en un aumento de los pesos alineados y una disminución relativa de los pesos no alineados [57].

El vector resultante se pasa a través de una capa de activación ReLU y una convolución  $1 \times 1$  que reduce sus dimensiones a  $1 \times 32 \times 32$ . Luego, se aplica una capa sigmoide al vector, que lo escala en el rango de  $[0,1]$ . Estos valores resultantes son los coeficientes de atención (pesos), donde los coeficientes más cercanos a 1 indican características más relevantes [57].

Finalmente, los coeficientes de atención se redimensionan mediante interpolación trilineal para que coincidan con las dimensiones originales del vector  $x$  ( $64 \times 64$ ). Luego, se multiplican elemento a elemento con el vector  $x$  original, lo que escala el vector según la relevancia de las características. Este resultado se transmite a lo largo de la conexión condicionada como de costumbre [57].

### 3.4.4 Columna vertebral de las arquitecturas

Gracias a la transferencia de aprendizaje, se puede construir una arquitectura para la detección y segmentación de objetos sobre otra CNN que fue estrenada originalmente para la detección de objetos. En este caso, se usa la CNN como extractor de características, y es, de hecho, la columna vertebral del modelo para la segmentación de las lesiones (Figura 3.11) [58].

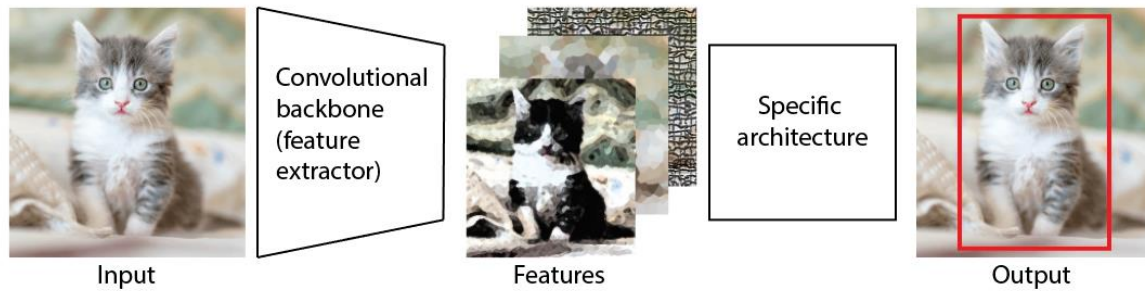


Figura 3.11 Ejemplo de la implementación de una columna vertebral [58].

En general, el término "columna vertebral" se utiliza para describir la red neuronal encargada de extraer características a partir de los datos de entrada y transformarlos en una representación de características específicas. Estas redes de extracción de características suelen ser efectivas cuando se utilizan de forma independiente en tareas más simples, y se integran como una parte fundamental en la extracción de características de modelos más complejos [58].

Existen numerosas arquitecturas populares de CNN que se utilizan como columna vertebral en redes neuronales. No obstante, en este trabajo se seleccionó una arquitectura específica basándose en la compatibilidad y disponibilidad de la librería "keras-unet-collection" [53], utilizada en el proyecto. Además, se consideró su precisión en la tarea de detección de objetos, evaluada mediante el rendimiento del modelo en el conjunto de datos de validación de ImageNet (Anexo I).

La elección de la columna vertebral recayó en DenseNet201 debido a su alta exactitud. Es importante destacar que EfficientNetB7 muestra valores aún más altos, sin embargo, se encontraron problemas de incompatibilidad durante el proceso de guardado, lo que llevó a descartar su utilización (ver Anexo I).

### 3.4.5 Parámetros utilizados en las arquitecturas

Los parámetros implementados para cada arquitectura son los siguientes:

- **U-NET:**

Tamaño de entrada: (256, 256, 256).

Número de filtros: [64, 128, 256, 512, 1024].

Número de clases:1.

Número de capas convolucionales por bloque de reducción de resolución: 2.

Número de capas convolucionales (después de la concatenación) por bloque: 2.

Función de activación: ReLU.

Función de activación de salida: Sigmoide.

Normalización por lotes: Verdadera.

Capa de agrupación (pool): Máxima.

Unpooling: Falso.

Columna vertebral: DenseNet201.

Pesos: ImageNet.

Congelar columna vertebral: Verdadero.

Congelar normalización por lotes: Verdadero.

- **U-Net++:**

Tamaño de entrada: (256, 256, 256).

Número de filtros: [64, 128, 256, 512, 1024].

Número de clases:1.

Número de capas convolucionales por bloque de reducción de resolución: 2.

Número de capas convolucionales (después de la concatenación) por bloque: 2.

Función de activación: ReLU.

Función de activación de salida: Sigmoide.

Normalización por lotes: Verdadera.

Capa de agrupación (pool): Máxima.

Unpooling: Falso.

Columna vertebral: DenseNet201.

Pesos: ImageNet.

Congelar columna vertebral: Verdadero.

Congelar normalización por lotes: Verdadero.

- **Attention U-NET:**

Tamaño de entrada: (256, 256, 256).

Número de filtros: [64, 128, 256, 512, 1024].

Número de clases:1.  
Número de capas convolucionales por bloque de reducción de resolución: 2.  
Número de capas convolucionales (después de la concatenación) por bloque: 2.  
Función de activación: ReLU.  
Función de atención de activación: ReLU.  
Tipo de atención: Aditiva.  
Función de activación de salida: Sigmoide.  
Normalización por lotes: Verdadera.  
Capa de agrupación (pool): Máxima.  
Unpooling: Falso.  
Columna vertebral: DenseNet201.  
Pesos: ImageNet.  
Congelar columna vertebral: Verdadero.  
Congelar normalización por lotes: Verdadero.

### **3.5 Evaluación**

Durante esta etapa, se llevan a cabo evaluaciones para seleccionar el modelo con el mejor rendimiento en la cuantificación automática de la carga metabólica tumoral. Para esta tarea se consideraron varios aspectos importantes, incluyendo, la selección de la arquitectura óptima de la red neuronal y su capacidad de generalización.

A continuación, se describe el método implementado para evaluar estos aspectos:

#### **3.5.1 Validación cruzada anidada en k particiones**

La validación cruzada es un procedimiento utilizado para estimar el rendimiento de los modelos de aprendizaje automático al realizar predicciones en datos que no se utilizaron durante el entrenamiento [59].

Este enfoque se puede emplear tanto para optimizar los hiperparámetros de un modelo en un conjunto de datos, como para comparar y seleccionar un modelo adecuado para dicho conjunto. Sin embargo, al utilizar el mismo procedimiento de validación cruzada y el mismo conjunto de datos para ajustar y seleccionar un modelo, es posible que se produzca una evaluación optimista y sesgada del rendimiento del modelo [60].

Para superar este sesgo, se puede adoptar un enfoque conocido como validación cruzada doble o validación cruzada anidada. En este caso, se realiza una anidación del procedimiento de optimización de selección del modelo. Esta estrategia se considera la más adecuada para evaluar y comparar modelos de aprendizaje automático (Figura 3.12).

El proceso de validación cruzada con  $k$  particiones consta de varios pasos clave para determinar la mejor arquitectura y la capacidad de generalización de un modelo seleccionado. Estos pasos se describen a continuación [59]:

1. **Partición de los datos:** En primer lugar, se realiza una partición de los datos en tres conjuntos: entrenamiento, validación y prueba. Esta partición se realiza con el objetivo de utilizar cada conjunto para una tarea específica. El conjunto de entrenamiento se utiliza para ajustar el modelo, el conjunto de validación se utiliza para evaluar el rendimiento y seleccionar la arquitectura más adecuada, y finalmente, el conjunto de prueba se utiliza para evaluar la capacidad de generalización del modelo.

Recordar que los datos de las pruebas se utilizan exclusivamente para evaluar el desempeño al final del proceso y nunca se utilizarán en el proceso de aprendizaje. Los datos de validación se utilizan explícitamente para seleccionar los parámetros/modelos con el mejor rendimiento según una estimación del error de generalización. Finalmente, los datos de entrenamiento se utilizan para aprender la instancia del modelo de una clase de modelo.

2. **División en  $k$  particiones externas:** El conjunto de datos externo se divide en  $k$  particiones, generando  $k$  conjuntos de entrenamiento y prueba.
3. **División de  $k$  particiones internas:** El conjunto de entrenamiento externo se divide en  $k$  particiones internas adicionales, creando un conjunto de entrenamiento y validación interna en cada una de ellas.
4. **Ajuste del modelo:** Una vez que se ha realizado la partición de los datos, se procede a ajustar el modelo utilizando el conjunto de entrenamiento. Durante este proceso, el modelo se entrena utilizando los datos disponibles y se ajustan los parámetros del modelo para minimizar el error en las predicciones.
5. **Selección de la arquitectura:** Se selecciona el modelo con la mejor arquitectura en base a los valores promedio de las diferentes métricas de evaluación obtenidas a lo largo del conjunto de datos.
6. **Evaluación de la capacidad de generalización:** Ahora, cada modelo es entrenado con el conjunto de datos de entrenamiento externo y se utiliza el conjunto de prueba para evaluar la capacidad de generalización del modelo seleccionado. Este conjunto contiene datos que el modelo no ha visto durante el entrenamiento ni la selección de la arquitectura. Se utilizan las métricas de evaluación para medir el rendimiento del modelo en este conjunto y determinar cómo se generaliza a nuevos datos.



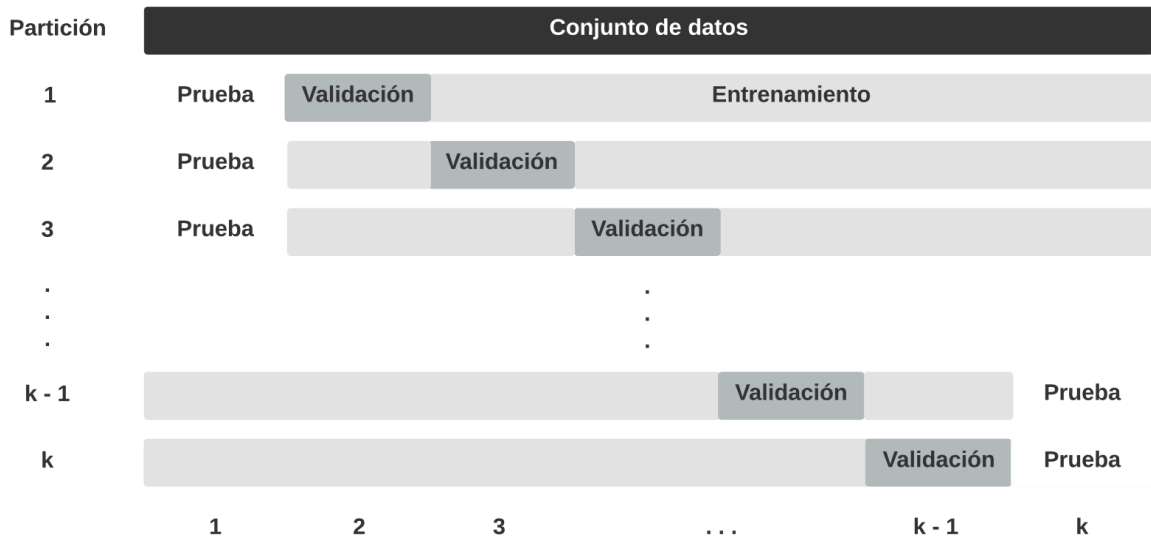


Figura 3.12 Validación cruzada anidada en k particiones.

El proceso de validación cruzada anidada en k particiones se lleva a cabo como paso inicial para seleccionar la arquitectura más adecuada entre U-NET, U-Net++ y Attention U-NET, así como para evaluar su capacidad de generalización, como se muestra en la Figura 3.13. Cada bucle de este proceso se diseña para abordar una de estas incógnitas.

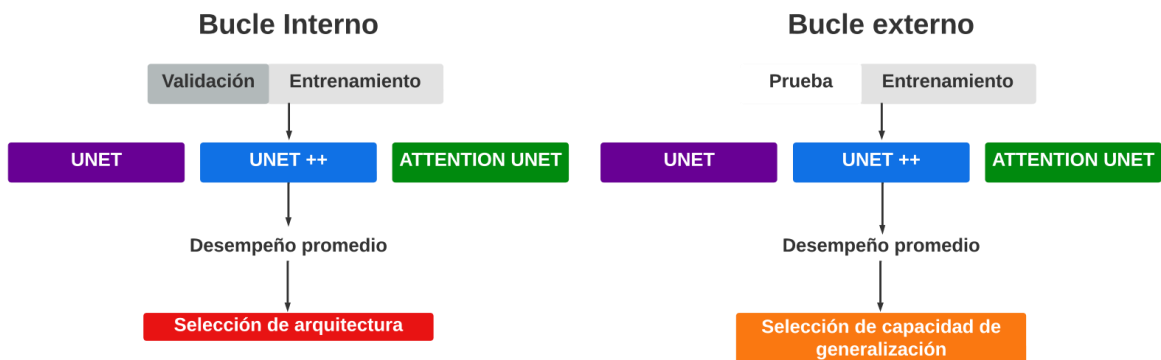


Figura 3.13 Bucles del proceso de evaluación.

### 3.5.2 Índices

A continuación, se detallan los índices utilizados en el proceso de evaluación. Es importante destacar que, en el proceso de segmentación semántica, los píxeles se clasifican mediante aprendizaje supervisado, lo que nos permite obtener una matriz de confusión que

muestra el rendimiento de los algoritmos. En esta matriz, cada columna representa el número de predicciones de cada clase, mientras que cada fila representa una instancia de clase verdadera (Tabla 3.4).

Tabla 3.4 Matriz de confusión.

		Clase Predicha		
		Positivo	Negativo	
Clase Real	Positivo	TP (Verdadero Positivo)	FN (Falso Negativo)	$N^+ = TP + FN$
	Negativo	FP (Falso Positivo)	TN (Verdadero Negativo)	$N^- = FP + TN$
		$\hat{N}^+ = TP + FP$	$\hat{N}^- = FN + TN$	

Donde el verdadero positivo (TP) representa el número total de píxeles asociados a la captación del radiofármaco detectados correctamente y verdadero negativo (TN) representa el número total de píxeles donde no existe captación asociada al carcinoma de próstata detectados correctamente. Falso positivo (FP) y Falso negativo (FN) representa el número total de errores positivos y errores negativos.

### 3.5.2.1 Exactitud (Accuracy)

Muestra qué tan bien el algoritmo logra clasificar correctamente todos los píxeles en la imagen, independientemente de la clase a la que pertenezcan. Una puntuación alta de exactitud indica que el algoritmo está realizando predicciones precisas y consistentes en todas las clases presentes. Su expresión es:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \quad (3.1)$$

### 3.5.2.2 Sensibilidad (Recall)

También conocido como recall o tasa de verdaderos positivos, es una métrica que evalúa la proporción de píxeles positivos correctamente identificados (verdaderos

positivos) en relación con el total de píxeles que realmente pertenecen a la clase objetivo (verdaderos positivos más falsos negativos).

El recall en la segmentación semántica indica qué tan bien el algoritmo puede encontrar y recuperar todos los píxeles relevantes de una clase específica en una imagen. Una puntuación alta de recall implica que hay menos falsos negativos, es decir, el algoritmo está identificando la mayoría de los píxeles de la clase objetivo.

$$Se = \frac{TP}{TP + FN} \quad (3.2)$$

### 3.5.2.3 Precisión (Precision)

Muestra qué tan preciso es el algoritmo al identificar los píxeles que pertenecen a las clases que estamos interesados en segmentar. Una puntuación alta de precisión indica que hay una menor cantidad de falsos positivos, es decir, menos píxeles incorrectamente clasificados como pertenecientes a una clase específica. Su expresión es:

$$Pr = \frac{TP}{TP + FP} \quad (3.3)$$

### 3.5.2.4 Área bajo la curva ROC (AUC)

En la segmentación semántica, el Área bajo la Curva (AUC, por sus siglas en inglés) se refiere a una métrica utilizada para evaluar el rendimiento global de un modelo de segmentación en términos de su capacidad para realizar predicciones precisas en varias clases.

El AUC se calcula mediante la representación gráfica de la curva ROC (Receiver Operating Characteristic). La curva ROC muestra la tasa de verdaderos positivos (sensibilidad) en función de la tasa de falsos positivos (1 - especificidad) a través de diferentes umbrales de clasificación. La especificidad se define así:

$$Especificidad = \frac{TN}{TN + FP} \quad (3.4)$$

La AUC mide el área bajo la curva ROC, y se interpreta como la capacidad del modelo para distinguir entre clases positivas y negativas. Cuanto más cerca esté el valor del AUC de 1, mejor será el rendimiento del modelo en la segmentación semántica. Se obtiene así:

$$AUC = \int_0^1 Pr[TP](v)dv \quad (3.5)$$

Donde la sensibilidad  $Pr[TP]$  es función de la tasa de falsos positivos  $v = Pr[FP]$ .

### 3.5.2.5 Coeficiente de Sørensen-Dice

El coeficiente ó índice de Sørensen-Dice, también conocido por otros nombres tales como el índice de Sørensen, coeficiente de Dice, es un estadístico utilizado para comparar la similitud de dos muestras. Fue desarrollado independientemente por los botánicos Thorvald Sørensen y Lee Raymond Dice, que publicaron en 1948 y 1945 respectivamente. Es una medida ampliamente utilizada para evaluar algoritmos de segmentación de imágenes médicas y ofrece una medida estandarizada de la precisión de la segmentación [61].

En la segmentación semántica, es común que haya una clase de interés que ocupe solo una pequeña proporción del área total de la imagen. Por ejemplo, en la segmentación de tumores en imágenes médicas, el tumor puede representar sólo una fracción pequeña del área total de la imagen. Esto crea un desequilibrio en la distribución de las clases, donde la clase de interés está subrepresentada en comparación con otras regiones o clases en la imagen [61].

El índice de Dice aborda esta desventaja de los índices desbalanceados al considerar tanto los verdaderos positivos como los falsos positivos y los falsos negativos en la evaluación de la segmentación. Al tener en cuenta tanto los aciertos como los errores en la clasificación, el índice de Dice proporciona una medida más equilibrada y sensible a la similitud entre la segmentación obtenida y la segmentación de referencia.

La fórmula original de Sørensen estaba destinada a ser aplicada a datos de presencia/ausencia, y se define de la siguiente forma:

$$QS = \frac{2C}{A + B} = \frac{2|A \cap B|}{|A| + |B|} \quad (3.6)$$

A y B son el número de elementos en las muestras A y B, respectivamente, y C es el número de elementos compartidas por las dos muestras; QS es el coeficiente de similitud y varía de 0 a 1. El índice de Sørensen-Dice es igual al doble del número de elementos comunes a ambos conjuntos dividido por la suma del número de elementos de cada conjunto. Este concepto se puede aplicar a tareas de segmentación semántica como se ilustra en la Figura 3.14.

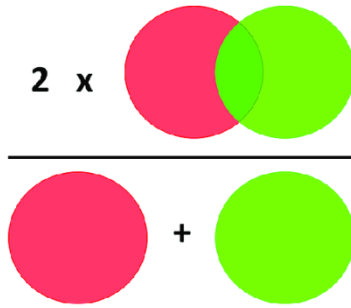


Figura 3.14 Coeficiente Sørensen-Dice [62].

Se toma los píxeles de la intersección presente en ambas imágenes que se está comparando, en este caso de la segmentación original hecha por el médico nuclear y la predicha por el modelo y se multiplica por 2. Luego se divide por el total de píxeles en ambas segmentaciones.

### 3.5.2.6 Coeficiente de Jaccard

El coeficiente de Jaccard, también conocido como intersección sobre unión (IoU), es una medida estándar de rendimiento utilizada en el problema de segmentación. Su objetivo es evaluar la similitud entre la región predicha y la región real de un objeto en un conjunto de imágenes. Esta medida describe el nivel de superposición entre las segmentaciones, es decir, entre la segmentación de la predicción y la segmentación real. Cuanto mayor sea el grado de superposición, mayor será el valor de IoU [63]. Su cálculo se realiza mediante la siguiente expresión:

$$IoU = \frac{TP}{TP + FP + FN}, \quad (3.7)$$

donde TP, FP y FN representan la cantidad de verdaderos positivos, falsos positivos y falsos negativos, respectivamente.

Además, se puede utilizar la función de pérdida para determinar el grado de error cometido. Esta función se expresa de la siguiente manera [63]:

$$Los IoU = 1 - \frac{TP}{TP + FP + FN} \quad (3.8)$$

De esta forma, cuanto menor sea el valor de IoU, menor será la pérdida y, por lo tanto, menor será el error en la segmentación.

### 3.6 Despliegue

El propósito de la fase de despliegue consiste en presentar el modelo final de una forma adecuada para que los usuarios finales puedan utilizarlo. Para lograr esto, se planteó la arquitectura del denominado “Componente inteligente” (Figura 3.15), el cual se encarga de analizar los exámenes PET/CT para determinar y cuantificar la carga metabólica tumoral asociada al carcinoma de próstata. Esta carga es expresada en mililitros y es desplegada en pantalla, además de visualizarse la segmentación volumétrica predicha por el componente inteligente sobre cada corte axial del examen, para que finalmente sea valorado por el médico nuclear.

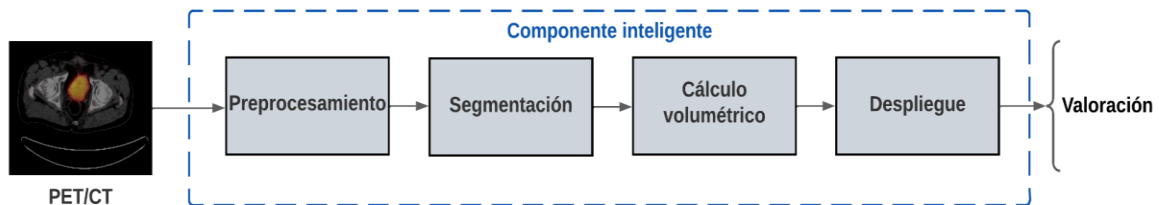


Figura 3.15 Arquitectura del componente inteligente.

## Capítulo 4.

### Resultados y discusión

#### 4.1 Conjunto de datos

Se recolectó un total de 100 exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA de acuerdo con el protocolo descrito en la sección 3.2. La interpretación y anotación de las lesiones (proceso de segmentación) se llevaron a cabo bajo la supervisión del médico nuclear. A continuación, se describen las características de los pacientes incluidos en el conjunto de datos (ver Tabla 4.1).

Tabla 4.1 Demografía de los pacientes sometidos al examen PET/CT.

No Pacientes	Edad (años)	Peso (kg)	Actividad administrada (mCi)
100	44 - 90	50 - 120	3 - 4

Se llevó a cabo una exploración de los datos, en la cual se identificaron un total de 33 pacientes con lesiones relacionadas con el carcinoma de próstata. Los demás exámenes fueron descartados, ya que los pacientes no mostraron captaciones del radiofármaco asociadas al carcinoma de próstata. Por esta razón el 33% de los exámenes se utilizaron para la etapa de modelado.

Los exámenes utilizados consisten de 8.210 imágenes de cortes axiales. De estas 311 contienen las lesiones que fueron utilizadas para alimentar el algoritmo.

Una vez determinados los cortes con lesiones, se realizó un análisis de la ubicación anatómica de las lesiones, dividiéndolas en tórax, abdomen y pelvis. No se encontraron lesiones en el resto de las regiones del cuerpo. Notablemente, el 87% de las imágenes de estas lesiones se localizaron en la pelvis, en concordancia con la distribución habitual del carcinoma de próstata. Los resultados específicos para cada una de estas regiones se detallan en la Figura 4.1.

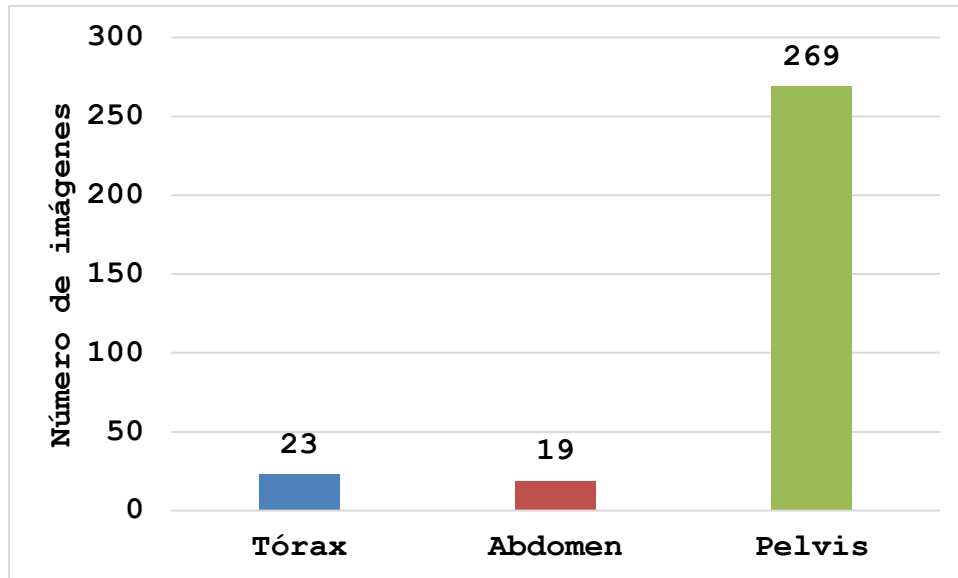


Figura 4.1 Distribución de imágenes por región anatómica.

Tras examinar la Figura 4.1, se llegó a la conclusión de que únicamente se dispone de una cantidad significativa de imágenes para abordar la cuantificación del MTB en la región anatómica de la pelvis.

Además, se realizó una exploración de los órganos en los cuales se distribuyen estas lesiones en la región de la pelvis. Se encontró que la afectación para este conjunto de datos se concentra principalmente en la próstata (117) y en diferentes ganglios linfáticos (138). La Figura 4.2 ilustra un ejemplo de las lesiones en la región de la pelvis.

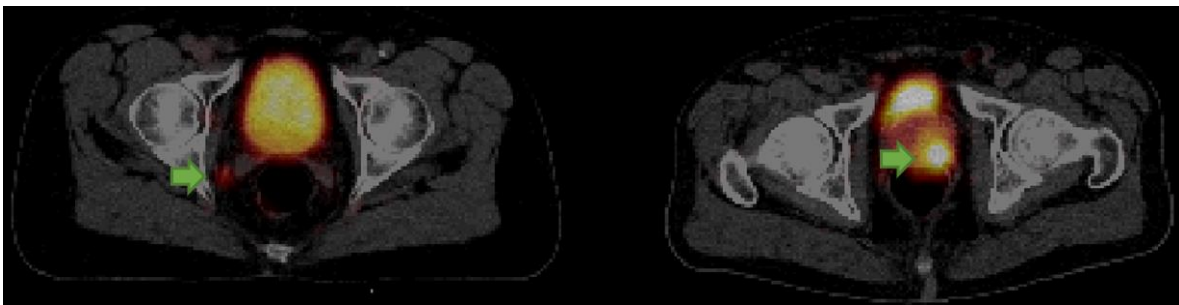


Figura 4.2 Ejemplo de lesiones en la región de la pelvis. Las flechas verdes indican las lesiones en el corte axial, a la izquierda en un ganglio y a la derecha en la próstata.

Como se mencionó anteriormente, el conjunto de datos colectado no representa todas las lesiones que pueden presentarse en diversas regiones anatómicas. Por lo tanto,



este trabajo, se focalizó únicamente en las lesiones de la próstata y los ganglios linfáticos de la zona de la pelvis.

## 4.2 Modelado

Según la metodología propuesta, la etapa de modelado se llevó a cabo utilizando la validación cruzada anidada en  $k$  particiones. Inicialmente, se determinó y seleccionó la arquitectura entre U-NET, U-NET++, y attention U-NET, basándose en el mejor rendimiento según los diversos índices propuestos en la sección 3.5.2. Posteriormente, se evaluó la capacidad de generalización de la arquitectura elegida y, finalmente, el modelo se entrenó utilizando todo el conjunto de datos, lo que se implementó en la etapa de despliegue. A continuación, se presentan los resultados para cada una de estas etapas.

### 4.2.1 Segmentación del examen PET/CT de $^{68}\text{Ga}$ -PSMA

La Figura 4.3 ilustra el rendimiento de forma cualitativa del modelo en la segmentación de lesiones en la región de la próstata. La imagen A muestra un corte axial de un paciente con captaciones fisiológicas en la vejiga y captaciones asociadas al carcinoma en la próstata. En la imagen B, se observa la segmentación realizada por el modelo, con un coeficiente de Dice de 0,95, lo que indica una delimitación casi perfecta. La región de color verde corresponde a los píxeles verdaderos positivos, mientras que los píxeles en rojo representan los falsos positivos detectados por el modelo. Como puede observarse, el modelo también es capaz de diferenciar las captaciones fisiológicas no relacionadas con el carcinoma, como es el caso de la vejiga.

En la imagen C, se presenta un caso de captación leve y de baja intensidad en la próstata. En la imagen D, se muestra la segmentación realizada por el modelo con un coeficiente de Dice de 0,63, lo que indica una delimitación con algunos errores e ilustra algunas de las dificultades del método para delinear completamente la región. Estos desafíos se deben principalmente a la baja captación del radiofármaco, y es uno de los ejemplos en los que el modelo enfrenta mayores dificultades. El color amarillo representa la región de los falsos negativos.

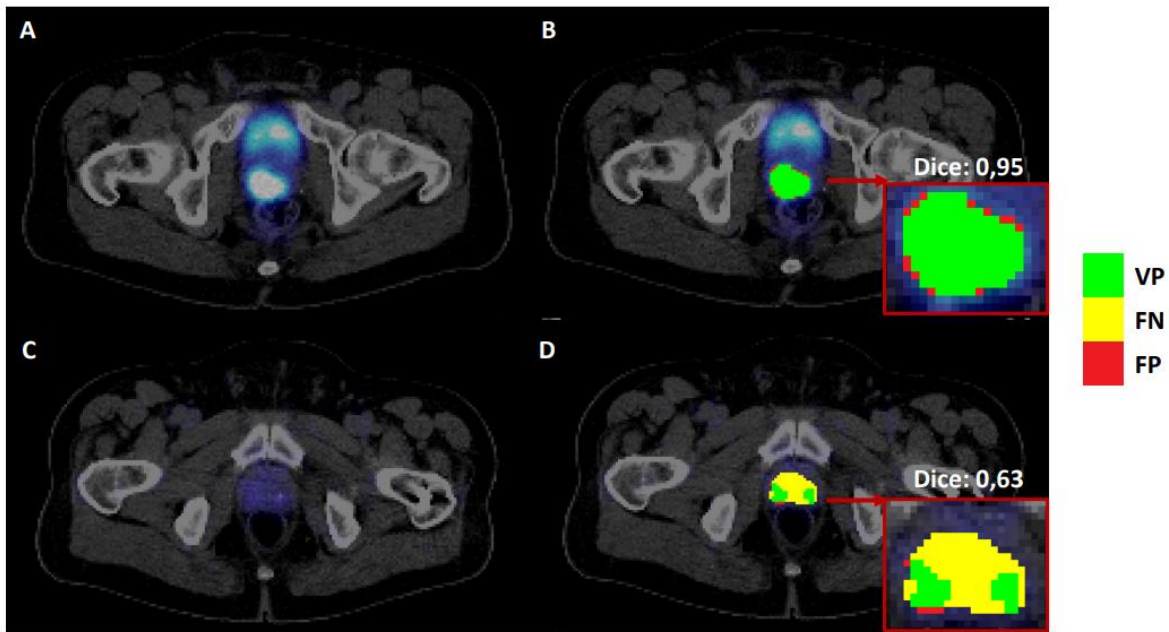


Figura 4.3 Comparación cualitativa de la segmentación en la próstata.

La Figura 4.4 ilustra los resultados obtenidos por el modelo propuesto en la segmentación, se exploró el rendimiento de forma cualitativa en la segmentación de lesiones en los ganglios linfáticos de la pelvis. En la imagen A, se muestra un corte axial de un paciente con captaciones fisiológicas en la vejiga y captaciones asociadas al carcinoma en un ganglio linfático. La imagen B muestra la segmentación realizada por el modelo, con un coeficiente de Dice de 0,65, indicando una delimitación moderada de la lesión. Es importante notar que el modelo presenta limitaciones en este contexto para diferenciar las captaciones de baja intensidad, como es común en los ganglios linfáticos.

En la imagen C, se presenta un caso de captación en un ganglio linfático en la región de la pelvis. En la imagen D, se muestra la segmentación realizada por el modelo con un coeficiente de Dice de 0,64, lo que indica una delimitación con ciertos errores y dificultades para completar la región. Además, se pueden observar varios falsos positivos asociados a las captaciones fisiológicas de la vejiga, lo que representa un desafío adicional para el modelo en este contexto. En la imagen también se visualiza en color amarillo la región que representa los falsos negativos, aunque la cantidad de píxeles no es tan pronunciada en comparación con la región de falsos positivos.

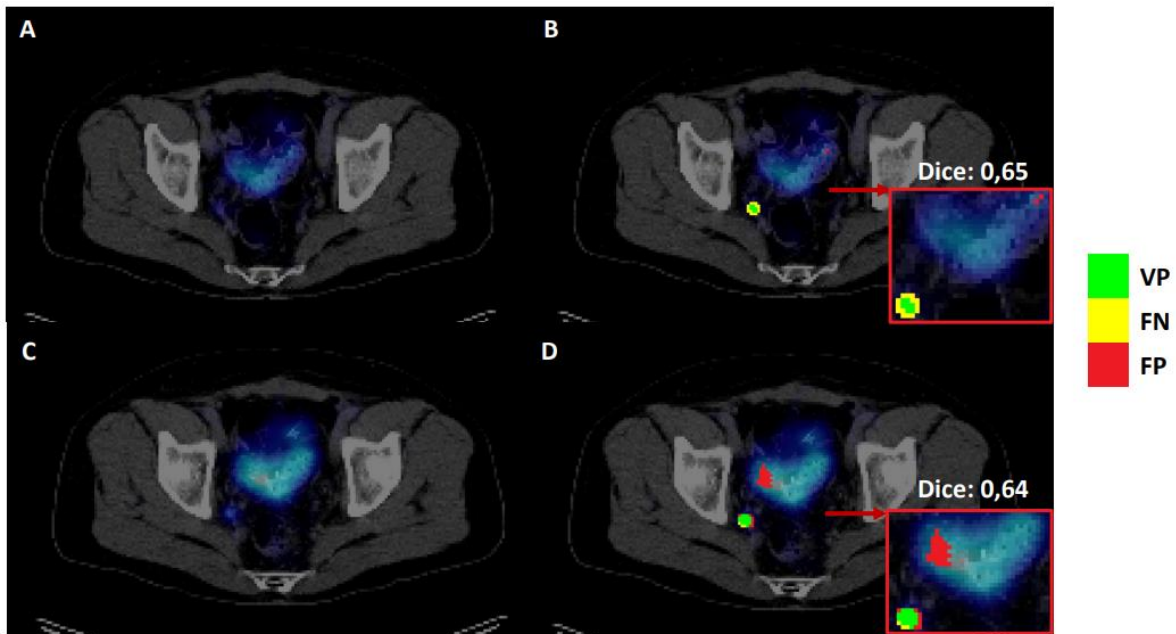


Figura 4.4 Comparación cualitativa de la segmentación en los ganglios de la pelvis.

#### 4.2.2 Segmentación del examen PET/CT $^{68}\text{Ga}$ -PSMA a nivel de grupo

El modelo seleccionado obtuvo un coeficiente de Dice promedio de 0,99 en el conjunto de entrenamiento y 0,53 en el conjunto de prueba, del cual se procedió a inspeccionar su comportamiento para realizar una comparación cualitativa frente a cada una de las imágenes implementados en la prueba en diferentes contextos como se observa en la sección 4.2.1.

Tras el análisis cualitativo, se observó una clara diferencia en el desempeño del modelo frente a los contextos anatómicos de la próstata y los ganglios linfáticos. Para una evaluación más cuantitativa, se analizó el comportamiento del modelo a lo largo de los diferentes contextos anatómicos utilizando los coeficientes de Dice para el conjunto de prueba, representados en un diagrama de violín (Figura 4.5). El diagrama de violín representa la densidad de probabilidad de los datos, donde cada punto representa el coeficiente de Dice obtenido por la segmentación en una imagen anatómica axial.

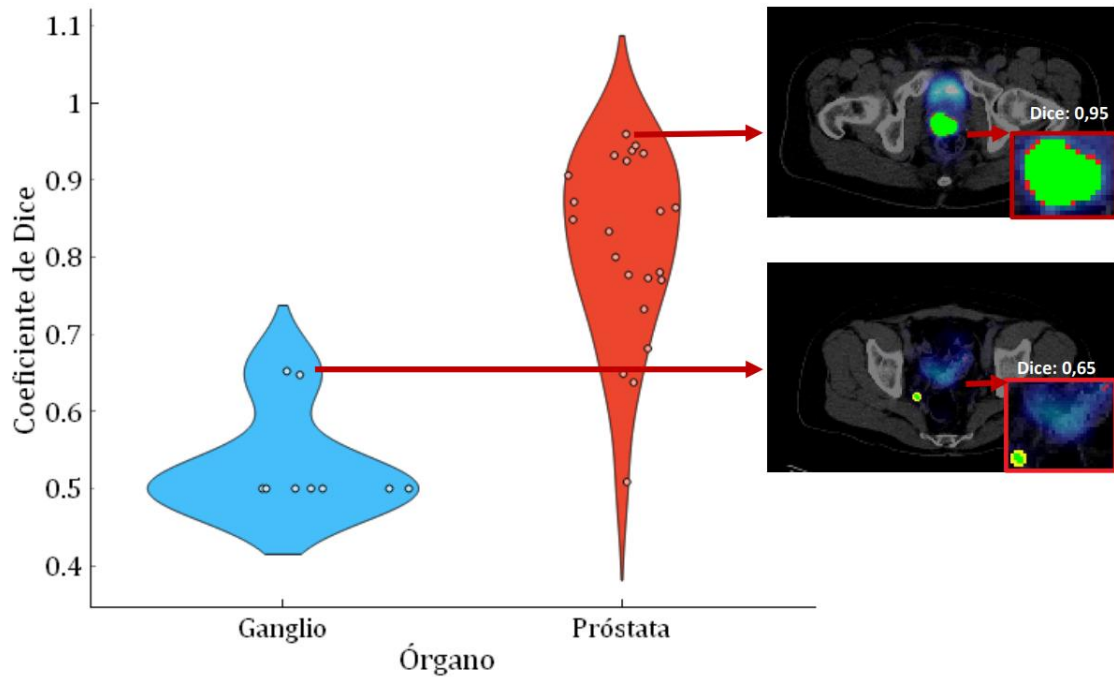


Figura 4.5 Diagrama de violín del conjunto de datos de prueba.

El diagrama de violín revela una distribución multimodal, evidenciada por la presencia de múltiples picos con diferentes alturas y posiciones, como se ilustra en la Figura 4.5. Se identifican dos modas distintas: los coeficientes de Dice extraídos de las segmentaciones de lesiones asociadas a los ganglios linfáticos de la pelvis exhiben valores inferiores en comparación con las lesiones de la próstata. Esta diferencia es lo suficientemente significativa como para dividir las en dos categorías distintas. Estos hallazgos sugieren la necesidad de abordar el problema en dos contextos separados. En consecuencia, se recomienda la construcción de modelos independientes para abordar, por un lado, la segmentación de la próstata y, por otro, la segmentación de los ganglios linfáticos, considerando las disparidades en las características de las lesiones detectadas en cada región anatómica.

También, se evaluó el rendimiento del modelo por medio de la curva ROC que representa la relación entre la tasa de verdaderos positivos (Sensibilidad) y la tasa de falsos positivos (1 - Especificidad). El punto óptimo del modelo suele ser aquel en el que se maximiza la sensibilidad y, al mismo tiempo, se minimiza la tasa de falsos positivos. Para ello, se trazó la curva en cada una de las imágenes evaluadas y se estimó el valor promedio para el modelo completo. Además, se calculó el área bajo la curva AUC, obteniendo un valor promedio de 0,7, siendo superior al 0,5 lo que indica que el modelo es mejor que una clasificación aleatoria (ver Figura 4.6).

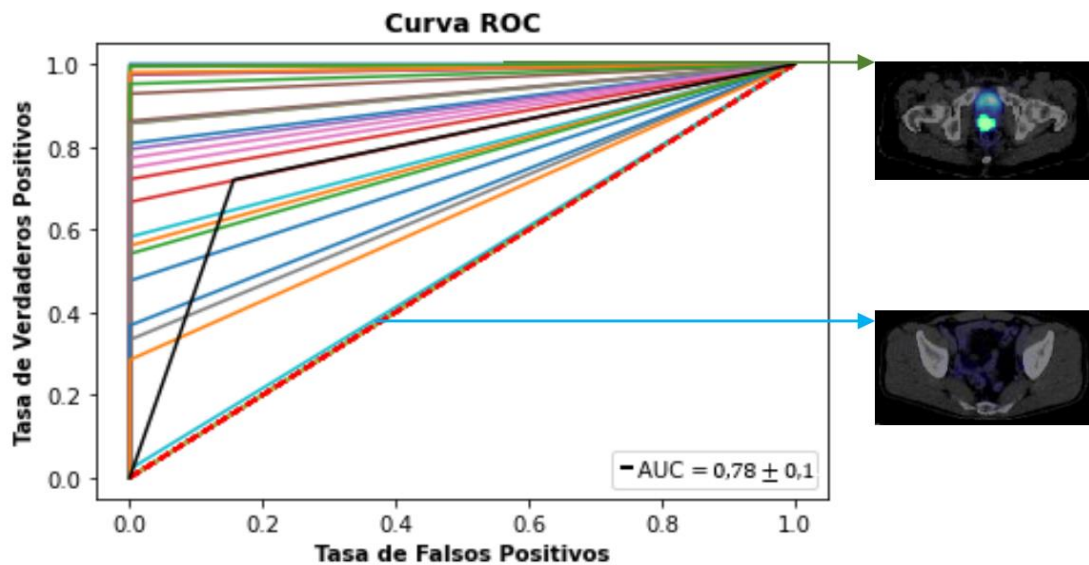


Figura 4.6 Curva ROC del modelo evaluado.

### 4.2.3 Determinación de la arquitectura

Para determinar la arquitectura con el mejor rendimiento, se priorizaron el coeficiente de Dice y la función de pérdida de Jaccard. Esto se debió a la presencia de un desequilibrio en el número de píxeles relacionados con la segmentación médica vs el total de píxeles analizados en las imágenes de los cortes axiales, lo que afecta a índices clásicos como la exactitud, precisión, recall y el AUC, ya que son afectados por este desequilibrio.

En este sentido, como se mencionó en la sección 3.5.2.5 y 3.5.2.6, el índice de Dice y la función de pérdida de Jaccard aborda esta desventaja de los índices desbalanceados, ya que se enfoca en medir la similitud entre la segmentación obtenida y la segmentación de referencia. Por lo tanto, se consideró como el indicador más adecuado para evaluar el desempeño de las arquitecturas en este caso específico.

En la Figura 4.7, se presenta una imagen de prueba que muestra la captación del radiofármaco en la próstata, asociada al carcinoma de próstata, la cual es utilizada para evaluar el desempeño del algoritmo. La imagen del medio muestra la segmentación de la región de interés o referencia, realizada por el médico nuclear. Finalmente, en la imagen de la derecha se muestra la segmentación predicha por el algoritmo. Cada uno de los índices evaluados a continuación, se basa en esta comparación, para medir las diferencias entre la imagen de referencia o segmentación médica y la predicción del algoritmo.

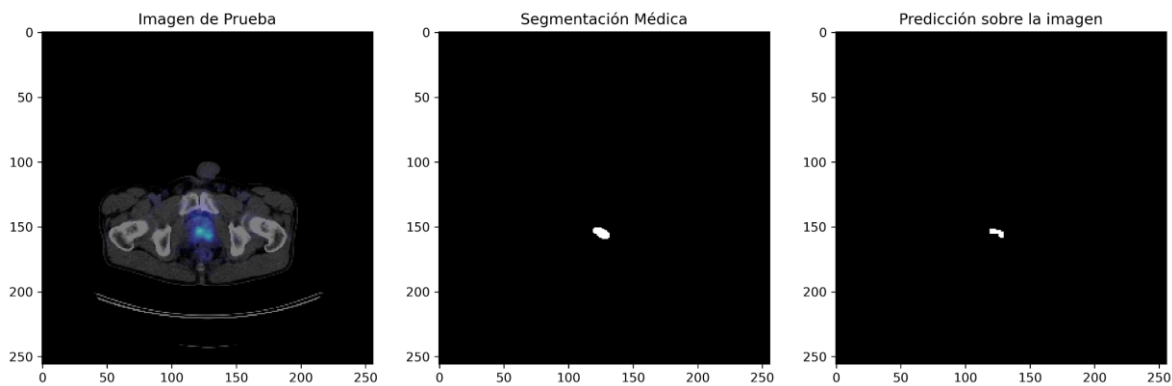


Figura 4.7 Imagen de prueba.

Los resultados obtenidos en la elección de la arquitectura para cada uno de los diferentes índices se presentan en el Anexo II.

Para estimar el rendimiento de la arquitectura a lo largo de todo el conjunto de datos, se promediaron los valores obtenidos en cada iteración de la partición  $k$ . En este sentido, la arquitectura Attention U-NET mostró un coeficiente de Dice promedio de 0,92 en el entrenamiento y 0,48 en la validación, junto con un valor promedio de la función de pérdida de Jaccard de 0,11 en el entrenamiento y 0,66 en la validación (ver Anexo II).

Estos resultados superaron a las arquitecturas U-NET y U-NET++ y, en consecuencia, se seleccionó la arquitectura Attention U-NET como la más adecuada para el trabajo.

#### 4.2.4 Evaluación de la capacidad de generalización

La capacidad de generalización de la red se evalúa para determinar cuán bien la red puede generalizar su aprendizaje a datos no vistos previamente. En otras palabras, se quiere saber si la red es capaz de hacer predicciones precisas en situaciones y datos que no ha encontrado durante su proceso de entrenamiento. Para esto, como se mencionó anteriormente el coeficiente de Dice y la función de pérdida de Jaccard son los utilizados para evaluar la capacidad de generalización de las diferentes arquitecturas como se consigna en el Anexo III.

Independientemente de la métrica utilizada, una red que generaliza bien debe mostrar un rendimiento sólido tanto en el conjunto de entrenamiento como en el conjunto de prueba, evitando el sobreajuste. La capacidad de generalización es crucial en el desarrollo de modelos de aprendizaje profundo, ya que garantiza que el modelo sea útil y

confiable en situaciones del mundo real, donde se enfrenta a datos nuevos y no vistos previamente.

En este contexto, la arquitectura Attention U-NET se destaca por su rendimiento en términos de generalización en comparación con las otras arquitecturas propuestas. Se obtuvo un coeficiente de Dice promedio de 0,99 en el conjunto de entrenamiento y 0,53 en el conjunto de prueba a lo largo de las k particiones planteadas (ver Anexo III). Además, se logró un valor máximo de 0,74 con la partición 8 (ver Figura 4.8).

Estos resultados indican que la arquitectura Attention U-NET es eficaz en la tarea de generalización, lo que la convierte en una opción sólida para abordar problemas del mundo real con datos desconocidos.

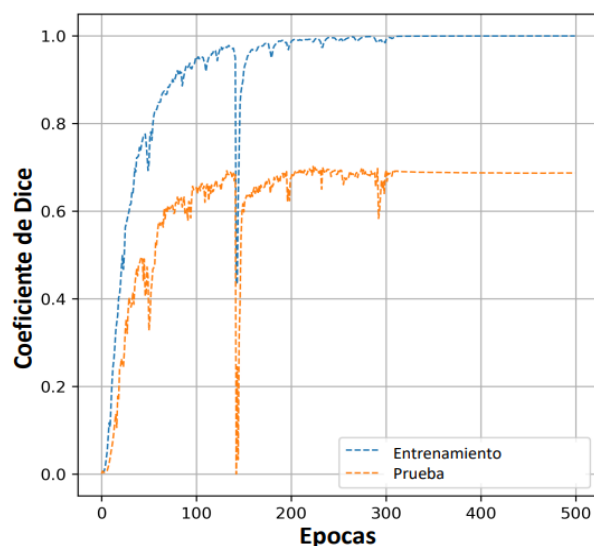


Figura 4.8 Entrenamiento de la arquitectura Attention U-NET para la octava partición.

En la Figura 4.8, se muestra la evolución del coeficiente de Dice durante el entrenamiento de la arquitectura Attention U-NET utilizando la octava partición del conjunto de datos. Se destaca un valor máximo de 0,74 en la prueba. Este patrón de rendimiento se repite consistentemente en todos los procesos de entrenamiento y evaluación de los diversos modelos en el conjunto de datos, aunque con valores máximos inferiores en otras particiones.

Después de evaluar cuidadosamente la capacidad de generalización, se tomó la decisión de implementar la arquitectura Attention-UNET para su posterior uso en un entorno clínico. El modelo fue entrenado con todo el conjunto de datos y, posteriormente, se llevó a cabo la fase de despliegue.



#### 4.2.5 Comparación con método de umbralización

También, se realizó una comparación cualitativa del modelo entrenado en comparación con el método clásico de umbralización que es altamente difundido [64], como se observa en la Figura 4.9.

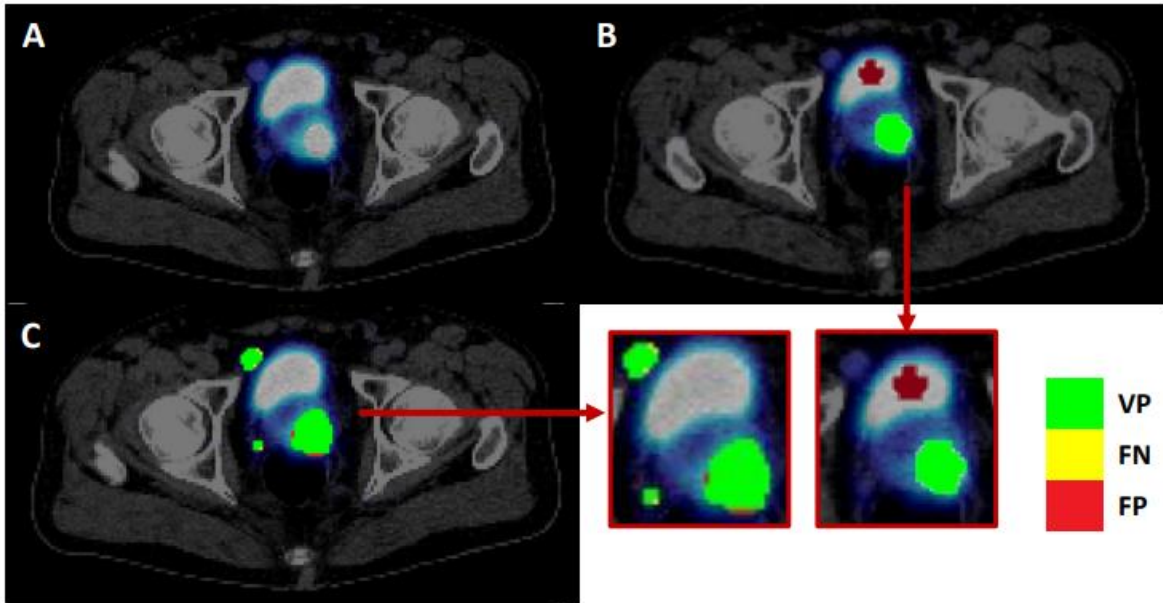


Figura 4.9 Análisis cualitativo del método de umbralización.

En la Figura 4.9 se examinó de manera cualitativa el desempeño del método de umbralización en la segmentación de lesiones en la próstata y los ganglios de la pelvis. En la imagen A, se muestra un corte axial de un paciente con captaciones fisiológicas en la vejiga y captaciones asociadas al carcinoma tanto en un ganglio linfático como en la próstata. La imagen B presenta la segmentación realizada mediante el método de umbralización, mostrando una delimitación moderada de la lesión en la próstata, pero ignorando por completo la captación en el ganglio y además incluyendo captaciones fisiológicas de la vejiga que son incorrectas. Los píxeles verdaderos positivos se representan en color verde, mientras que los píxeles en rojo representan los falsos positivos detectados por el modelo. Es importante notar que el modelo de umbralización presenta limitaciones en este contexto para diferenciar las captaciones de baja intensidad, como es común en los ganglios. Por el contrario, en la imagen C se muestra la segmentación realizada por el modelo propuesto, el cual presenta resultados destacables en todas las lesiones y es capaz de distinguir las captaciones fisiológicas de forma precisa, en comparación con el método de umbralización. En la imagen también se visualiza en color amarillo la región que representa los falsos negativos.



La principal diferencia y limitación del método de umbralización radica en su incapacidad para diferenciar las captaciones fisiológicas normales, lo que puede resultar en la segmentación incorrecta de regiones con captaciones normales, como en el caso de la vejiga. Además, este método enfrenta dificultades cuando se trata de captaciones bajas de radiofármaco, como en el caso de los ganglios linfáticos, debido a la baja diferencia de intensidad entre las lesiones y el fondo, lo que dificulta la elección de un umbral adecuado.

En contraste, el modelo propuesto no solo localiza con precisión los órganos, sino que también produce límites coherentes, incluso en casos con captaciones de baja intensidad en los ganglios linfáticos. Además, el modelo es capaz de distinguir captaciones fisiológicas normales que no están asociadas al carcinoma de próstata, lo que mejora significativamente la precisión y confiabilidad de la segmentación en comparación con el método de umbralización. La utilización del modelo propuesto brinda resultados más precisos y robustos, lo que resulta fundamental para el análisis adecuado y el diagnóstico preciso de las lesiones relacionadas con el cáncer de próstata.

### **4.3 Discusión de los resultados**

En el desarrollo de este trabajo se propone una herramienta novedosa y completamente automática para la segmentación y cuantificación de la carga metabólica tumoral en pacientes con carcinoma de próstata utilizando imágenes PET/CT de  $^{68}\text{Ga}$ -PSMA. Esta herramienta, está diseñada para facilitar la tarea al médico nuclear, requiere únicamente un previo acondicionamiento de las imágenes antes de ser sometidas al análisis. Para ello, se utiliza un examen de imágenes axiales PET/CT fusionadas en formato TIFF. El componente inteligente de la herramienta realiza un preprocesamiento y procesamiento sin intervención humana de cada corte, obteniendo finalmente una imagen de las lesiones delimitadas y una cuantificación en ml de dichos volúmenes.

Es crucial destacar que el análisis y la evaluación se llevaron a cabo empleando los coeficientes mencionados en la sección 3.5.1. Además, subrayar que los errores y las mediciones realizadas en píxeles guardan una relación directa con las medidas en mililitros. En consecuencia, la magnitud de los errores se equipará directamente con la unidad primaria de medida, es decir, los mililitros. También, es necesario destacar que cuantificar los errores en la representación de píxeles permite describir más detalladamente las discrepancias detectadas en la representación de las estructuras analizadas.

Es significativo destacar que esta herramienta solo fue entrenada en lesiones de la próstata y ganglios linfáticos de la pelvis debido a la falta de información suficiente en el conjunto de datos para abordar otras regiones anatómicas. Por lo tanto, se requiere una actualización y entrenamiento para las regiones faltantes.

En contraste, hasta donde se tiene conocimiento, actualmente la única herramienta disponible en la literatura para evaluar la carga tumoral en todo el cuerpo de pacientes con cáncer de próstata es el software qPSMA [22]. Sin embargo, esta herramienta presenta ciertas deficiencias. En primer lugar, es semiautomática y requiere la intervención de un especialista en medicina nuclear. Además, se basa en el método de umbralización, lo que resulta en una variabilidad y una fuerte dependencia del umbral establecido para distinguir las captaciones benignas de las malignas. El método propuesto, en cambio, no requiere intervención humana en la tarea directa de cuantificación de la carga tumoral.

En el caso de las lesiones en tejido blando, se utiliza el umbral hepático para segmentar las lesiones. Se establece un volumen esférico sobre el cual se mide el SUV (valor estandarizado de captación) del lóbulo hepático derecho, considerando cualquier valor por encima de este umbral como maligno. Sin embargo, la presencia de metástasis hepáticas puede interferir con el umbral basado en el hígado, incluyendo lesiones hepáticas o pasando por alto lesiones en otras regiones con una captación más baja que la actividad de fondo del hígado.

Para las lesiones óseas, se utiliza un algoritmo de detección de hueso basado en la imagen de tomografía computarizada (CT) para determinar la región a analizar y luego se contrasta con la imagen de PET asociada a la lesión. En este caso, se establece un umbral de SUV igual a 3 para determinar las lesiones, lo que puede provocar un desalineamiento entre las imágenes de PET y CT. Además, algunas lesiones óseas pueden ubicarse fuera de la máscara ósea y, como resultado, se segmentan incorrectamente como lesiones en tejidos blandos. Por otro lado, el método propuesto en este trabajo utiliza imágenes fusionadas de PET/CT, lo que evita falencias en la localización anatómica de las lesiones.

Otra limitación de qPSMA es la necesidad de una gran intervención humana, lo que requiere un entrenamiento y familiarización con la herramienta. Por lo general, se requieren correcciones manuales para delinear el intestino y los ganglios linfáticos positivos para el ligando PSMA abdominal, así como para eliminar la captación de falsos positivos en estructuras con captación inespecífica, como la aorta, el esófago, el uréter y el recto. En contraste, el método propuesto sólo requiere una adaptación previa del examen al formato TIFF lo que puede reducir en gran medida el tiempo que tarda el médico en la cuantificación de la MTB.

Además, qPSMA está diseñada para facilitar la delimitación de las lesiones y no se evalúa en la predicción de lesiones, por lo que no implementa coeficientes de evaluación de la calidad de la segmentación contrastados con una segmentación manual realizada por el médico. Por el contrario, el método propuesto sí realiza dicho análisis.

La solución propuesta no requiere intervención humana específica para la cuantificación de la carga metabólica tumoral, una tarea que suele depender de métodos manuales de morfología o segmentación. Esta solución no interrumpe los procesos

diagnósticos llevados a cabo por el médico. En lugar de eso, actúa como un apoyo a esta tarea al delinear las lesiones en las áreas de interés designadas. Es crucial aclarar que el usuario debe proporcionar previamente la región de interés, como se destacó previamente solo puede ser utilizada en la pelvis; la herramienta en sí no realiza la selección de regiones de forma automática. Su función se centra en asistir en la delimitación precisa de las áreas identificadas por el médico, dejando la decisión final en manos del usuario en cuanto a qué áreas se consideran relevantes para el análisis.

El método propuesto presenta gran versatilidad en diferentes contextos al no requerir intervención humana, lo que garantiza la repetibilidad de los resultados, además, es capaz de distinguir captaciones fisiológicas benignas. Sin embargo, es importante tener en cuenta que todavía no se ha establecido ningún método de segmentación para el PET como el estándar de oro, por lo que se requieren estudios futuros en este campo.

En general, el modelo aborda diferentes escenarios de segmentación en la próstata, mostrando su capacidad para lograr resultados precisos y los desafíos que enfrenta en ciertas situaciones específicas como lo es una captación leve del radiofármaco.

#### **4.4 Despliegue del componente inteligente**

De acuerdo con la arquitectura propuesta en la Figura 3.15, se elaboró un código fuente para el análisis de los exámenes PET/CT con la siguiente estructura:

- Se requiere un examen PET/CT tomado con el protocolo descrito en la sección 3.2 de imágenes fusionadas.
- Las imágenes fusionadas requieren un recorte a una matriz 256 x 256 como se describe en la sección 3.2 y deben ser convertidas al formato TIFF.
- Una vez conocida la ruta de almacenamiento, esta debe ser ingresada en el código fuente.
- La primera etapa con la que se encuentra el examen es el módulo de preprocesamiento, en el cual son binarizadas y convertidas a escala de grises para su procesamiento.
- Posteriormente en el módulo de segmentación el modelo entrenado se encarga de realizar la predicción y segmentación de la carga metabólica tumoral, es importante

notar que solo es capaz de determinar lesiones de próstata y algunas de los ganglios linfáticos en la pelvis.

- Una vez delimitado, se realiza la estimación del volumen en mililitros de la MTB.
- Finalmente, el código fuente despliega imágenes axiales con su respectiva delimitación de la carga metabólica tumoral como se observa en la Figura 4.6, adicionalmente genera un archivo .TIFF el cual puede ser utilizado para su análisis en diferentes herramientas de procesamiento de imágenes e incluso ser desplegadas en 3D como se observa en la Figura 4.10.

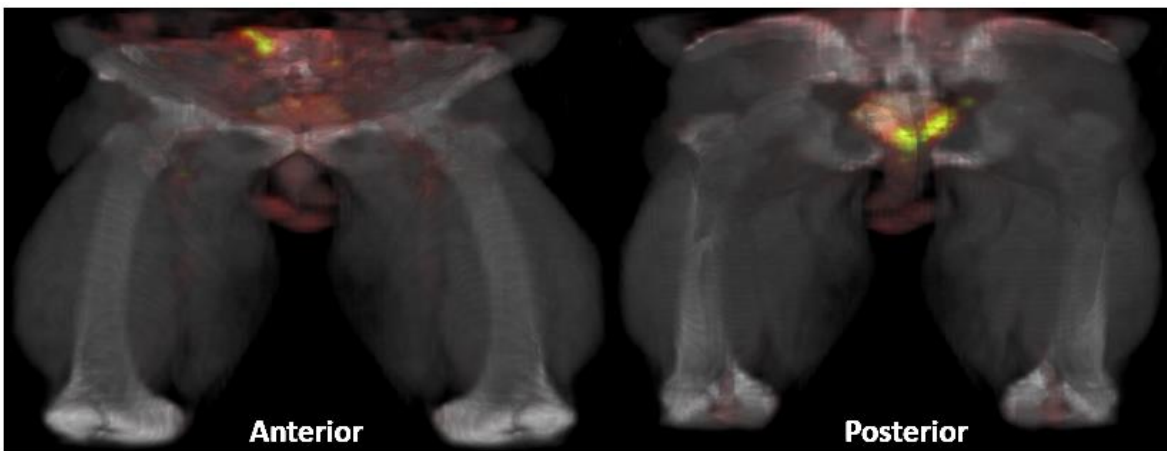


Figura 4.10 Visualización 3D de la MTB. El color verde representa el volumen delimitado por el componente inteligente.

El código fuente fue desarrollado en lenguaje python 3.0.x y requiere las siguientes librerías:

- Keras
- Tensor Flow
- Numpy
- Matplotlib
- cv2
- PIL
- Pandas
- tiff file

Para garantizar un funcionamiento óptimo, se necesitan al menos 16GB de memoria RAM y una tarjeta de video NVIDIA RTX 2060 como requisitos mínimos.

El tiempo de análisis de una imagen de un corte axial es de aproximadamente 30 ms, mientras que, para un examen promedio, teniendo en cuenta solo los cortes de la pelvis, el tiempo de análisis puede llevar alrededor de 4 segundos.

## Capítulo 5.

### Conclusiones y trabajos futuros

#### 5.1 Conclusiones

Los resultados obtenidos en este trabajo presentan una perspectiva alentadora en la cuantificación de la carga metabólica tumoral en exámenes PET/CT de  $^{68}\text{Ga}$ -PSMA. La implementación de redes neuronales convolucionales permite realizar la tarea de forma automática y con mínima intervención humana, reduciendo significativamente los tiempos requeridos en comparación con los métodos manuales y de umbralización. Así mismo, demuestran un potencial para mejorar la eficiencia y precisión en el diagnóstico y seguimiento del carcinoma de próstata. Además de estos beneficios, el trabajo ha proporcionado las siguientes conclusiones:

- La CNN implementada en el componente inteligente ha sido capaz de segmentar y cuantificar automáticamente la carga metabólica tumoral en dos tipos de lesiones asociadas al carcinoma de próstata, específicamente en los ganglios linfáticos y la próstata. El coeficiente de Dice promedio obtenido fue de 0,53.
- Además, la CNN ha demostrado habilidad para reconocer el contexto anatómico de las imágenes axiales, permitiéndole diferenciar captaciones fisiológicas benignas, como en el caso de la vejiga. Esto evita incluir dichas regiones en la cuantificación de la carga metabólica tumoral, a diferencia de las herramientas basadas en el método de umbralización.
- Se destaca la importancia de la recolección adecuada del conjunto de datos, incluyendo un proceso de anotación/segmentación precisa. La calidad de los datos afecta directamente la capacidad de generalización del modelo, y en este trabajo, se obtuvieron buenos resultados debido al éxito en este proceso.
- Aunque el conjunto de datos recolectado permitió abordar el problema de la cuantificación automática de la carga tumoral, se sugiere que, para mejorar el modelo, se necesitaría incluir una mayor variedad de lesiones tanto en tejido blando como en lesiones óseas en regiones anatómicas fuera de la pelvis.

- Se implementaron tres arquitecturas de redes neuronales: U-NET, UNET++, y Attention U-NET. Tras la evaluación a lo largo del conjunto de datos, se determinó que la red Attention U-NET presentó el mejor desempeño. Por lo tanto, esta arquitectura es prometedora para implementaciones futuras en otros contextos del análisis de imágenes PET/CT.

## 5.2 Trabajos futuros

El trabajo realizado deja una serie de líneas abiertas y sesgos para su extensión futura:

- El proceso de segmentación y anotación realizado por el médico requiere exámenes en los cuales no se conozca previamente el paciente. Explorar la posibilidad de reducir esta dependencia es un área de interés para futuras extensiones.
- Se requiere aumentar la cantidad de médicos nucleares que realicen el proceso de anotación/segmentación, para reducir la dependencia del observador.
- Se asumió independencia entre las imágenes de los cortes axiales, debido a que existen cambios anatómicos entre cortes contiguos en un mismo paciente, sin embargo, se requiere explorar si el sistema presenta alguna dependencia entre ellas.
- Dado que las metástasis son características del cáncer avanzado, sería valioso incluir pacientes en esos estadios de la enfermedad para obtener un estudio más integral y representativo.
- Recopilar datos adicionales que abarquen metástasis en huesos, pulmones, hígado y ganglios linfáticos fuera de la pelvis es esencial para mejorar la eficacia del modelo en diferentes regiones anatómicas.
- La necesidad de alimentar modelos por separado para segmentar las captaciones de los ganglios linfáticos y la región de la próstata debe abordarse, ya que cada tipo de lesión presenta características únicas que afectan el rendimiento de un solo modelo.

- Aumentar el conjunto de datos para el entrenamiento y mejorar la capacidad de generalización del modelo es un aspecto importante a considerar.
- Explorar la generación de datos mediante transformaciones no rígidas a partir del conjunto de datos original puede ayudar a simular nuevos contextos y mejorar la capacidad de generalización del modelo.
- Establecer un protocolo estandarizado para la reconstrucción de imágenes PET/CT es crucial, ya que la variabilidad en el número y distancia entre imágenes afecta la estimación de la carga metabólica tumoral.
- Realizar experimentos para ajustar mejor los hiperparámetros relacionados con el problema de segmentación semántica de imágenes médicas permitirá mejorar la precisión y eficacia del modelo.
- Explorar arquitecturas alternativas de redes neuronales convolucionales puede conducir a mejoras en el desempeño del modelo, brindando la oportunidad de seleccionar la más adecuada para el contexto específico.
- Desarrollar un módulo que permita la adecuación previa de los exámenes sin intervención del usuario es una meta importante para agilizar el proceso y aumentar la automatización.
- Una interfaz amigable con el usuario para interactuar con la herramienta y corregir posibles errores en las segmentaciones realizadas por el componente inteligente será útil para mejorar la precisión y confianza en los resultados.
- Implementar el algoritmo en un entorno clínico que permita su integración con sistemas PACS facilitará su aplicación en un contexto real y su adopción en la práctica médica.



## Referencias

- [1] P. Rawla, «Epidemiology of Prostate Cancer», *World J Oncol*, vol. 10, n.º 2, p. 63, 2019, doi: 10.14740/WJON1191.
- [2] R. B. dos Reis *et al.*, «Prostate Cancer in Latin America: Challenges and Recommendations», *Cancer Control*, vol. 27, n.º 1, ene. 2020, doi: 10.1177/1073274820915720.
- [3] «Páginas - Incidencia cáncer de próstata en Colombia». Accedido: 9 de agosto de 2022. [En línea]. Disponible en: <https://www.sispro.gov.co/observatorios/oncancer/indicadores/Paginas/C%C3%A1ncer-de-pr%C3%B3stata.aspx>
- [4] T. Matoba *et al.*, «Impact of tumor burden on survival in patients with recurrent or metastatic head and neck cancer treated with immune checkpoint inhibitors», *Scientific Reports 2022 12:1*, vol. 12, n.º 1, pp. 1-8, ago. 2022, doi: 10.1038/s41598-022-18611-z.
- [5] C. Hurley *et al.*, «Comparison of 18F-FDG-PET-CT and bone scintigraphy for evaluation of osseous metastases in newly diagnosed and recurrent osteosarcoma», *Pediatr Blood Cancer*, vol. 63, n.º 8, p. 1381, ago. 2016, doi: 10.1002/PBC.26014.
- [6] F. Dondi, D. Albano, F. Bertagna, y G. Treglia, «Bone Scintigraphy versus PSMA-Targeted PET/CT or PET/MRI in Prostate Cancer: Lessons Learned from Recent Systematic Reviews and Meta-Analyses», *Cancers (Basel)*, vol. 14, n.º 18, sep. 2022, doi: 10.3390/CANCERS14184470.
- [7] I. Tselchlidis y A. Vrachimis, «PSMA PET in Imaging Prostate Cancer», *Front Oncol*, vol. 12, p. 170, ene. 2022, doi: 10.3389/FONC.2022.831429/BIBTEX.
- [8] J. Kuten *et al.*, «Head-to-Head Comparison of 68 Ga-PSMA-11 with 18 F-PSMA-1007 PET/CT in Staging Prostate Cancer Using Histopathology and Immunohistochemical Analysis as a Reference Standard», *J Nucl Med*, vol. 61, n.º 4, pp. 527-532, abr. 2020, doi: 10.2967/JNUMED.119.234187.
- [9] S. Hoberück *et al.*, «Intraindividual comparison of [68 Ga]-Ga-PSMA-11 and [18F]-F-PSMA-1007 in prostate cancer patients: a retrospective single-center analysis», *EJNMMI Res*, vol. 11, n.º 1, pp. 1-18, oct. 2021, doi: 10.1186/S13550-021-00845-Z/TABLES/9.
- [10] S. Schmuck *et al.*, «Initial Experience with Volumetric 68Ga-PSMA I&T PET/CT for Assessment of Whole-Body Tumor Burden as a Quantitative Imaging Biomarker in

Patients with Prostate Cancer», *J Nucl Med*, vol. 58, n.º 12, pp. 1962-1968, dic. 2017, doi: 10.2967/JNUMED.117.193581.

- [11] H. J. Im, T. Bradshaw, M. Solaiyappan, y S. Y. Cho, «Current Methods to Define Metabolic Tumor Volume in Positron Emission Tomography: Which One is Better?», *Nucl Med Mol Imaging*, vol. 52, n.º 1, pp. 5-15, feb. 2018, doi: 10.1007/S13139-017-0493-6.
- [12] F. Barbato *et al.*, «PSMA-PET for the assessment of metastatic hormone-sensitive prostate cancer volume of disease», *J Nucl Med*, vol. 62, n.º 12, pp. 1747-1750, dic. 2021, doi: 10.2967/JNUMED.121.262120.
- [13] C. J. Sweeney *et al.*, «Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer», *N Engl J Med*, vol. 373, n.º 8, pp. 737-746, ago. 2015, doi: 10.1056/NEJMOA1503747.
- [14] C. C. Parker *et al.*, «Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial», *The Lancet*, vol. 392, n.º 10162, pp. 2353-2366, dic. 2018, doi: 10.1016/S0140-6736(18)32486-3/ATTACHMENT/20C2D10F-7890-48AB-ABB8-E0BBEF754C89/MMC1.PDF.
- [15] R. Seifert *et al.*, «Semiautomatically Quantified Tumor Volume Using 68 Ga-PSMA-11 PET as a Biomarker for Survival in Patients with Advanced Prostate Cancer», *J Nucl Med*, vol. 61, n.º 12, pp. 1786-1792, dic. 2020, doi: 10.2967/JNUMED.120.242057.
- [16] R. Seifert *et al.*, «Total tumor volume reduction and low PSMA expression in patients receiving Lu-PSMA therapy», *Theranostics*, vol. 11, n.º 17, pp. 8143-8151, 2021, doi: 10.7150/THNO.60222.
- [17] C. Parkinson, C. Matthams, K. Foley, y E. Spezi, «Artificial intelligence in radiation oncology: A review of its current status and potential application for the radiotherapy workforce», *Radiography*, vol. 27, pp. S63-S68, oct. 2021, doi: 10.1016/J.RADI.2021.07.012.
- [18] H. Zaidi, I. E. N.-E. journal of nuclear medicine and molecular, y undefined 2010, «PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques», *Springer*, vol. 37, n.º 11, pp. 2165-2187, nov. 2010, doi: 10.1007/s00259-010-1423-3.
- [19] M. Evans *et al.*, «Head and neck target delineation using a novel PET automatic segmentation algorithm», *Elsevier*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0167814016344255>

- [20] Z. Zhong *et al.*, «Simultaneous cosegmentation of tumors in PET-CT images using deep fully convolutional networks», *Med Phys*, vol. 46, n.º 2, pp. 619-633, feb. 2019, doi: 10.1002/MP.13331.
- [21] Y. Guo *et al.*, «Automatic lung tumor segmentation on PET/CT images using fuzzy Markov random field model», *hindawi.com*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://www.hindawi.com/journals/cmmm/2014/401201/>
- [22] A. Gafita *et al.*, «qPSMA: Semiautomatic Software for Whole-Body Tumor Burden Assessment in Prostate Cancer Using 68 Ga-PSMA11 PET/CT», *J Nucl Med*, vol. 60, n.º 9, pp. 1277-1283, 2019, doi: 10.2967/JNUMED.118.224055.
- [23] D. Jin, D. Guo, T. Ho, A. Harrison, ... J. X.-M. I., y undefined 2021, «DeepTarget: Gross tumor and clinical target volume segmentation in esophageal cancer radiotherapy», *Elsevier*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1361841520302735>
- [24] A. Groendahl, I. Knudtsen, ... B. H.-P. in M., y undefined 2021, «A comparison of methods for fully automatic segmentation of tumors and involved nodes in PET/CT of head and neck cancers», *iopscience.iop.org*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://iopscience.iop.org/article/10.1088/1361-6560/abe553/meta>
- [25] Y. Yu, P. Decazes, ... J. L.-L.-... M. I. and, y undefined 2018, «Semi-automatic lymphoma detection and segmentation using fully conditional random fields», *Elsevier*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0895611118301307>
- [26] A. Kerhet, C. Small, H. Quon, T. Riauka, L. S.-C. oncology, y undefined 2010, «Application of machine learning methodology for PET-based definition of lung cancer», *mdpi.com*, Accedido: 6 de noviembre de 2021. [En línea]. Disponible en: <https://www.mdpi.com/941962>
- [27] R. Fonti, M. Conson, y S. Del Vecchio, «PET/CT in radiation oncology», *Semin Oncol*, vol. 46, n.º 3, pp. 202-209, jun. 2019, doi: 10.1053/J.SEMINONCOL.2019.07.001.
- [28] «What Is Prostate Cancer?| Prostate Cancer Types | American Cancer Society». Accedido: 22 de julio de 2023. [En línea]. Disponible en: <https://www.cancer.org/cancer/types/prostate-cancer/about/what-is-prostate-cancer.html>
- [29] C. L. Rock *et al.*, «American Cancer Society guideline for diet and physical activity for cancer prevention», *CA Cancer J Clin*, vol. 70, n.º 4, pp. 245-271, jul. 2020, doi: 10.3322/CAAC.21591.

- [30] «Pruebas para diagnosticar y determinar la etapa del cáncer de próstata | American Cancer Society». Accedido: 22 de julio de 2023. [En línea]. Disponible en: <https://www.cancer.org/es/cancer/tipos/cancer-de-prostata/deteccion-diagnostico-clasificacion-por-etapas/como-se-diagnostica.html>
- [31] S. Schmuck *et al.*, «Initial Experience with Volumetric 68Ga-PSMA I&T PET/CT for Assessment of Whole-Body Tumor Burden as a Quantitative Imaging Biomarker in Patients with Prostate Cancer», *J Nucl Med*, vol. 58, n.º 12, pp. 1962-1968, dic. 2017, doi: 10.2967/JNUMED.117.193581.
- [32] M. Unterrainer *et al.*, «Recent advances of PET imaging in clinical radiation oncology», *Radiation Oncology*, vol. 15, n.º 1, abr. 2020, doi: 10.1186/S13014-020-01519-1.
- [33] F. G. Dall’Olio *et al.*, «Tumour burden and efficacy of immune-checkpoint inhibitors», *Nature Reviews Clinical Oncology* 2021 19:2, vol. 19, n.º 2, pp. 75-90, oct. 2021, doi: 10.1038/s41571-021-00564-3.
- [34] J. Xiang *et al.*, «Metabolic tumor burden: a new promising way to reach precise personalized therapy in PDAC», *Cancer Lett*, vol. 359, n.º 2, pp. 165-168, abr. 2015, doi: 10.1016/J.CANLET.2015.01.023.
- [35] F. C. Gaertner *et al.*, «Uptake of PSMA-ligands in normal tissues is dependent on tumor load in patients with prostate cancer», *Oncotarget*, vol. 8, n.º 33, p. 55094, ago. 2017, doi: 10.18632/ONCOTARGET.19049.
- [36] F. C. Gaertner *et al.*, «Uptake of PSMA-ligands in normal tissues is dependent on tumor load in patients with prostate cancer», *Oncotarget*, vol. 8, n.º 33, p. 55094, ago. 2017, doi: 10.18632/ONCOTARGET.19049.
- [37] S. R. Cherry, J. A. Sorenson, y M. E. Phelps, «Physics in nuclear medicine», p. 523, 2012.
- [38] S. R. Banerjee *et al.*, «68Ga-labeled inhibitors of prostate-specific membrane antigen (PSMA) for imaging prostate cancer», *J Med Chem*, vol. 53, n.º 14, pp. 5333-5341, jul. 2010, doi: 10.1021/JM100623E.
- [39] E. Demirci, O. E. Sahin, M. Ocak, B. Akovali, J. Nematyazar, y L. Kabasakal, «Normal distribution pattern and physiological variants of 68Ga-PSMA-11 PET/CT imaging», *Nucl Med Commun*, vol. 37, n.º 11, pp. 1169-1179, nov. 2016, doi: 10.1097/MNM.0000000000000566.
- [40] «PET/TC (tomografía por emisión de positrones – tomografía computada)». Accedido: 23 de julio de 2023. [En línea]. Disponible en: <https://www.radiologyinfo.org/es/info/pet>

- [41] A. Géron, «Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow (2019, O’reilly)», *Hands-On Machine Learning with R*, p. 510, 2017, Accedido: 23 de julio de 2023. [En línea]. Disponible en: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>
- [42] S. J. Russell y P. Norvig, «Inteligencia artificial: un enfoque moderno (2a. ed.)», *Revista ICONO14 Revista científica de Comunicación y Tecnologías emergentes*, pp. 19-21, 2013, Accedido: 23 de julio de 2023. [En línea]. Disponible en: <https://bv.unir.net:3555/es/ereader/unir/45310>
- [43] S. J. (Stuart J. Russell y P. Norvig, «Artificial intelligence : a modern approach», p. 1167, Accedido: 23 de julio de 2023. [En línea]. Disponible en: [https://books.google.com/books/about/Artificial\\_Intelligence\\_A\\_Modern\\_Approac.html?hl=es&id=cb0qEAAAQBAJ](https://books.google.com/books/about/Artificial_Intelligence_A_Modern_Approac.html?hl=es&id=cb0qEAAAQBAJ)
- [44] D. Cielen, A. Meysman, y M. Ali, «Introducing data science : big data, machine learning, and more, using Python tools», Accedido: 23 de julio de 2023. [En línea]. Disponible en: [https://books.google.com/books/about/Introducing\\_Data\\_Science.html?hl=es&id=bTozEAAAQBAJ](https://books.google.com/books/about/Introducing_Data_Science.html?hl=es&id=bTozEAAAQBAJ)
- [45] A. Géron, «Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow : concepts, tools, and techniques to build intelligent systems», p. 821, Accedido: 24 de julio de 2023. [En línea]. Disponible en: [https://books.google.com/books/about/Hands\\_On\\_Machine\\_Learning\\_with\\_Scikit\\_Le.html?hl=es&id=HHetDwAAQBAJ](https://books.google.com/books/about/Hands_On_Machine_Learning_with_Scikit_Le.html?hl=es&id=HHetDwAAQBAJ)
- [46] «Ch. 9 - Object Detection and Segmentation». Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://manipulation.csail.mit.edu/segmentation.html>
- [47] J. Heaton, «Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016, 800 pp, ISBN: 0262035618 Article», *Genet Program Evolvable Mach*, vol. 19, n.º 1-2, pp. 305-307, 2018, Accedido: 24 de julio de 2023. [En línea]. Disponible en: [https://books.google.com/books/about/Deep\\_Learning.html?hl=es&id=Np9SDQAAQBAJ](https://books.google.com/books/about/Deep_Learning.html?hl=es&id=Np9SDQAAQBAJ)
- [48] «A Comprehensive Introduction to Different Types of Convolutions in Deep Learning | by Kunlun Bai | Towards Data Science». Accedido: 18 de noviembre de 2023. [En línea]. Disponible en: <https://towardsdatascience.com/a-comprehensive-introduction-to-different-types-of-convolutions-in-deep-learning-669281e58215>
- [49] «Convolutional Neural Networks (LeNet) — DeepLearning 0.1 documentation». Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://deeplearningtutorials.readthedocs.io/en/latest/lenet.html>

- [50] «The pipeline of the general CNN architecture [30]. | Download Scientific Diagram». Accedido: 24 de julio de 2023. [En línea]. Disponible en: [https://www.researchgate.net/figure/The-pipeline-of-the-general-CNN-architecture-30\\_fig1\\_329389588](https://www.researchgate.net/figure/The-pipeline-of-the-general-CNN-architecture-30_fig1_329389588)
- [51] «Transfer learning & fine-tuning». Accedido: 24 de julio de 2023. [En línea]. Disponible en: [https://keras.io/guides/transfer\\_learning/](https://keras.io/guides/transfer_learning/)
- [52] P. Chapman *et al.*, «CRISP-DM 1.0: Step-by-step data mining guide». 2000. Accedido: 1 de septiembre de 2019. [En línea]. Disponible en: <https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman-Clinton/54061a4aa0cb241a726f54d0569efae1c13aab3a?tab=abstract&citingPapersSort=year&citingPapersLimit=10&citingPapersOffset=0&citationIntent=methodology&year%5B0%5D=0&year%5B1%5D=0&citedPapersSort=is-influential&citedPapersLimit=10&citedPapersOffset=0>
- [53] Y. (Kyle) Sha, «yingkaisha/keras-unet-collection: v0.1.12», sep. 2021, doi: 10.5281/ZENODO.5449801.
- [54] O. Ronneberger, P. Fischer, y T. Brox, «U-net: Convolutional networks for biomedical image segmentation», *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234-241, 2015, doi: 10.1007/978-3-319-24574-4\_28/COVER.
- [55] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, y J. Liang, «Unet++: A nested u-net architecture for medical image segmentation», *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, pp. 3-11, 2018, doi: 10.1007/978-3-030-00889-5\_1/COVER.
- [56] O. Oktay *et al.*, «Attention U-Net: Learning Where to Look for the Pancreas», abr. 2018, Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://arxiv.org/abs/1804.03999v3>
- [57] «A detailed explanation of the Attention U-Net | by Robin Vinod | Towards Data Science». Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://towardsdatascience.com/a-detailed-explanation-of-the-attention-u-net-b371a5590831>
- [58] «What Does Backbone Mean in Neural Networks? | Baeldung on Computer Science». Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://www.baeldung.com/cs/neural-network-backbone>

- [59] T. L. A. Dinh y F. Aires, «Nested leave-two-out cross-validation for the optimal crop yield model selection», *Geosci Model Dev*, vol. 15, n.º 9, pp. 3519-3535, may 2022, doi: 10.5194/GMD-15-3519-2022.
- [60] «Nested Cross-Validation | Guide to Nested Cross Validation». Accedido: 24 de julio de 2023. [En línea]. Disponible en: <https://www.analyticsvidhya.com/blog/2021/03/a-step-by-step-guide-to-nested-cross-validation/>
- [61] A. Carass *et al.*, «Evaluating White Matter Lesion Segmentations with Refined Sørensen-Dice Analysis», *Scientific Reports 2020 10:1*, vol. 10, n.º 1, pp. 1-19, may 2020, doi: 10.1038/s41598-020-64803-w.
- [62] F. Greco y C. A. Mallio, «Artificial intelligence and abdominal adipose tissue analysis: A literature review», *Quant Imaging Med Surg*, vol. 11, n.º 10, oct. 2021, doi: 10.21037/QIMS-21-370.
- [63] M. A. Rahman y Y. Wang, «Optimizing intersection-over-union in deep neural networks for image segmentation», *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10072 LNCS, pp. 234-244, 2016, doi: 10.1007/978-3-319-50835-1\_22/COVER.
- [64] S. Jardim, J. António, y C. Mora, «Image thresholding approaches for medical image segmentation - short literature review», *Procedia Comput Sci*, vol. 219, pp. 1485-1492, ene. 2023, doi: 10.1016/J.PROCS.2023.01.439.

## Anexos

### Anexo I. Columnas vertebrales.

A continuación, se presentan las arquitecturas compatibles con la librería keras\_unet\_collection (<https://keras.io/api/applications/>):

Tabla A.1 Columnas vertebrales compatibles con keras-unet-collection

Modelo	Tamaño (MB)	Exactitud Top-1 (%)	Exactitud Top-5 (%)	Parámetros (Millones)
VGG16	528	71,3	90,1	138,4
VGG19	549	71,3	90,0	143,7
ResNet50	98	74,9	92,1	25,6
ResNet101	171	76,4	92,8	44,7
ResNet152	232	76,6	93,1	60,4
ResNet50v2	98	76,0	93,0	25,6
ResNet101v2	171	77,2	93,8	44,7
ResNet152v2	232	78,0	94,2	60,4
DenseNet121	33	75,0	92,3	8,1
DenseNet169	57	76,2	93,2	14,3
<b>DenseNet201</b>	<b>80</b>	<b>77,3</b>	<b>93,6</b>	<b>20,2</b>
EfficientNetB7	256	84,3	97,0	66,7

En la Tabla A.1, se presenta la métrica de exactitud Top-1 y Top-5. La exactitud Top-1 se refiere a la precisión convencional, lo cual implica que la respuesta del modelo (aquella con la probabilidad más alta) debe coincidir exactamente con la respuesta esperada. Por otro lado, la exactitud Top-5 implica que una de las cinco respuestas con mayor probabilidad debe coincidir con la respuesta esperada.



## Anexo II. Resultados de la selección de la arquitectura.

Las tablas A.2-7 presentan los resultados obtenidos en la elección de la arquitectura para cada uno de los diferentes índices. En la tabla A.2 se consignan los coeficientes de Dice para cada arquitectura.

Tabla A.2 Coeficientes de Dice para el bucle interno

Coeficiente de Dice						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,858	0,488	<b>0,931</b>	<b>0,387</b>	0,956	0,531
C2	0,882	0,499	0,751	0,348	0,962	0,534
C3	0,811	0,442	0,762	0,259	0,942	0,460
C4	0,758	0,388	0,914	0,306	0,848	0,424
C5	0,811	0,466	0,899	0,360	0,978	0,543
C6	0,920	0,539	0,931	0,387	0,955	0,501
C7	0,948	0,558	0,917	0,339	0,896	0,522
C8	0,843	0,508	0,930	0,288	0,845	0,390
<b>Promedio</b>	0,853	0,486	0,872	0,334	<b>0,923</b>	<b>0,488</b>

Tabla A.3 Función de pérdida de Jaccard para el bucle interno.

Función de pérdida de Jaccard						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,195	0,666	0,129	0,747	0,075	0,632
C2	0,182	0,659	0,335	0,779	0,071	0,629
C3	0,264	0,697	0,312	0,837	0,103	0,693
C4	0,308	0,743	0,156	0,810	0,184	0,714
C5	0,228	0,672	0,179	0,767	0,042	0,625
C6	0,124	0,625	0,129	0,748	0,084	0,651
C7	0,093	0,603	0,151	0,788	0,132	0,623
C8	0,240	0,653	0,130	0,823	0,218	0,746
<b>Promedio</b>	0,204	0,665	0,190	0,787	<b>0,114</b>	<b>0,664</b>

En las Tablas A.4-10 se observan resultados destacados para todos los índices en todas las arquitecturas, con diferencias cuantitativas mínimas. Estas pequeñas variaciones se deben a la sensibilidad que tienen estos índices al desbalance o mayor proporción de zonas donde no se encuentran las lesiones. En otras palabras, todos los modelos demuestran la habilidad de distinguir estas áreas. Por lo tanto, la elección de la arquitectura se centró principalmente en el coeficiente de Dice y la función de pérdida de Jaccard, como métricas fundamentales para el análisis y la selección de la arquitectura más adecuada.

Tabla A.4 Exactitud para el bucle interno.

Exactitud						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,9993	0,9981	0,9999	0,9980	0,9999	0,9985
C2	0,9998	0,9986	0,9996	0,9983	0,9999	0,9987
C3	0,9996	0,9985	0,9995	0,9980	0,9999	0,9983
C4	0,9996	0,9985	0,9998	0,9981	0,9997	0,9985
C5	0,9996	0,9985	0,9998	0,9978	0,9999	0,9985
C6	0,9999	0,9985	0,9999	0,9980	0,9999	0,9984
C7	0,9999	0,9990	0,9999	0,9984	<b>0,9998</b>	<b>0,9990</b>
C8	0,9996	0,9984	0,9999	0,9979	0,9995	0,9976
<b>Promedio</b>	<b>0,9997</b>	<b>0,9985</b>	0,9998	0,9981	0,9998	0,9984

Tabla A.5 Precisión para el bucle interno.

Precisión						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,8803	0,6337	0,9600	0,4876	0,9960	0,7037
C2	0,9539	0,6269	0,9232	0,4763	0,9916	0,6567
C3	0,9770	0,5596	0,8340	0,3769	0,9878	0,6263
C4	0,8372	0,5527	0,9399	0,3686	0,8704	0,5606

C5	0,8634	0,6138	0,9437	0,4179	0,9956	0,7023
C6	0,9759	0,6613	0,9513	0,4716	0,9814	0,6739
C7	0,9941	0,7194	0,9403	0,4157	0,9946	0,7102
C8	0,9214	0,6445	0,9616	0,4171	0,8914	0,4791
<b>Promedio</b>	0,9254	0,6265	0,9318	0,4290	<b>0,9636</b>	<b>0,6391</b>

Tabla A.6 Recall para el bucle interno.

Recall						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,9724	0,5619	0,9677	0,3912	0,9686	0,4896
C2	0,9339	0,5032	0,8062	0,3348	0,9793	0,5091
C3	0,8385	0,4339	0,8008	0,2404	0,9589	0,4591
C4	0,7793	0,3360	0,9750	0,3177	0,8505	0,4128
C5	0,8248	0,4444	0,9477	0,3609	0,9897	0,4825
C6	0,9764	0,5311	0,9831	0,3893	0,9812	0,4632
C7	0,9530	0,5110	0,9707	0,3290	0,8951	0,4844
C8	0,9321	0,5241	0,9716	0,2674	0,9690	0,5162
<b>Promedio</b>	<b>0,9013</b>	<b>0,4807</b>	0,9279	0,3288	0,9490	0,4771

Tabla A.7 Área bajo la curva ROC (AUC) para el bucle interno.

Área bajo la curva ROC (AUC)						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Validación	Entrenamiento	Validación	Entrenamiento	Validación
C1	0,9996	0,8482	0,9998	0,7524	0,9995	0,7975
C2	0,9988	0,8575	0,9451	0,7507	0,9998	0,8176
C3	0,9792	0,8139	0,9415	0,6903	0,9993	0,7893
C4	0,9349	0,7495	0,9993	0,7226	0,9966	0,8129
C5	0,9332	0,7836	0,9997	0,7574	0,9999	0,7909
C6	0,9998	0,8213	0,9999	0,7495	0,9999	0,7947
C7	0,9996	0,8073	0,9997	0,7210	0,9876	0,7999
C8	0,9988	0,8684	0,9998	0,6932	0,9994	0,8278
<b>Promedio</b>	<b>0,9805</b>	<b>0,8187</b>	0,9856	0,7296	0,9977	0,8038

### Anexo III. Evaluación de la capacidad de generalización.

A continuación, se presentan los resultados de la evaluación de la capacidad de generalización:

Tabla A.8 Coeficiente de Dice para la evaluación de la capacidad de generalización.

Coeficiente de Dice						
Partición	U-NET		U-NET++		Attention U-NET	
	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	0,99	0,46	0,99	0,22	0,99	0,51
C2	0,98	0,44	0,94	0,18	0,99	0,51
C3	0,99	0,52	0,97	0,48	0,99	0,56
C4	0,98	0,50	0,99	0,37	0,99	0,59
C5	0,99	0,45	0,99	0,08	0,99	0,49
C6	0,99	0,45	0,98	0,33	0,99	0,54
C7	0,99	0,18	0,98	0,10	0,99	0,34
C8	0,98	0,67	0,99	0,49	0,99	0,74
<b>Promedio</b>	0,99	0,46	0,98	0,28	<b>0,99</b>	<b>0,53</b>

Tabla A.9 Función de pérdida de Jaccard para la evaluación de la capacidad de generalización.

Función de pérdida de Jaccard						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	1,46	0,69	6,96	0,87	0,18	0,74
C2	0,02	0,71	0,10	0,90	3,46	0,73
C3	2,26	0,64	0,04	0,67	0,04	0,67
C4	0,03	0,65	7,07E-01	0,77	6,13	0,58
C5	5,83	0,70	1,01	0,95	5,28	0,67
C6	06,50	0,70	0,02	0,79	1,87	0,63
C7	0,01	0,90	0,02	0,94	8,44	0,90
C8	0,02	0,48	4,07	0,67	9,64	0,41
<b>Promedio</b>	1,46	0,69	0,11	0,82	<b>0,02</b>	<b>0,67</b>

En las Tablas A.10-13 se observan resultados destacados para todos los índices en todas las arquitecturas, con diferencias cuantitativas mínimas, de forma análoga a los resultados obtenidos en la elección de la arquitectura. Estas pequeñas variaciones, como se mencionó anteriormente, se deben a la sensibilidad que tienen estos índices al desbalance o mayor proporción de zonas donde no se encuentran las lesiones. Por lo tanto, la capacidad de generalización se centró principalmente en el coeficiente de Dice y la función de pérdida de Jaccard.

Tabla A.10 Exactitud para la evaluación de la capacidad de generalización.

Exactitud						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	1,0000	0,9993	1,0000	0,9988	0,9998	0,9991
C2	1,0000	0,9978	0,9999	0,9968	1,0000	0,9977
C3	1,0000	0,9987	0,9999	0,9986	1,0000	0,9988
C4	1,0000	0,9981	1,0000	0,9979	1,0000	0,9984
C5	1,0000	0,9995	1,0000	0,9993	1,0000	0,9994
C6	1,0000	0,9990	1,0000	0,9984	1,0000	0,9991
C7	1,0000	0,9955	1,0000	0,9953	1,0000	0,9956
C8	1,0000	0,9993	1,0000	0,9989	1,0000	0,9994
<b>Promedio</b>	1,0000	0,9984	1,0000	0,9980	1,0000	0,9984

Tabla A.11 Precisión para la evaluación de la capacidad de generalización.

Precisión						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	1,0000	0,3712	0,9978	0,1653	0,9325	0,3002
C2	0,9950	0,7777	0,9635	0,3341	0,9992	0,7634
C3	1,0000	0,5381	0,9742	0,5057	0,9991	0,6345
C4	0,9834	0,7183	1,0000	0,7101	1,0000	0,8565



C5	1,0000	0,6424	1,0000	0,2136	1,0000	0,5685
C6	1,0000	0,4858	0,9924	0,2782	1,0000	0,5309
C7	0,9983	0,6339	0,9990	0,4642	1,0000	0,6809
C8	0,9981	0,8096	1,0000	0,6429	1,0000	0,8499
<b>Promedio</b>	0,9969	0,6221	0,9909	0,4143	0,9914	0,6481

Tabla A.12 Recall para la evaluación de la capacidad de generalización.

Recall						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	1,0000	0,6423	0,9994	0,3528	0,9630	0,6887
C2	0,9897	0,3147	0,9872	0,1246	0,9999	0,2901
C3	0,9997	0,5085	0,9965	0,4725	0,9758	0,4105
C4	0,9992	0,3958	1,0000	0,2537	1,0000	0,4503
C5	1,0000	0,3570	1,0000	0,0505	1,0000	0,4281
C6	0,9999	0,4298	0,9980	0,4358	1,0000	0,5466
C7	0,9951	0,1048	0,9866	0,0586	1,0000	0,0988
C8	0,9914	0,5900	1,0000	0,3968	1,0000	0,6530
<b>Promedio</b>	0,9969	0,4179	0,9960	0,2681	0,9924	0,4458

Tabla A.13 AUC para la evaluación de la capacidad de generalización.

Área bajo la curva ROC (AUC)						
	U-NET		U-NET++		Attention U-NET	
Partición	Entrenamiento	Prueba	Entrenamiento	Prueba	Entrenamiento	Prueba
C1	1,0000	0,8569	1,0000	0,7089	0,9999	0,9466
C2	0,9998	0,6900	1,0000	0,5914	1,0000	0,6738
C3	1,0000	0,7862	1,0000	0,7649	0,9998	0,7698
C4	1,0000	0,7397	1,0000	0,6422	1,0000	0,7441
C5	1,0000	0,6968	1,0000	0,5336	1,0000	0,7239
C6	1,0000	0,7332	1,0000	0,7651	1,0000	0,7926
C7	1,0000	0,5699	1,0000	0,5366	1,0000	0,5540
C8	1,0000	0,8262	1,0000	0,7182	1,0000	0,8350
<b>Promedio</b>	1,0000	0,7374	1,0000	0,6576	1,0000	0,7550