



**Implementación de un flujo de trabajo para la evaluación de segmentación de imágenes de tomografía de abdomen en población SURA**

Jhonatan Andres Velez Giraldo

Trabajo de grado presentado para optar al título de Bioingeniero

Asesores

Jon Edison Duque, Bioingeniero, Magíster en Ingeniería

Isabella Ariza Cuberos, Bioingeniera

Universidad de Antioquia

Facultad de Ingeniería

Bioingeniería

Medellín, Antioquia, Colombia

2024

---

<b>Cita</b>	Velez Giraldo [1]
-------------	-------------------

---

<b>Referencia</b>	[1]	J.A Velez Giraldo, "Implementación de un flujo de trabajo para la evaluación de segmentación de imágenes de tomografía de abdomen en población SURA", Trabajo de grado profesional, Bioingeniería, Universidad de Antioquia, Medellín, Antioquia, Colombia, 2024.
Estilo IEEE (2020)		

---



Centro de Documentación de la Facultad de Ingeniería (CENDOI)

**Repositorio Institucional:** <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - [www.udea.edu.co](http://www.udea.edu.co)

**Rector:** John Jairo Arboleda Céspedes.

**Decano/Director:** Julio Cesar Saldarriaga.

**Jefe departamento:** John Fredy Ochoa Gomez.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

## **Dedicatoria**

*A mi madre, a quien considero el pilar fundamental en mi desarrollo profesional, le agradezco su resiliencia y su apoyo incondicional, que han sido un faro de luz en los momentos más difíciles.*

*Expreso mi gratitud a mis líderes de proceso, Jon Edinson Duque, Juan Camilo Zapata y Yesika Alexandra Londoño, por su constante colaboración, paciencia, motivación y los valiosos aprendizajes que me han brindado en todo momento.*

*A todas las personas que han creído en mí a lo largo de este camino, les agradezco su apoyo y confianza, los cuales han sido un motor que me ha impulsado a superar desafíos.*

## **Agradecimientos**

*Agradezco a todos mis amigos y familiares que han sido parte de mi proceso formativo a lo largo de los años y que continúan siendo una parte importante en mi vida.*

*A la Universidad de Antioquia, por brindarme una educación de calidad y por todos los conocimientos adquiridos en esta prestigiosa institución.*

*Expreso mi gratitud hacia mis asesores Jazmín Ximena Suarez, Juan Gabriel Castrillón, Jorge Vélez, Juan Pablo Valencia e Isabella Ariza Cuberos, por su constante acompañamiento y por permitirme formar parte del grupo de investigación en imágenes e inteligencia artificial.*

*A la compañía SURA, por confiar en mis habilidades y potencial, y por brindarme una valiosa oportunidad de crecimiento personal y profesional.*

## TABLA DE CONTENIDO

<b>RESUMEN</b> .....	<b>11</b>
<b>ABSTRACT</b> .....	<b>13</b>
<b>I. INTRODUCCIÓN</b> .....	<b>15</b>
<b>II. PLANTEAMIENTO DEL PROBLEMA</b> .....	<b>17</b>
A. Antecedentes.....	20
<b>III. JUSTIFICACIÓN</b> .....	<b>22</b>
<b>IV. OBJETIVOS</b> .....	<b>24</b>
A. Objetivo general.....	24
B. Objetivos específicos.....	24
<b>V. MARCO TEÓRICO</b> .....	<b>25</b>
Tomografía Computarizada .....	25
Casos clínicos.....	27
Segmentación en imágenes .....	28
UNETR .....	31
XNAT.....	31
Quality check(QC) .....	32
Modelos normativos .....	33
Valores atípicos (Outliers).....	35
Control de calidad de los datos .....	39
Modelos de regresión.....	39
Validación cruzada.....	42
Métricas.....	42
Delta de edad.....	43
<b>VI. METODOLOGÍA</b> .....	<b>45</b>
1. Identificar y seleccionar una muestra de sujetos sanos.....	47
• Selección, caracterización e importación de imágenes de tomografía computarizada de abdomen contrastado a la plataforma de procesamiento XNAT...47	
2. Desarrollar un proceso de control de calidad de la segmentación.....	51
• Recibir y comprender el flujo de trabajo para la extracción de volúmenes de los órganos corporales que se realizó previamente donde se integró además la herramienta de MONAI en XNAT.....	51
• Validación de la segmentación realizada previamente del proyecto MONAI a nivel local para la definición y extracción de los volúmenes de: riñones, hígado, bazo, páncreas, estómago, esófago, adrenales entre otros órganos. ....	56
3. Filtración de los datos, para la construcción de un modelo normativo robusto.....	61
• Revisar y aplicar técnicas que permitan detectar valores atípicos para su posterior tratamiento o eliminación en caso de que afecte significativamente la estadística de nuestro modelo normativo. ....	61

- Creación del modelo normativo poblacional SURA. .... 64
- Implementación del modelo en sujetos sanos y patológicos para determinar su funcionalidad. .... 65

**VII. RESULTADOS..... 67**

- 1. Identificar y seleccionar una muestra de sujetos sanos..... 67
- 2. Desarrollar un proceso de control de calidad de la segmentación..... 70
- 3. Filtración de los datos para la construcción de un modelo normativo robusto..... 77

**VIII. DISCUSIÓN..... 91**

**IX. CONCLUSIONES..... 99**

**X. RECOMENDACIONES..... 102**

**REFERENCIAS..... 103**

## LISTA DE TABLAS

Tabla 1 Etiquetas para los diferentes órganos del proyecto .....	46
Tabla 2. Cantidad de datos recopilados de patologías de interés. ....	50
Tabla 3. Distribución de sujetos con los que se implementó el modelo normativo por rango de edad en la población masculina.....	80
Tabla 4. Distribución de sujetos con los que se implementó el modelo normativo por rango de edad en la población femenina.....	81

## LISTA DE FIGURAS

Fig. 1 Flujo de trabajo diseñado por el grupo de investigación.....	45
Fig. 2 Sujeto perteneciente a la población de sanos con nefrectomía. ....	47
Fig. 3. Criterios de inclusión y exclusión aplicados en la conclusión del reporte para recuperación de nuevos estudios.....	49
Fig. 4 Flujo de trabajo implementado para desarrollar un proceso de control de calidad de la segmentación.....	51
Fig. 5 Curva de pérdida del modelo. ....	53
Fig. 6 Quality check inicial en el proyecto de volumetría de abdomen.....	57
Ecuación 12. Cálculo de los diferentes cortes para el Quality Check.....	58
Fig. 7 Quality check implementado para validación. ....	59
Fig. 8 Flujo de trabajo implementado para filtrar los datos y la construcción de un modelo normativo robusto.....	61
Fig. 9 N-grama de las conclusiones de los estudios de la población validada sana. ..	68
Fig. 10 Distribución de la población y edades de los sujetos sanos por género.....	69
Fig. 11 Resultados de la validación del Quality Check .....	71
Fig. 12 Resultados de la validación de la segmentación en el visor Mricron. ....	72
Fig. 13 Sujeto perteneciente a la población de sanos con problemas de segmentación.....	73
Fig 14. Prueba pareada para comparar algunos volúmenes entre sujetos sanos y con hepatomegalia.....	74
Fig 15. Prueba pareada para comparar algunos volúmenes entre sujetos sanos y con esplenomegalia.....	74
Fig.16 Análisis del porcentaje de sujetos patológicos con problemas de segmentación.....	75
Fig.17 QC de sujeto patológico con problemas en su segmentación.....	76
Fig.18 QC de sujeto patológico con problema parcial en su segmentación.....	76
Fig.19 Visualización en el visor MRicron de segmentación perteneciente a la población patológica.....	77
Fig. 20 Distribución la población masculina según el número de outliers.....	78
Fig. 21 Distribución la población femenina según el número de outliers.....	78
Fig. 22 Distribución porcentual de outliers por órgano en población masculina .....	79
Fig. 23 Distribución porcentual de outliers por órgano en población femenina....	80
Fig.24 Distribución de los volúmenes de hígado en la población masculina. ....	82
Fig.25 Distribución de los volúmenes de hígado en la población femenina.....	82
Fig.26 Distribución de los volúmenes de bazo en la población masculina. ....	83
Fig. 27 Distribución de los volúmenes de bazo en la población femenina. ....	83

Fig.28 Resultados comparativos de las regresiones con los volúmenes de hígado en población masculina .....	84
Fig.29 Resultados comparativos de las regresiones con los volúmenes de hígado en población femenina.....	84
Fig.30 Curvas normativas generadas para el hígado en ambas poblaciones.....	85
Fig.31 Resultados comparativos de las regresiones con los volúmenes de bazo en población masculina.....	86
Fig.32 Resultados comparativos de las regresiones con los volúmenes de bazo en población femenina.....	86
Fig.33 Curvas normativas generadas para el bazo en ambas poblaciones.....	87
Fig.34 Prueba de delta de edad considerando el volumen del hígado para la predicción de la edad . .....	88
Fig.35 Prueba de delta de edad considerando el volumen del bazo para la predicción de la edad.....	89
Fig.36 Prueba de delta de edad considerando los diez órganos para la predicción de la edad en sujetos sanos y con hepatomegalia.....	90
Fig.37 Prueba de delta de edad considerando los diez órganos para la predicción de la edad en sujetos sanos y con esplenomegalia. ....	90
Fig. 38 Comparación porcentual de los resultados de la validación de sujetos catalogados como sanos en ambas consultas.....	92
Fig.39 Sujeto identificado con outliers en 10 órganos.....	96
Fig.40 Curva normativa de la arteria aorta.....	97
Fig.41 Relación encontrada entre la población con esplenomegalia y hepatomegalia. ....	98



## LISTA DE ECUACIONES

Ecuación 1. Cálculo del valor Z para la detección de outliers.....	36
Ecuación 2. Determinación de outliers con el método de IQR.....	36
Ecuación 4. Cálculo del límite superior en el método de tukey.....	37
Ecuación 6. Cálculo de la desviación absoluta de la mediana.....	38
Ecuación 7. Cálculo del Z-score a partir de la mediana.....	38
Ecuación 8. Determinación de los outliers con MAD.....	38
Ecuación 9. Cálculo del error cuadrático medio.....	43
Ecuación 10. Cálculo del error cuadrático medio estandarizado.....	43
Delta de edad.....	43
Ecuación 11. Cálculo de delta de edad.....	44
Ecuación 12. Cálculo de los diferentes cortes para el Quality Check.....	58

## **SIGLAS, ACRÓNIMOS Y ABREVIATURAS**

<b>TC</b>	<b>Tomografía Computarizada</b>
<b>IA</b>	<b>Inteligencia Artificial</b>
<b>IMC</b>	<b>Índice de Masa Corporal</b>
<b>MAD</b>	<b>Median Absolute Deviation</b>
<b>UdeA</b>	<b>Universidad de Antioquia</b>
<b>QC</b>	<b>Quality Check</b>
<b>IQR</b>	<b>Rango Intercuartílico</b>
<b>LOESS</b>	<b>Suavizado de diagrama de dispersión estimado localmente</b>
<b>GP</b>	<b>Procesos Gaussianos</b>
<b>RMSE</b>	<b>Error Cuadrático Medio</b>
<b>SMSE</b>	<b>Error Cuadrático Medio Estandarizado</b>

---

## **RESUMEN**

La segmentación en imágenes médicas es fundamental en la radiología y la evaluación de enfermedades. Esta técnica se utiliza para identificar y delinear estructuras anatómicas de interés en imágenes médicas, como tomografías computarizadas (TC) y resonancias magnéticas (RM). Uno de los análisis más comunes realizados con la segmentación es el cálculo del volumen de órganos específicos. Esta medición es esencial en una amplia variedad de evaluaciones clínicas, como la detección temprana de tumores, el seguimiento de la progresión de enfermedades crónicas como la esclerosis múltiple o la evaluación de la respuesta al tratamiento. La capacidad de cuantificar el volumen de órganos con precisión a partir de imágenes médicas ha revolucionado la práctica clínica y ha brindado a los profesionales de la salud una herramienta invaluable para el diagnóstico y el monitoreo de condiciones médicas, es así como la volumetría a partir de segmentación de imágenes e IA es un campo en rápido desarrollo que tiene el potencial de revolucionar la práctica clínica puesto que permite automatizar el proceso de segmentación, lo que puede mejorar la precisión y la eficiencia. Además, la IA permite utilizar imágenes de alta resolución, lo que puede proporcionar mediciones de volumen más precisas.

En la búsqueda de un modelo normativo para el volumen de órganos abdominales a partir de imágenes de tomografía computarizada (TC), se plantea la necesidad de establecer criterios precisos que permitan a los profesionales de la salud evaluar su tamaño confiable. Es allí donde la segmentación de imágenes desempeña un papel crítico en aplicaciones médicas, ya que los resultados obtenidos son fundamentales para el diagnóstico y tratamiento adecuados de algunas patologías.

Esta detección permitirá identificar errores que puedan surgir durante la segmentación de los órganos, entendiendo que la aplicación efectiva de este control de calidad no solo mejorará la precisión del modelo normativo, sino que contribuirá a la reducción de diagnósticos incorrectos y mejorará la toma de decisiones terapéuticas.

Por lo cual se buscó implementar un flujo de trabajo que permita evaluar la calidad de la segmentación de imágenes de tomografía abdominal en la población SURA, precisamente por la necesidad de establecer criterios rigurosos y confiables al

---

evaluar el volumen de órganos abdominales para que se realice de manera inequívoca, considerando que pueda implementarse en el ámbito médico, como herramienta crucial en futuros diagnósticos y tratamientos de patologías.

***Palabras clave* — Volumetría, tomografía, datos atípicos, segmentación, control de calidad, automatización.**

---

**ABSTRACT**

Image segmentation plays a crucial role in the field of radiology and disease assessment in medical imaging. This technique is used to identify and outline anatomical structures of interest in medical images, such as computed tomography (CT) scans and magnetic resonance imaging (MRI). One of the most common analyses performed with segmentation is the calculation of the volume of specific organs. This measurement is essential in a wide range of clinical assessments, including early tumor detection, monitoring the progression of chronic diseases like multiple sclerosis, or evaluating treatment response. The ability to accurately quantify organ volume from medical images has revolutionized clinical practice and provided healthcare professionals with an invaluable tool for diagnosis and monitoring medical conditions. Volumetry from image segmentation and AI is a rapidly developing field with the potential to revolutionize clinical practice by automating the segmentation process, thereby improving accuracy and efficiency. Additionally, AI enables the use of high-resolution images, which can provide more precise volume measurements.

In the quest to develop a normative model for abdominal organ volume from CT images, there is a need to establish precise criteria that allow healthcare professionals to reliably evaluate the size of these organs. This is where image segmentation plays a critical role in medical applications, as the results obtained are crucial for proper diagnosis and treatment of certain pathologies.

This detection will allow the identification of errors that may arise during organ segmentation, understanding that the effective implementation of this quality control will not only improve the accuracy of the normative model but also contribute to the reduction of incorrect diagnoses and enhance therapeutic decision-making.

Therefore, there is a need to implement a workflow that allows for the evaluation of the quality of abdominal CT image segmentation in the SURA population, precisely due to the need to establish rigorous and reliable criteria when evaluating the volume of various abdominal organs to be performed unequivocally, considering at all times that it

---

can be implemented in the medical field as a crucial tool in future diagnoses and treatments of various pathologies.

***Keywords* — Volumetry, tomography, outliers, segmentation, quality control, automation.**

---

## I. INTRODUCCIÓN

En medicina, la segmentación de órganos y estructuras a partir de imágenes médicas, como tomografías computarizadas (CT) o resonancias magnéticas (RM), es fundamental en el diagnóstico y la planificación de tratamientos [1].

La segmentación implica la identificación y delimitación precisa de estructuras de interés en la imagen. Para abordar este desafío, se emplean diversas técnicas de procesamiento de imágenes y aprendizaje automático, como algoritmos de segmentación basados en regiones, contornos o redes neuronales convolucionales (CNN), que pueden contribuir a mejorar la precisión de la segmentación de diferentes órganos [2]. Lograr una segmentación perfecta puede resultar difícil en todas las circunstancias.

Por lo tanto, a pesar de los avances tecnológicos y los métodos automatizados de segmentación de imágenes, la validación y verificación de los resultados siguen siendo responsabilidad del médico especialista. Aunque los algoritmos de segmentación demuestran precisión y eficiencia en la obtención de resultados, solo un médico experimentado puede evaluar integralmente la relevancia clínica de los segmentos identificados y garantizar que los resultados sean coherentes con lo esperado en función de sus conocimientos [3]. Por ello, este estudio presenta un enfoque que complementa la precisión de los métodos de segmentación con la experiencia y el juicio clínico del médico especialista, asegurando así resultados seguros y confiables en el diagnóstico y tratamiento médico.

En este contexto, se evidencia la necesidad de establecer protocolos y procedimientos específicos dentro del flujo de trabajo. Los protocolos buscan identificar y abordar posibles errores o inconsistencias que surjan durante la segmentación de órganos. La razón detrás de esto es garantizar que los resultados obtenidos sean de la más alta calidad y confiabilidad.

Este proceso se asemeja entonces a la implementación de un riguroso control de calidad que permita evaluar y validar los datos generados mediante la segmentación. Uno de los objetivos clave es la detección de valores atípicos o anómalos en el conjunto

---

de datos resultante de este proceso. Estos valores atípicos pueden incluir segmentaciones incorrectas, incoherencias en las mediciones de volumen de los órganos o incluso sujetos patológicos que inicialmente se consideraban como sanos.



## **II. PLANTEAMIENTO DEL PROBLEMA**

La medicina se encuentra actualmente frente a uno de sus mayores desafíos como lo es la obtención de mediciones precisas y coherentes del volumen de los órganos. Esta dificultad surge debido a que los profesionales de la salud suelen realizar estas mediciones de manera aproximada, en gran parte debido a la complejidad estructural del cuerpo humano, que se complica aún más cuando se tiene en consideración una serie de factores que incluyen entre otras, formas irregulares como las presentes en los riñones [4].

En este contexto, nos hemos centrado en abordar esta complejidad al analizar los órganos y estructuras ubicados en la región abdominal, donde la irregularidad es evidente en múltiples órganos o estructuras. Por lo tanto, este enfoque nos plantea el desafío de lograr la precisión en la medición de volúmenes de órganos como el bazo, los riñones, el hígado, entre otros.

Allí cabe destacar además que, al examinar los efectos del envejecimiento en el organismo humano, se revelan una serie de cambios tanto morfológicos como fisiológicos en todos los tejidos. Comprender estas transformaciones es fundamental para distinguir entre las condiciones patológicas que afectan a adultos mayores, individuos con enfermedades y la población en general [5]. En varios órganos, el proceso de envejecimiento implica la pérdida de células sin su reposición, lo que resulta en una reducción del número celular y alteraciones en el volumen o la forma de las estructuras corporales. Un ejemplo representativo de este proceso se observa en el riñón, donde investigaciones previas han documentado una disminución moderada en la velocidad de filtración glomerular, así como cambios en la estructura vascular y glomerular [5].

Estos cambios también pueden manifestarse en otras estructuras, como el cerebro, que experimenta una reducción en su volumen con la edad, lo que se traduce en una menor capacidad de atención, memoria de trabajo y posibles trastornos

---

motores. Además, la disminución de la masa muscular conlleva a un aumento en la infiltración de grasa, lo que resulta en una pérdida progresiva de fuerza [5].

Esta variabilidad en los volúmenes de órganos asociada con la edad se pretende analizar estadísticamente, por lo que nuevamente caemos ante la imperiosa necesidad de controlar la calidad de los volúmenes entregados por el algoritmo, lo que añade un factor crítico, ya que si no se hace un correcto control de calidad se pueden tener consecuencias significativas en el diagnóstico al caer en sesgos, siendo relevante que diversas patologías pueden desencadenar cambios notables en los volúmenes de órganos presentes en el modelo de segmentación. Por tanto, es crucial resaltar la importancia de detectar con precisión dichos cambios en los órganos afectados, evitando cualquier sesgo, lo que, por consiguiente, al tener en cuenta la variación de estos volúmenes a partir de un análisis de datos y posterior depuración se debería mejorar la certeza y eficacia de los diagnósticos médicos a partir de los resultados de la segmentación.

La tomografía computarizada (TC) emerge como una herramienta fundamental en este contexto. Su capacidad para realizar mediciones de volúmenes con precisión facilita a los médicos el diagnóstico preciso de enfermedades, permitiendo identificar con mayor certeza anomalías, lesiones o condiciones patológicas en los órganos o tejidos del paciente [6]. La combinación de la TC con técnicas de segmentación mejora aún más esta precisión al delimitar con exactitud las estructuras anatómicas específicas [7], lo que facilita la detección de anomalías y la evaluación de cambios morfológicos asociados con el envejecimiento o enfermedades.

Además, estas mediciones precisas de volúmenes no solo permiten anticipar el curso de una enfermedad y diseñar el tratamiento más adecuado, sino que también posibilitan supervisar la progresión de enfermedades y evaluar la efectividad de las intervenciones médicas. La precisión es especialmente relevante en enfermedades progresivas o crónicas [8], donde las mediciones exactas pueden influir en la elección de terapias y en la evaluación de su efectividad a lo largo del tiempo.

Por tanto, la precisión en la medición de volúmenes de órganos y estructuras abdominales resulta fundamental para mejorar el diagnóstico y el tratamiento en

---

medicina. La integración de la tomografía computarizada con técnicas de segmentación representa un avance significativo en este objetivo, permitiendo una atención médica más efectiva y personalizada [9].

Sin embargo, es importante destacar que la verificación visual de los resultados del algoritmo de segmentación no es suficiente, ya que como se mencionó anteriormente, la segmentación perfecta no existe. Se requiere una revisión exhaustiva de los resultados de la segmentación, que incluya un análisis detallado de los datos obtenidos, por lo tanto, los diferentes volúmenes obtenidos por cada estructura de cada sujeto de la población deben examinarse minuciosamente mediante un control de calidad. Este análisis profundo puede ayudar a identificar de manera más precisa las posibles incoherencias en los resultados del algoritmo.

---

**A. Antecedentes**

Recientemente, se ha aplicado la segmentación y estimación volumétrica de diversos órganos del cuerpo mediante tomografía computarizada que permitirá así el seguimiento de tratamientos en pacientes. No obstante, la segmentación manual conlleva un componente subjetivo que varía significativamente según la experiencia del radiólogo. Para abordar este desafío, en este estudio se empleó el proyecto MONAI, un marco de trabajo de código abierto que ofrece herramientas para la implementación de aprendizaje automático en aplicaciones médicas [10].

Es crucial destacar que, hasta el momento, existen pocos normogramas a nivel de volumetría corporal que permitan predecir y estimar patologías en órganos abdominales. Para abordar esta carencia, se llevó a cabo un flujo de trabajo destinado a la segmentación automática de imágenes de tomografía y la extracción de volúmenes de distintos órganos abdominales, generando así una fase inicial de una base de datos con información de la población colombiana. El modelo normativo resultante fue creado utilizando un módulo desarrollado por el grupo de investigación en imágenes médicas SURA [10].

La creación de un modelo normativo de volumetría corporal específicamente adaptado a la población colombiana abre nuevas posibilidades en el análisis y diagnóstico de enfermedades abdominales, ofreciendo un enfoque más preciso y personalizado en la evaluación clínica y el seguimiento de tratamientos [10].

En estudios previos, se ha demostrado con éxito la eficacia del modelo de segmentación UNET en órganos abdominales específicos, como la arteria aorta y las glándulas suprarrenales. No obstante, hasta la fecha no se han encontrado investigaciones que aborden un modelo de segmentación que incluya múltiples órganos abdominales. Este aspecto no solo enriquece nuestro proyecto, sino que también plantea un desafío significativo durante la depuración de datos.

Es imprescindible realizar un análisis exhaustivo para determinar las segmentaciones que pueden ser incluidas dentro de la curva del modelo normativo, asegurando que no se vean afectados de manera significativa los resultados

---

generalizados de los 10 órganos. Este proceso de depuración garantizará la coherencia y fiabilidad de los resultados, permitiendo una interpretación precisa y confiable de los datos obtenidos.

---

### **III. JUSTIFICACIÓN**

Se busca desarrollar modelos normativos que puedan implementarse en la práctica clínica para mejorar y agilizar el diagnóstico de diversas patologías a través del análisis de volúmenes estructurales. Estos modelos representan marcos de referencia que describen las características, dimensiones o comportamientos considerados como "normales" en una población específica. Tienen aplicaciones en campos como la medicina, psicología e ingeniería, estableciendo estándares de lo que se considera común o esperado en ese grupo particular de individuos.

Su utilidad principal radica en proporcionar un punto de comparación para evaluar el rendimiento, las características o medidas individuales en relación con el estándar establecido. Esto resulta especialmente relevante en la detección de desviaciones o anomalías, que podrían indicar condiciones médicas o comportamientos atípicos que requieren una atención y análisis profundos desde una perspectiva estadística. Por lo tanto, se vuelve fundamental implementar un proceso de control de calidad riguroso que permita una depuración exitosa de los datos. Esto, a su vez, facilita la construcción de modelos normativos realmente valiosos en este estudio, garantizando la fiabilidad y precisión de los resultados obtenidos.

En esta línea, se propone implementar una curva normativa para estandarizar los volúmenes de los órganos abdominales, lo que es un recurso valioso para el diagnóstico inicial de los profesionales de la salud. No obstante, es fundamental considerar otras variables relevantes. Estas pueden abarcar desde la variabilidad individual en la anatomía y la función de los órganos, hasta la influencia de factores como la edad, el sexo y el estado de salud previo.

En este contexto, es importante destacar que la construcción de esta curva se basa en datos provenientes de sujetos con estructuras abdominales saludables. Para garantizar la precisión y fiabilidad de los resultados, es esencial asegurar la integridad de los datos recopilados por medio de un control de calidad riguroso, que permitirá en caso de que se detectan posibles inconsistencias, una depuración minuciosa para asegurar la coherencia y replicabilidad de los hallazgos obtenidos.

---

Estos modelos normativos son herramientas cruciales para los profesionales de la salud, ya que facilitan la toma de decisiones clínicas informadas, mejorando el diagnóstico, tratamiento y seguimiento de los pacientes. Además, son fundamentales en la investigación científica, ya que sirven de base para estudios, investigaciones y desarrollo de tecnologías, al dar un marco de comprensión de lo considerado dentro de la variabilidad normal en una población determinada.

Esto subraya la importancia de realizar mediciones de volúmenes con mayor precisión a través de la tomografía computarizada (TC), dada su amplia utilización en la actualidad como una de las principales técnicas de diagnóstico por imágenes. Su fácil acceso y disponibilidad la convierten en una opción viable. Además, numerosos estudios respaldan su eficacia como una herramienta destacada para la visualización detallada de las estructuras abdominales [11], permitiendo ver a los médicos y al equipo involucrado en la investigación inconsistencias dentro de los resultados de la segmentación.

Se optó por tomografías de abdomen contrastadas, ya que la creación de un modelo de segmentación basado en este tipo de imágenes ofrece una visualización más nítida de las estructuras anatómicas y posibles anomalías gracias al contraste, facilitando la detección de lesiones y áreas anómalas por el algoritmo de inteligencia artificial [12]. La automatización de la segmentación con este tipo de imágenes agiliza el proceso diagnóstico, posibilitando una evaluación más rápida y precisa, lo que puede traducirse en una atención médica más eficaz y oportuna. Además, estas tomografías permiten realizar mediciones precisas de los volúmenes de los órganos, lo que a su vez posibilita la personalización del tratamiento para cada paciente, sobre todo en intervenciones quirúrgicas y tratamientos que demandan una comprensión detallada de la anatomía abdominal.

---

## **IV. OBJETIVOS**

### ***A. Objetivo general***

Implementar un flujo de trabajo para la evaluación de segmentación de imágenes de tomografía de abdomen en población SURA.

### ***B. Objetivos específicos***

- Identificar y seleccionar una muestra de sujetos sanos con estudios de tomografía computacional de abdomen utilizando el repositorio de imágenes médicas de SURA.
- Desarrollar un proceso de control de calidad de la segmentación de las imágenes de abdomen.
- Realizar un control de calidad automático en las imágenes y posterior filtración de los datos, para la construcción de un modelo normativo robusto.



---

## **V. MARCO TEÓRICO**

### ***Tomografía Computarizada***

La tomografía computarizada (TC), es una técnica de imagen médica de amplio uso en la actualidad, principalmente por su utilidad al momento de obtener imágenes detalladas y tridimensionales del interior del cuerpo humano [13], convirtiéndose en una herramienta fundamental en la medicina diagnóstica y el monitoreo de enfermedades.

Su funcionamiento se basa en el uso de rayos X y un ordenador para crear imágenes transversales de una o varias zonas del cuerpo, donde un tubo de rayos X gira alrededor del paciente emitiendo haces de rayos X en diferentes ángulos, mientras unos detectores miden la cantidad de rayos X que atraviesan el cuerpo y envían esta información a una computadora, que la procesa para crear las imágenes [6].

A diferencia de las radiografías convencionales, que proporcionan imágenes bidimensionales, la tomografía computarizada (TC) genera imágenes tridimensionales que permiten una visualización detallada de las estructuras internas desde diversos ángulos. Entre sus ventajas, destaca su rapidez y precisión, siendo una herramienta útil para obtener volúmenes de distintos órganos [14].

Por otro lado, es esencial abordar el concepto de series en imágenes de tomografía computarizada (TC), las cuales constan de varios cortes o rebanadas almacenadas como archivos de imágenes digitales. Estos archivos se presentan de manera secuencial y, al visualizarse en conjunto, proporcionan una representación tridimensional del área o estructura anatómica en estudio. En este punto, surge el concepto de fases, que se refiere a distintos momentos en el tiempo durante la adquisición de imágenes. Estas fases son particulares en estudios específicos donde se busca visualizar un proceso en evolución o capturar detalles en momentos específicos.

En este contexto, nos enfocamos en la identificación de series, que son secuencias de imágenes obtenidas durante un procedimiento de tomografía computarizada (TC). La adquisición de imágenes en una TC implica capturar múltiples cortes o secciones del cuerpo mientras el escáner de CT gira alrededor del paciente. Cada rotación del escáner de tomografía computarizada (TC) produce una "imagen

---

axial" o "corte". Estos cortes representan secciones transversales del área específica del cuerpo bajo examen y se obtienen en intervalos regulares a lo largo del cuerpo, generando múltiples imágenes consecutivas. Comúnmente, la colección de estos cortes adquiridos durante el escaneo se denomina "serie".

Las series en tomografía computarizada pueden variar según la técnica y el propósito del escaneo. Por ejemplo, una serie de TC abdominal puede incluir imágenes axiales a lo largo de esta área, mientras que una serie de TC cerebral mostrará imágenes axiales del cerebro. Además, la configuración específica de las series de TC puede variar según el protocolo utilizado por el médico o el técnico radiólogo para obtener la información necesaria sobre la condición o área de interés del paciente. En el contexto de nuestro proyecto, se priorizaron las series volumétricas axiales en fase portal para la posterior creación del modelo de segmentación en Monai.

La rapidez de los tomógrafos permite la realización de exámenes dinámicos trifásicos. En otras palabras, es posible obtener imágenes en la fase arterial precoz, en la fase tardía (también conocida como fase portal) y en una fase parenquimatosa. Este enfoque dinámico proporciona una información detallada y secuencial que resulta fundamental en la interpretación clínica y el diagnóstico preciso [15].

- **Fase arterial:** Antes de la fase arterial, al paciente se le administra un medio de contraste yodado a través de una vena. Este agente de contraste contiene yodo, que absorbe los rayos X y mejora la visibilidad de los vasos sanguíneos. El contraste inyectado realza las arterias, haciendo que se vuelvan más visibles en las imágenes de TC. Esto permite evaluar la perfusión arterial y detectar posibles anomalías, como estenosis, aneurismas u otras patologías vasculares. Es útil para estudios de trasplante, quimioembolización o estudios angiográficos. La fase arterial precoz será de utilidad para detectar anomalías vasculares, como por ej. las fístulas arterioportales. La fase arterial tardía permite diagnosticar hepatomas hipervasculares y las reconstrucciones multiplanares facilitan su comprensión espacial [15].

- **Fase portal:** Durante esta fase, el agente de contraste administrado previamente se ha distribuido más ampliamente en el sistema venoso portal del

---

abdomen, lo que permite evaluar la circulación sanguínea en las venas hepáticas y otras estructuras asociadas, lo que la convierte en la fase ideal para la obtención de las diferentes series que se utilizaron en el proyecto [16].

- **Fase parenquimatosa:** La fase parenquimatosa es una etapa específica durante la adquisición de imágenes que se centra en la visualización detallada del parénquima de los órganos. Esta fase sigue a la fase portal en estudios abdominales y está destinada a evaluar la estructura y las características internas de los órganos, como el hígado, muestra mucha utilidad para la detección de metástasis hipovasculares [15].

### **Casos clínicos**

La consideración del volumen de órganos desempeña un papel crucial en el diagnóstico, tratamiento y seguimiento de diversas condiciones médicas lo que justifica la necesidad de la tecnología de imágenes médicas y el software de procesamiento de imágenes como herramientas clave para realizar estas mediciones con la precisión necesaria, algunos casos clínicos son:

- **Oncología:** En el diagnóstico y tratamiento del cáncer, es fundamental medir el volumen de tumores y órganos afectados. Esto ayuda a determinar el estado del cáncer, planificar tratamientos como la radioterapia o la cirugía, y evaluar la respuesta a las terapias [16].

- **Cirugía de trasplante:** En el trasplante de órganos, como el corazón, el hígado o los riñones, se requiere medir con precisión el volumen del órgano donante y del órgano receptor para asegurar un ajuste adecuado y prevenir complicaciones [17]

- 
- **Cardiología:** En enfermedades cardíacas, la evaluación del volumen ventricular es crucial para determinar la función cardíaca, calcular el gasto cardíaco y diagnosticar afecciones como la insuficiencia cardíaca [9].
  - **Neurología:** En el estudio de trastornos neurológicos, como tumores cerebrales o aneurismas, se mide el volumen intracraneal para evaluar la presión intracraneal y planificar cirugías [18].
  - **Nefrología:** Para evaluar la función renal, se miden los volúmenes renales y se calcula el filtrado glomerular. Esto es esencial para diagnosticar enfermedades renales y ajustar la dosis de medicamentos en pacientes con disfunción renal [19].
  - **Gastroenterología:** En trastornos como la enfermedad inflamatoria intestinal (EII) o la cirrosis hepática, se mide el volumen del intestino o el hígado para evaluar el grado de daño [20].
  - **Ortopedia:** En el diseño de prótesis y en la planificación de cirugías ortopédicas, se mide el volumen de huesos y articulaciones para garantizar un ajuste adecuado [21].
  - **Estudios de densidad ósea:** Para evaluar la salud ósea en condiciones como la osteoporosis, se mide el volumen y la densidad ósea para calcular el índice T-score [22].

### ***Segmentación en imágenes***

La segmentación en imágenes médicas se lleva a cabo mediante diversos métodos, cada uno con enfoques específicos para identificar y delinear áreas de interés dentro de las imágenes. Algunas de las técnicas más comunes incluyen:

---

- **Segmentación basada en umbrales (Thresholding):** Este método implica asignar píxeles a una región específica basándose en un umbral predefinido de intensidad. Los píxeles por encima o debajo del umbral se asignan a diferentes regiones, siendo útil cuando las estructuras de interés presentan una clara diferencia de intensidad con respecto al fondo [23].

- **Métodos de crecimiento de regiones (Region Growing):** Comienza con un conjunto de píxeles semilla y expande iterativamente la región conectando píxeles adyacentes con similitud en intensidad. Se emplea para segmentar áreas con patrones homogéneos de intensidad [23].

- **Contornos activos (Active Contour Models o Snakes):** Son contornos deformables que se ajustan a los bordes de las estructuras utilizando fuerzas internas y externas. Útiles para segmentar estructuras con bordes no definidos o deformaciones [24].

- **Segmentación por regiones activas (Active Region Models):** Similar a los contornos activos, pero enfocado en segmentar áreas en la imagen en lugar de un contorno. Combina métodos de contornos activos con técnicas de segmentación por crecimiento de regiones [25].

- **Métodos de aprendizaje automático (Machine Learning):** Utilizan algoritmos, como redes neuronales convolucionales (CNN) o clustering, para aprender patrones y segmentar imágenes de forma adaptativa [26].

Los médicos radiólogos son expertos en interpretar imágenes médicas y tienen un sólido conocimiento en anatomía, fisiología, patología y tecnología de imágenes. Su experiencia les permite evaluar la precisión de la segmentación y realizar ajustes necesarios.

---

En la segmentación manual con contornos activos, los radiólogos intervienen directamente delineando bordes anatómicos en las imágenes. Esta técnica manual permite un control preciso, aunque puede ser laboriosa y consumir tiempo.

El rápido desarrollo de algoritmos de inteligencia artificial (IA) ha facilitado procesos que antes parecían demasiado complejos como por ejemplo el aprendizaje automático, logrando el desarrollo de algoritmos que pueden aprender de datos y mejorar su rendimiento con el tiempo, uno de estos casos ocurre con las redes neuronales artificiales que son modelos de aprendizaje automático inspirados en la estructura y funcionamiento del cerebro humano [27]. Estas redes están compuestas por capas de nodos interconectados que procesan información y pueden utilizarse para tareas como clasificación de imágenes convirtiéndola en una herramienta valiosa para procesar imágenes de forma automática ayudando a tareas tales como la identificación y segmentación de los diferentes órganos [28].

MONAI es un marco de trabajo de código abierto [29], diseñado específicamente para procesar imágenes médicas y la investigación de aprendizaje profundo en el dominio médico [30].

La integración de MONAI (Medical Open Network for AI) y XNAT (XNAT Imaging Platform) es así un enfoque valioso para aprovechar las capacidades de procesamiento de imágenes médicas de MONAI [31], además de la gestión de datos de imágenes de XNAT en aplicaciones de investigación médica.

En el proceso de segmentar automáticamente órganos y tejidos en las imágenes de TC, se ha podido facilitar la identificación y medición de estructuras específicas permitiendo así la detección de anomalías como tumores, lesiones o calcificaciones proporcionando diagnósticos asistidos por computadora a partir de la inteligencia artificial, donde los médicos reciben recomendaciones basadas en análisis previos de casos similares, pero además la inteligencia artificial se utiliza para predecir resultados clínicos, como el riesgo de progresión de enfermedades o la eficacia de tratamientos [32].

---

Es importante resaltar que parte fundamental para aplicar estos algoritmos de inteligencia artificial es la recolección de datos donde se requieren muchos datos etiquetados para entrenar modelos de IA de manera efectiva [33], lo que puede ser costoso y requerir tiempo. Además, como se dijo antes, la segmentación perfecta puede ser difícil de lograr y es allí donde hay que tomar medidas y consideraciones para conformar un modelo normativo.

### ***UNETR***

Para este proyecto se utilizó UNETR que es un modelo de red neuronal utilizado para la segmentación de imágenes médicas, especialmente en el ámbito de la visión por computadora y el procesamiento de imágenes médicas. Combina dos arquitecturas populares en el campo del aprendizaje profundo: la red U-Net y los Transformers [34].

La arquitectura U-Net es conocida por su eficiencia en la segmentación de imágenes. Se compone de una estructura en forma de U que consta de un codificador (reducción de la resolución de la imagen y extracción de características) y un decodificador que restaura la resolución original y ayuda en la segmentación [35].

Por otro lado, los Transformers son modelos efectivos en tareas de procesamiento de lenguaje natural y se han adaptado con éxito para aplicaciones de visión por computadora. Estos modelos son conocidos por capturar relaciones a larga distancia en secuencias y modelar contextos complejos [34].

### ***XNAT***

XNAT es una plataforma versátil y poderosa utilizada en entornos de investigación médica para almacenar, analizar y colaborar en datos de imágenes médicas. Ofrece una amplia gama de herramientas y funcionalidades para gestionar eficientemente datos complejos en el campo de la neurociencia y otras disciplinas de la salud [36]. Además, se adhiere a varios estándares de la industria, como el formato DICOM, lo que facilita la compatibilidad con sistemas y datos médicos ya existentes.

Esta plataforma se emplea para el almacenamiento, organización y compartición de datos de imágenes médicas como resonancias magnéticas (RM), tomografías computarizadas (TC), imágenes PET y otros tipos. Esto permite una gestión eficiente de conjuntos de datos complejos y grandes volúmenes de imágenes, convirtiéndose en un software seguro para el almacenamiento y acceso a datos médicos. Proporciona herramientas para controlar los niveles de acceso, garantizando así la privacidad y la seguridad de los datos del paciente [37].

XNAT cuenta con una interfaz de usuario intuitiva que permite a los usuarios cargar, visualizar, analizar y descargar datos de imágenes médicas. Facilita enormemente el proceso de importación de datos que alimentarán nuestro modelo normativo al automatizar procesos mediante algoritmos de procesamiento de datos directamente en la plataforma. Esta herramienta es altamente personalizable y extensible, lo que agiliza la integración de módulos y funcionalidades personalizadas según las necesidades del usuario o del proyecto. Esto permite la integración con varias herramientas y programas de análisis de imágenes médicas, facilitando el análisis de datos de imágenes utilizando diferentes algoritmos y técnicas.

Es importante destacar que esta plataforma ha facilitado la colaboración entre los investigadores implicados en el proyecto al permitir el intercambio de datos y el acceso compartido a recursos. Esta función ha fomentado la colaboración y la investigación multidisciplinaria.

### ***Quality check(QC)***

- **Matriz de transformación:** Es una herramienta esencial en geometría y gráficos por computadora, utilizada para describir transformaciones geométricas en un espacio tridimensional. Esta matriz representa diferentes tipos de transformaciones, como rotaciones, traslaciones, escalado, reflexiones y proyecciones en un sistema de coordenadas 3D. Al aplicarla a un objeto tridimensional, se multiplican las coordenadas del objeto (presentadas como un vector) por la matriz, lo que produce nuevas



---

coordenadas transformadas del objeto. Este proceso posibilita diversas operaciones y manipulaciones geométricas en el espacio 3D, siendo esencial en campos como gráficos por computadora, diseño asistido por computadora (CAD), renderizado 3D y animación.

- **Step:** Durante el procesamiento de imágenes de distintas dimensiones, se requiere estandarizar un algoritmo para visualizar efectivamente los órganos de interés en todas las imágenes. Dentro de la matriz de transformación, el "step" se relaciona con la escala o el factor de escala en el eje Z (profundidad) de la imagen. Controla el cambio de escala en esa dimensión, particularmente si la transformación implica escalado en dicha dimensión. En el contexto de imágenes médicas, como resonancia magnética (RM) o tomografía computarizada (TC), el "step" representa el intervalo entre cortes o imágenes adyacentes en un conjunto, indicando la separación o distancia entre cada corte o imagen en el conjunto de slices.

- **convención radiológica:** se refiere a las normas y convenciones utilizadas en el campo de la radiología para interpretar, informar y comunicar los hallazgos radiológicos [38]. Estas normas son cruciales para garantizar una comunicación efectiva y uniforme entre radiólogos, profesionales de la salud y especialistas. Una característica esencial de esta convención radica en la presentación de imágenes médicas en una orientación anatómica estándar, donde las estructuras se visualizan como si el observador estuviera mirando al paciente desde el pie de la cama hacia la cabeza. Esto significa que el lado derecho del paciente se visualiza en el izquierdo de la imagen y, a su vez, el izquierdo se representa en el derecho de la imagen, como se observa en la Fig. 3.

### ***Modelos normativos***

Los modelos normativos son una herramienta útil para predecir y evaluar las medidas de un individuo en relación con una población. Proporcionan así una estimación de la totalidad o parte de una distribución de probabilidad condicional para

---

una población de referencia de individuos normales o en nuestro caso específico sujetos sanos, permitiendo así identificar como determinado individuo podría desviarse de esta población de referencia. Esta distribución puede tener cualquier forma, usualmente con los sujetos sanos ubicados hacia el centro, donde cuanto más alejado se encuentre un sujeto de esta distribución menor será la probabilidad de pertenecer al grupo de sujetos sanos [38].

En investigaciones previas de este grupo, se ha constatado que el cambio en el volumen cerebral es un proceso natural a medida que las personas envejecen [39]. No obstante, diversas enfermedades neurodegenerativas pueden inducir alteraciones que se distinguen de los cambios típicos observados en individuos sanos. Una de las técnicas usadas para diagnosticar trastornos de la estructura y morfología cerebral es la resonancia magnética, donde se evalúa visualmente para discernir si las modificaciones cerebrales son atribuibles al envejecimiento o a desórdenes neurodegenerativos [38].

Por otro lado, está la tomografía de abdomen, que, aunque pretende evaluar los órganos y estructuras abdominales en lugar del cerebro, comparte similitudes en la adquisición de imágenes médicas y el análisis cuantitativo. En este contexto, la tomografía abdominal también puede requerir la comparación con modelos normativos para detectar posibles anomalías o enfermedades.

Para crear el modelo normativo, es esencial tener la materia prima en forma de datos, que consisten principalmente en tomografías de abdomen realizadas a individuos asegurados por SURA. Idealmente, estos individuos deberían estar libres de patologías abdominales, ya que la finalidad es establecer estándares volumétricos para el abdomen que posteriormente servirán como base para construir las curvas normativas de los distintos órganos abdominales.

Una vez que se han recopilado los datos, se procede a realizar un análisis minucioso para identificar y descartar aquellos sujetos que presenten valores atípicos en múltiples ocasiones, ya que la presencia de estos valores atípicos podría afectar negativamente la calidad de las curvas normativas [39].

---

### ***Valores atípicos (Outliers)***

Los outliers, también conocidos como valores atípicos, son puntos de datos que se desvían significativamente del resto de los datos en un conjunto. Estos valores se suponen inusualmente altos o bajos en comparación con los demás datos y pueden distorsionar el análisis estadístico y los modelos predictivos si no se manejan correctamente [40].

Los outliers pueden surgir por diversas razones, como errores de medición, errores de entrada de datos, eventos raros o inesperados en el fenómeno que se está estudiando, o simplemente debido a la variabilidad natural en los datos. Es importante identificar y manejar adecuadamente los outliers para evitar conclusiones erróneas o sesgadas en el análisis de datos. Dependiendo del contexto del problema, los outliers pueden ser eliminados, tratados o investigados más a fondo para determinar su validez y relevancia [40].

Para su detección se utilizan técnicas que básicamente son métodos utilizados para identificar observaciones atípicas o inusuales en un conjunto de datos. Estas técnicas pueden variar en complejidad y enfoque, pero generalmente se clasifican en métodos estadísticos, basados en distancias, modelos probabilísticos, y enfoques de aprendizaje automático, entre otros.

Algunas de las técnicas comunes de detección de outliers que se incluyeron en el trabajo fueron:

- **Z-score:** El Z-score, también conocido como puntaje estándar, es una medida estadística que proporciona información sobre la posición relativa de un valor dentro de un conjunto de datos en relación con la media y la desviación estándar de ese conjunto. Se calcula restando la media del valor y dividiendo el resultado por la desviación estándar [40]. El método del Z-score utiliza estos valores para calcular un puntaje estandarizado, que indica cuántas desviaciones estándar un valor particular está por encima o por debajo de la media.

El Z-score nos ayuda a comprender la posición relativa de un valor dentro de un conjunto de datos y a identificar valores atípicos (outliers) que se encuentren significativamente lejos de la media. Un Z-score alto o bajo indica que el valor está relativamente lejos de la media, mientras que un Z-score cercano a cero indica que el valor está cerca de la media. Los valores con Z-scores por encima de cierto umbral pueden considerarse outliers y pueden requerir un análisis adicional para comprender su naturaleza y posible impacto en los datos.

$$|Z| = \left| \frac{X - \mu}{\sigma} \right|$$

*Ecuación 1. Cálculo del valor Z para la detección de outliers*

Donde Z es el Z-score, X es el valor individual,  $\mu$  es la media del conjunto de datos y  $\sigma$  es la desviación estándar del conjunto de datos.

- **Rango Intercuartílico (IQR):** El rango intercuartílico (IQR) es una medida estadística que describe la dispersión de los datos en un conjunto. Se calcula como la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1) en un conjunto de datos ordenado. El IQR es menos sensible a valores extremos que otras medidas de dispersión, como la desviación estándar o el rango, lo que lo convierte en una medida robusta. El IQR proporciona información sobre cómo se distribuyen los datos alrededor de la mediana y se utiliza ampliamente en análisis estadísticos y en la detección de outliers [41].

$$(X < Q1 - \text{umbral} \times \text{IQR}) \mid (X > Q3 + \text{umbral} \times \text{IQR})$$

*Ecuación 2. Determinación de outliers con el método de IQR*

Donde X representa el valor individual, Q1 es el primer cuartil, que divide el conjunto de datos en dos partes iguales, dejando fuera el 25% más bajo de los datos,

Q3 es el tercer cuartil, que divide el conjunto de datos en dos partes iguales, dejando fuera el 25% más alto de los datos, IQR es el rango intercuartílico, que es la diferencia entre Q1 y Q3 y finalmente el umbral es el valor que se multiplica por IQR y que se ingresará dentro de los parámetros de la función para determinar el límite superior e inferior para identificar outliers. Si un valor está por debajo de  $Q1 - \text{umbral} \times \text{IQR}$  o por encima de  $Q3 + \text{umbral} \times \text{IQR}$ , se clasifica como un outlier.

- **Método de Tukey:** es una técnica utilizada en estadística para identificar valores atípicos en un conjunto de datos. Fue desarrollado por John Tukey y se basa en el rango intercuartílico (IQR) de los datos [42]. La prueba de Tukey (1977) para detectar valores atípicos implica establecer márgenes alrededor del primer y tercer cuartil de la distribución de datos históricos. Para determinar estos cuartiles, se considera un rango de 1.5 veces el rango intercuartílico a la derecha y a la izquierda de los cuartiles 1 y 3. Específicamente, se calculan los cuartiles correspondientes (Q3 y Q1), luego se determina el rango intercuartílico (IQR) como la diferencia entre estos dos cuartiles ( $\text{IQR} = Q3 - Q1$ ). Posteriormente, se establece un rango para identificar datos atípicos, aquellos que exceden el valor de  $Q3 + 1.5(\text{IQR})$ . En otras palabras, cualquier dato que se encuentre más allá de este margen se considera atípico, según el criterio de Tukey [43].

$$\text{lower\_bound} = Q1 - 1.5 \times \text{IQR}$$

Ecuación 3. Cálculo del límite inferior en el método de tukey

$$\text{upper\_bound} = Q3 + 1.5 \times \text{IQR}$$

Ecuación 4. Cálculo del límite superior en el método de tukey

$$\text{outliers} = (X < \text{lower\_bound}) \mid (X > \text{upper\_bound})$$

Ecuación 5. Determinación de los outliers con el método de tukey

---

Donde  $X$  representa el valor individual,  $Q1$  es el primer cuartil,  $Q3$  es el tercer cuartil y finalmente  $IQR$  es el rango intercuartílico.

- **Z- score modificado:** El z-score modificado, también conocido como MAD (Median Absolute Deviation), es una medida de cuántas desviaciones estándar un punto de datos está por encima o por debajo de la media de un conjunto de datos, pero a diferencia del z-score estándar, el z-score modificado utiliza una medida de dispersión más robusta, como la mediana de los datos, en lugar de la media, El z-score modificado es útil para identificar valores atípicos en conjuntos de datos que pueden contener valores extremos o distribuciones no normales. Al igual que el z-score estándar, los valores con un z-score modificado mayor que cierto umbral se consideran atípicos y pueden requerir un análisis adicional [44].

$$\text{med\_abs\_dev} = \text{median} (|X_i - \text{median}(X)|)$$

*Ecuación 6. Cálculo de la desviación absoluta de la mediana.*

$$\text{modified\_z\_scores} = 0.6745 \times \left( \frac{X_i - \text{median}}{\text{med\_abs\_dev}} \right)$$

*Ecuación 7. Cálculo del Z-score a partir de la mediana*

$$\text{outliers} = |\text{modified\_z\_scores}| > \text{umbral}$$

*Ecuación 8. Determinación de los outliers con MAD*

Donde  $X_i$  representa cada valor individual en el conjunto de datos,  $\text{median}$  es la mediana del conjunto de datos y 0,6745 es una constante encontrada en diversas fuentes literarias [44].

---

### ***Control de calidad de los datos***

El control de calidad de los datos es una etapa crítica en el procesamiento de imágenes médicas, ya que garantiza la confiabilidad y la precisión de los resultados obtenidos. El control de calidad aborda la identificación y corrección de problemas comunes que pueden surgir durante la adquisición o el procesamiento de los datos.

Previo a la detección y eliminación de datos atípicos, una revisión visual de las imágenes, para identificar problemas evidentes, como cortes incompletos o artefactos visuales [45], es de vital importancia. Así mismo, evaluar la calidad de la segmentación con ayuda de radiólogos especialistas, permitirá un mejor filtrado de nuestra población de muestra, evitando errores de etiquetado.

Así como, con la evolución de las soluciones de archivo de datos de investigación, se enfrenta el desafío de integrar datos distribuidos. Esto se complica debido a la necesidad de establecer acuerdos claros para manejar diferentes modelos de datos [46]. Su implementación puede presentar desafíos relacionados con la escalabilidad, el sesgo en el dominio de aplicación y la pérdida de acceso a datos detallados haciendo necesario controlar la calidad de los datos obtenidos.

### ***Modelos de regresión***

Los modelos de regresión son herramientas estadísticas utilizadas para investigar y modelar la relación entre una variable dependiente (también llamada variable de respuesta o variable explicada) y una o más variables independientes (también conocidas como variables predictoras o variables explicativas). Estos modelos se utilizan para entender cómo cambia la variable dependiente cuando se modifican las variables independientes [47].

La idea fundamental detrás de los modelos de regresión es ajustar una función matemática a los datos observados que describa la relación entre las variables de interés. Esta función puede ser lineal o no lineal, y puede ser elegida según las características de los datos y el tipo de relación que se espera entre las variables [47].

---

Los modelos de regresión se utilizan en una amplia variedad de campos, desde la economía y la sociología hasta la biología y la ingeniería. Algunos ejemplos comunes de modelos de regresión incluyen la regresión cuantil, la regresión centil, suavizado de diagrama de dispersión estimado localmente (LOESS), procesos gaussianos (GP), entre otros.

- **Regresión cuantil:** La regresión cuantílica, también conocida como regresión cuantil o regresión cuantil a cuantil, es un método estadístico utilizado para modelar la relación entre variables en un conjunto de datos. A diferencia de la regresión lineal tradicional que se enfoca en modelar la media condicional de la variable dependiente dado un conjunto de variables independientes, la regresión cuantílica modela diferentes cuantiles de la variable dependiente.

En la regresión cuantílica, se ajusta un modelo para cada cuantil de interés de la variable dependiente. Esto permite capturar las relaciones no lineales y heterocedasticidad en los datos, lo que puede ser útil cuando la relación entre las variables no es constante en todos los niveles de la variable dependiente [48]. Los cuantiles más comúnmente utilizados son el cuantil 0.25 (primer cuartil), el cuantil 0.50 (mediana) y el cuantil 0.75 (tercer cuartil), pero se pueden modelar otros cuantiles según la necesidad del análisis.

El objetivo de la regresión cuantílica es estimar los coeficientes del modelo que minimizan la suma ponderada de las diferencias absolutas entre los valores observados y predichos en cada cuantil [48]. Esto proporciona una descripción completa de la distribución condicional de la variable dependiente y permite capturar la variabilidad en diferentes partes de la distribución. La regresión cuantílica es especialmente útil en aplicaciones donde se desea examinar cómo cambian las relaciones entre variables en diferentes puntos de la distribución de la variable dependiente.

- **Regresión centil:** La regresión centil, también conocida como regresión centil a centil, es un método estadístico similar a la regresión cuantílica, pero en lugar de modelar diferentes cuantiles de la variable dependiente, se enfoca en modelar



---

puntos específicos de la distribución, generalmente centiles específicos. Al igual que en la regresión cuantílica, en la regresión centil se ajusta un modelo para cada punto de la distribución de interés. Esto permite capturar cómo cambian las relaciones entre las variables independientes y la variable dependiente en diferentes partes de la distribución de la variable dependiente.

- **Suavizado de diagrama de dispersión estimado localmente (LOESS):** es un método de regresión no paramétrica que se utiliza para suavizar una serie de datos mediante el ajuste de curvas locales [49]. En lugar de ajustar una única línea de regresión a través de todos los puntos de datos, LOESS ajusta múltiples polinomios de baja orden a subconjuntos de datos adyacentes, lo que permite capturar patrones complejos en los datos sin asumir una forma específica para la relación entre las variables. Este método se basa en un enfoque de ventana móvil, donde se ajusta un polinomio a los puntos de datos dentro de una ventana alrededor de cada punto, y luego se suavizan estos polinomios para obtener una estimación suave de la relación subyacente entre las variables. LOESS es útil para explorar relaciones no lineales en los datos y para suavizar series de tiempo u otras estructuras de datos con tendencias o patrones irregulares [49].

- **Procesos gaussianos(GP):** Los Procesos Gaussianos (GP) son un poderoso marco probabilístico utilizado en aprendizaje automático y estadísticas para modelar funciones y distribuciones de datos. Un GP es una generalización de la distribución normal multivariante a un número infinito de variables. Se caracteriza por una media y una matriz de covarianza, que describen completamente la distribución de cualquier conjunto de variables aleatorias [50]. En el contexto del modelado predictivo, un GP se utiliza para modelar una función desconocida que mapea las entradas a las salidas, donde se asume que cualquier conjunto finito de salidas sigue una distribución gaussiana condicionada a las entradas correspondientes.

Los GP son flexibles y pueden adaptarse a una amplia variedad de problemas de modelado. Permiten capturar incertidumbre en las predicciones y proporcionan una

---

forma natural de realizar interpolación y extrapolación. Además, los GP son altamente adaptables a diferentes tipos de datos y estructuras, lo que los hace útiles en una amplia gama de aplicaciones, desde la regresión hasta la clasificación y el aprendizaje por refuerzo [51].

### ***Validación cruzada***

La validación cruzada es una técnica utilizada en el aprendizaje automático y la estadística para evaluar el rendimiento de un modelo predictivo. Consiste en dividir el conjunto de datos en varios subconjuntos, denominados "folds" o "pliegues", y luego realizar múltiples iteraciones de entrenamiento y evaluación del modelo [52].

En cada iteración, se entrena el modelo en una combinación de los pliegues y se evalúa en el pliegue restante que no se utilizó para el entrenamiento. Este proceso se repite varias veces, de manera que cada pliegue sirve como conjunto de datos de prueba en al menos una ocasión [52].

La validación cruzada proporciona una estimación más robusta del rendimiento del modelo al promediar los resultados de todas las iteraciones [52]. Ayuda a detectar problemas de sobreajuste o subajuste al evaluar cómo se comporta el modelo con datos que no ha visto durante el entrenamiento [53].

### ***Métricas***

Las métricas para evaluar el rendimiento de modelos de regresión son medidas cuantitativas que se utilizan para determinar qué tan bien se ajusta un modelo a los datos observados. Estas métricas proporcionan una evaluación objetiva de la capacidad predictiva del modelo y permiten comparar diferentes modelos entre sí [54].

Algunas de las métricas más comunes para evaluar modelos de regresión que se incluyeron en el trabajo fueron:

---

- **Error cuadrático medio (RMSE):** El RMSE calcula la media de los cuadrados de las diferencias entre las predicciones del modelo y los valores reales. Es una medida de la dispersión de los errores y penaliza los errores más grandes de manera cuadrática [54].

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2}$$

*Ecuación 9. Cálculo del error cuadrático medio*

Donde  $n$  es el número total de observaciones en el conjunto de datos,  $y_i$  son los valores reales de la variable de respuesta para la  $i$ -ésima observación y finalmente  $\hat{y}_i$  son las predicciones del modelo para la  $i$ -ésima observación.

- **Error Cuadrático Medio Estandarizado (SMSE):** Es una métrica de evaluación que se utiliza para comparar el error de un modelo de regresión con la variabilidad de los datos reales. Esta métrica normaliza el Error Cuadrático Medio (RMSE) dividiéndolo por la desviación estándar de los valores reales en el conjunto de datos de prueba [54].

$$\text{SMSE} = \frac{\text{RMSE}}{\text{SD}(y_{\text{test}})}$$

*Ecuación 10. Cálculo del error cuadrático medio estandarizado.*

### ***Delta de edad***

Se trata de un modelo normativo que evalúa la diferencia entre la edad predicha y la edad real de un individuo. Esta discrepancia puede revelar información importante sobre el ritmo de envejecimiento, ya sea acelerado o retardado, en comparación con la población en general o con un grupo de referencia específico [39]. La predicción de la

---

edad de un sujeto puede ser a partir de variables tales como los volúmenes de diferentes órganos [39], generalmente en este tipo de algoritmos de machine learning se utilizan imágenes médicas como tomografías computarizadas (TC) o resonancias magnéticas (RM). De esta manera, el algoritmo determina la discrepancia entre los volúmenes observados y los volúmenes esperados para una determinada edad. Este tipo de análisis resulta útil en campos como la medicina preventiva, la geriatría y la investigación sobre el envejecimiento [39].

$$\Delta\text{Edad} = \text{Edad real} - \text{Edad predicha}$$

*Ecuación 11. Cálculo de delta de edad*

## VI.METODOLOGÍA

- **Revisión de la literatura - implementación de modelos normativos en volumetría de órganos del cuerpo a partir de imágenes médicas**

En la depuración de datos, se buscó en diversas bases bibliográficas disponibles en el repositorio de la Universidad de Antioquia. Entre estas fuentes, se exploró Scopus, reconocida como una de las mayores bases de datos bibliográficas multidisciplinarias y una destacada herramienta para el análisis de literatura científica y académica.

Durante la fase inicial del proyecto, se realizó una exhaustiva revisión del estado del arte en modelos normativos de volumetría corporal con el objetivo de enriquecer la metodología propuesta. En base a este análisis, se elaboró un plan detallado (Fig. 1) que delineó un flujo de trabajo preciso para la recopilación de estudios de pacientes sanos, fundamental para la construcción de la base de datos requerida. Se estableció un rango de edad óptimo para la población adulta, abarcando desde los 18 hasta los 80 años.

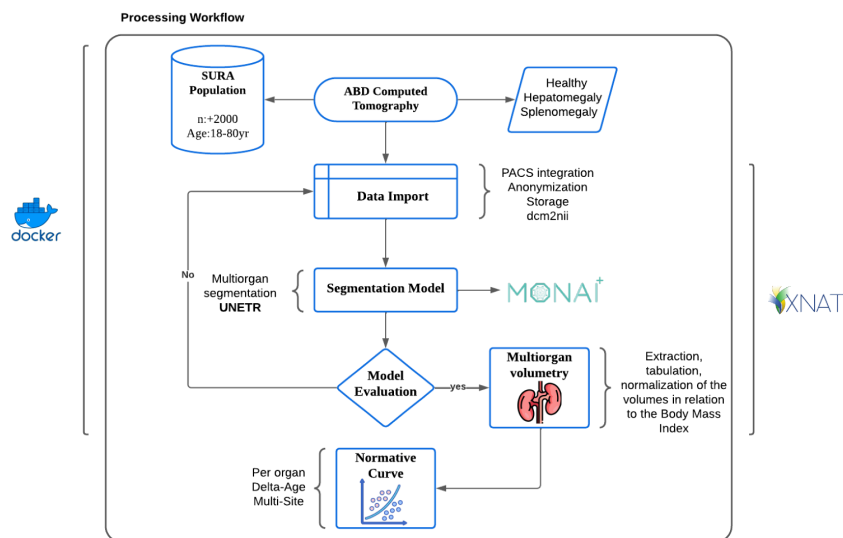


Fig. 1 Flujo de trabajo diseñado por el grupo de investigación.

Posteriormente, se optó por la aplicación de un algoritmo de segmentación automática basado en inteligencia artificial, previamente entrenado por nuestro grupo de investigación. Este algoritmo se integró de manera exitosa en nuestra plataforma de procesamiento XNAT, lo que permitió la importación directa de estudios desde los PACS mediante herramientas como Elastic Search y Findscu. Estos estudios se convirtieron al formato NIfTI para llevar a cabo la segmentación utilizando el algoritmo de IA implementado.

A partir de esta segmentación, surgió la idea de extraer, tabular y normalizar los volúmenes obtenidos. Conscientes de las posibles variaciones entre las características de cada sujeto y con el objetivo de evitar sesgos, se optó por normalizar los volúmenes utilizando el Índice de Masa Corporal (IMC). Esta medida fue seleccionada debido a su disponibilidad en la teradata de la organización, lo que facilitaba su recuperación. Medidas como por ejemplo el perímetro abdominal requerirían obtenerse a partir de las imágenes, lo que añadiría una capa de complejidad adicional, considerando que las imágenes adquiridas no tienen el mismo tamaño. Finalmente, se determina por concepto médico la realización de la curva normativa según los volúmenes normalizados por el IMC de 10 órganos, definiendo con aprobación médica los colores que iba a llevar cada órgano en la segmentación del algoritmo de IA para su identificación (Tabla 1).

Bazo = Rojo	Riñón Derecho= Azul	Riñón Izquierdo=Turquesa	Hígado=Amarillo	Aorta=Naranja
Vena cava inferior=Purpura	Vena portal y esplénica=Rosa	Páncreas= Marrón	Glándula suprarrenal Derecha= Verde	Glándula suprarrenal izquierda= Fucsia

*Tabla 1 Etiquetas para los diferentes órganos del proyecto*

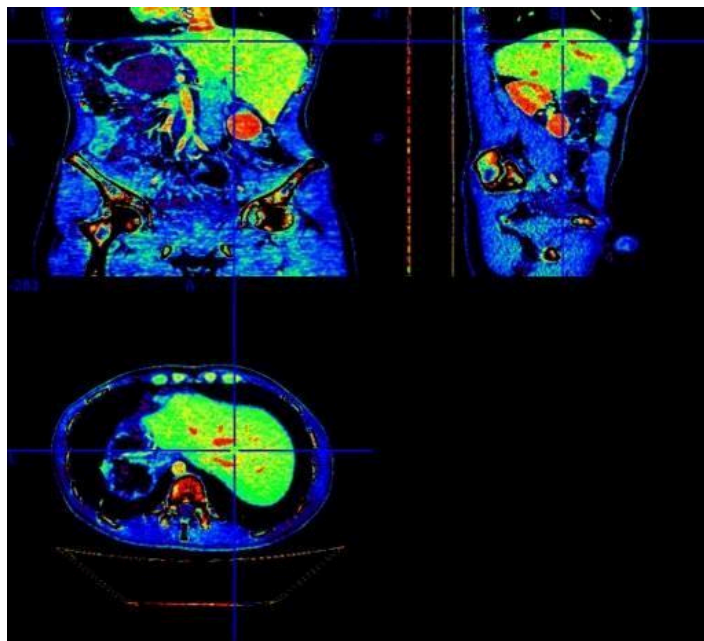
Es así como se realizaron algunas tareas específicas en base a los objetivos propuestos:

---

**1. Identificar y seleccionar una muestra de sujetos sanos**

- **Selección, caracterización e importación de imágenes de tomografía computarizada de abdomen contrastado a la plataforma de procesamiento XNAT.**

En el análisis exploratorio de los datos, se encontraron inconsistencias en varios individuos clasificados como sanos, algunos presentaban claros hallazgos patológicos, por lo que se prestó especial atención a estos descubrimientos ya que podrían sesgar negativamente la información recolectada, teniendo un efecto significativo en el análisis estadístico del modelo lo que resulta crucial para la formación.



*Fig. 2 Sujeto perteneciente a la población de sanos con nefrectomía.*

Se evidenció entonces la necesidad de validar los informes médicos de todos los sujetos que formarían la base de nuestro modelo normativo. Para abordar esta tarea, se diseñó un formato que permitiría al médico radiólogo, encargado de la validación, clasificar a los sujetos en cuatro categorías. Estas categorías se basaron en los

---

informes recuperados mediante Elasticsearch, los cuales posteriormente se estructuraron mediante algoritmos en indicaciones, conclusiones y hallazgos, para facilitar la tarea de validación. Las cuatro categorías establecidas fueron:

- **Sujetos Sanos:** Aquellos sujetos considerados aptos para la construcción del modelo normativo, contribuyendo a la representación de la población sana.
- **Sujetos patológicos:** Individuos cuyos informes concluían con problemas de salud considerablemente significativos, descartados del modelo normativo.
- **Sujetos con patologías de interés:** Sujetos que presentaban patologías con afectación directa en el volumen de un órgano específico. Estos casos se consideraron para contrastar el funcionamiento del modelo normativo y probar su eficacia.
- **No concluyentes:** Informes que no permitían una conclusión clara sobre el estado de salud del paciente, quedando en una categoría de incertidumbre.

Al revisar los resultados de la validación inicial realizada por el médico radiólogo, se identificó la necesidad de llevar a cabo una segunda evaluación. Aunque esta etapa adicional ya estaba planificada, proporcionó una serie de consideraciones importantes. El objetivo principal era mejorar la alimentación del modelo normativo y aumentar la disponibilidad de datos para su construcción de manera más eficiente. El equipo de investigación determinó que la población inicial de 574 sujetos, entre hombres y mujeres, era insuficiente, lo que llevó a buscar una ampliación.

Para abordar esta situación, se llevó a cabo una consulta en Elasticsearch para encontrar nuevos estudios de tomografía abdominal contrastada con el objetivo de aumentar la muestra de nuestro proyecto. Se realizó un filtrado meticuloso de la información para evitar la recuperación de estudios repetidos, y se generó un archivo CSV con los datos recopilados de todos los sujetos. Este archivo se utilizó para realizar



un seguimiento y una identificación más efectiva en el futuro, como estrategia para evitar la inclusión de los mismos estudios en futuras consultas y así optimizar el proceso y perfeccionar la selección de nuevos sujetos en la siguiente fase del proyecto.

Al centrarnos en las diversas clasificaciones dentro de la validación proporcionada por los médicos radiólogos, se nos proporcionó un marco claro para la validación y selección de sujetos, lo que optimizó el proceso de construcción y evaluación del modelo normativo. Iniciamos analizando las conclusiones de los estudios utilizando N-gramas de tres palabras en los sujetos sanos para refinar los criterios de inclusión (Fig. 9).

En esta segunda consulta, fue necesario aplicar criterios de inclusión específicos centrados en las conclusiones de los reportes médicos (Fig. 3), ya que esta sección resultó ser la más informativa y determinante para el diagnóstico. Durante este proceso, se cuidó de mantener un equilibrio en los rangos de edad tanto en la población masculina como en la femenina, abarcando entre 18 y 80 años.



Fig. 3. Criterios de inclusión y exclusión aplicados en la conclusión del reporte para recuperación de nuevos estudios

Luego, la información encontrada en la consulta de Elasticsearch con los criterios de inclusión y la metodología refinados, se consultó directamente en el superpacs con findscu, asegurándose de recuperar solo la información de estudios cuyas series fueran volumétricas, lo cual es crucial para nuestro proyecto. Una vez logrado esto, se

conformaron lotes con cinco sujetos cada uno para iniciar la importación de manera controlada y evitar saturar el servidor. La importación se llevó a cabo a través de una sesión de Tmux para ejecutar el proceso mientras se realizaban otras tareas específicas dentro del proyecto.

A la vez, se decidió eliminar del proyecto en la plataforma de procesamiento Xnat a aquellos sujetos clasificados como patológicos. Respecto a los sujetos catalogados como portadores de patologías de interés, se realizó un análisis adicional con la colaboración de un segundo médico radiólogo para clasificarlos en una patología específica que pudiera estudiarse.

Después de la validación inicial realizada por los médicos radiólogos, se procedió a clasificar sincrónicamente a los sujetos que presentaban patologías de interés en una de estas categorías. Las patologías de interés abordan cambios volumétricos en los órganos abdominales y se buscó aprovechar la importación previa de datos, reconociendo la complejidad del proceso. Se consideró necesario replantear el formato de validación, añadiendo una columna que definiera los casos de patología de interés con mayor especificidad, para ser precisos al indicar la patología y evitar la pérdida de información, pero a pesar de esto se hizo necesario consultar sujetos con las patologías de interés más comunes que fueron hepatomegalia, esplenomegalia y arteriosclerosis o también llamado ateroma calcificado (Tabla 2).

Patología	Importados	Nueva consulta
hepatomegalia	10 sujetos= 5M-5F	350
esplenomegalia	9 sujetos= 9M-0F	148
arteriosclerosis	10 sujetos= 5M-5F	286

*Tabla 2. Cantidad de datos recopilados de patologías de interés.*

**2. Desarrollar un proceso de control de calidad de la segmentación.**

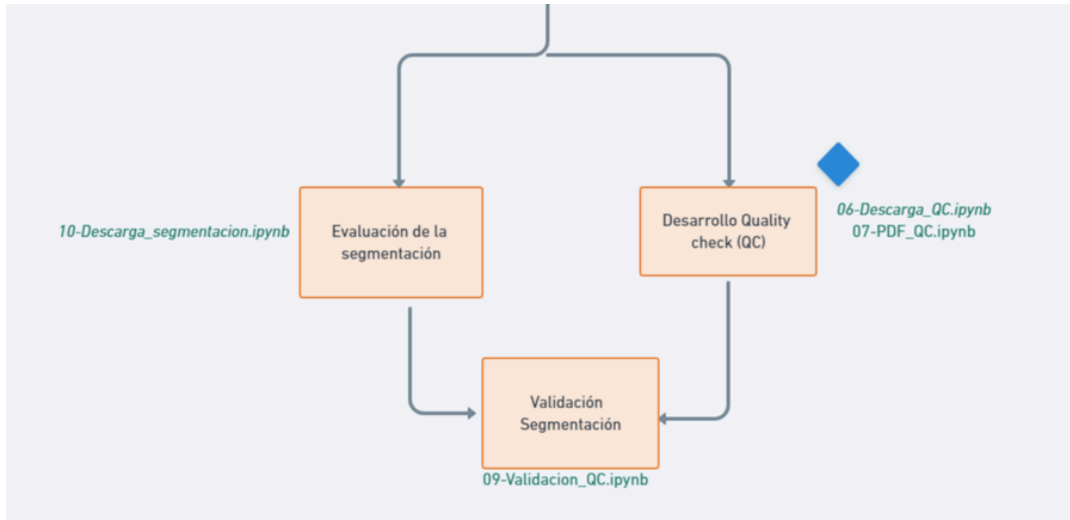


Fig. 4 Flujo de trabajo implementado para desarrollar un proceso de control de calidad de la segmentación

- **Recibir y comprender el flujo de trabajo para la extracción de volúmenes de los órganos corporales que se realizó previamente donde se integró además la herramienta de MONAI en XNAT.**

Antes de adentrarnos en el desarrollo de nuestro proceso de control de calidad, resultó necesario comprender los avances previos del proyecto. Para ello, se empleó el modelo de inteligencia artificial conocido como 3D Multi-organ Segmentation con UNETR [55], el cual desempeñó un papel fundamental como punto de referencia para el entrenamiento del modelo de segmentación. Esta arquitectura de red neuronal convolucional (CNN) se basa en las redes UNET, una estructura ampliamente reconocida en la segmentación semántica de imágenes médicas, la cual fue propuesta por Olaf Ronneberger, Philipp Fischer y Thomas Brox en 2015 [56].

La estructura de U-Net ha sido diseñada específicamente para enfrentar el desafío de segmentar imágenes médicas, donde se requiere una precisión notable en la delimitación de estructuras anatómicas o patológicas. El modelo UNETR, permite que el modelo capture características locales y aprenda relaciones globales en imágenes

---

médicas, lo que resulta particularmente útil en la segmentación precisa de órganos y estructuras anatómicas en la zona abdominal. Además, resalta su versatilidad para adaptarse a diversas técnicas de adquisición de imágenes médicas.

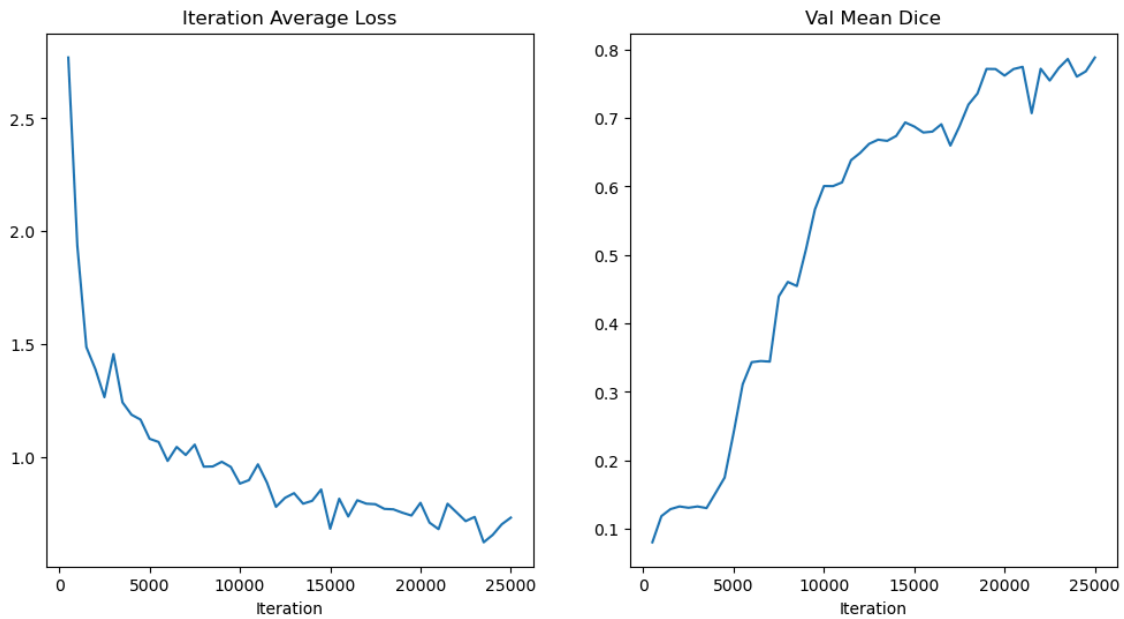
Para el entrenamiento, se utilizó una base de datos proveniente del BTCV Challenge [57], que consiste en un total de 24 imágenes. El proceso de entrenamiento se llevó a cabo en un servidor de SURA con una capacidad de procesamiento que incluye 8 núcleos de CPU y 16 GB de RAM, cada núcleo con 3.3 GHz de doble hilo. Este proceso de entrenamiento demandó aproximadamente 2 días para completar las 25,000 iteraciones establecidas en el tutorial.

Por otro lado, la validación del modelo, fue realizado previamente en el proyecto por el grupo de investigación siguiendo las pautas del tutorial, se realizó utilizando seis imágenes adicionales de la misma base de datos, garantizando así su coherencia y aplicabilidad. Este procedimiento nos permitió evaluar la efectividad del modelo UNETR en la segmentación exitosa de los órganos. Para llevar a cabo esta evaluación, se analizó la curva de pérdida del modelo (Fig. 5), que ofrece una medida de su precisión. Una pérdida menor indica un modelo más preciso. En el gráfico correspondiente (Fig. 5), se observa que la pérdida disminuye a medida que aumenta el número de épocas, lo que sugiere una mejora continua en la precisión del modelo durante el entrenamiento. No obstante, alrededor de las últimas iteraciones, se aprecia un ligero aumento en la pérdida.

Durante el pre-entrenamiento del modelo UNETR, se emplearon exploraciones con distintos tamaños de volumen y campos de visión. Los tamaños de volumen variaron entre  $512 \times 512 \times 85$  y  $512 \times 512 \times 198$ , mientras que los campos de visión oscilaron entre  $280 \times 280 \times 280 \text{ mm}^3$  y  $500 \times 500 \times 650 \text{ mm}^3$ . La resolución en el plano varió desde  $0,54 \times 0,54 \text{ mm}^2$  hasta  $0,98 \times 0,98 \text{ mm}^2$ , mientras que el grosor del corte se situó entre 2,5 mm y 5,0 mm

El modelo se entrenó con un tamaño de lote de 6 y se realizaron un total de 25,000 iteraciones tal como se observa en la figura 5, evaluando el rendimiento del modelo cada 500 iteraciones. Además, se guardó el modelo entrenado para su posterior utilización en la segmentación de los órganos abdominales. La métrica

"Iteración Media de Pérdida" representa la pérdida promedio calculada durante todas las iteraciones del entrenamiento, siendo su minimización el objetivo. Por otro lado, "Dice Medio de Validación" indica el valor promedio del coeficiente de Dice, que cuantifica la calidad y precisión de la segmentación automática, buscando maximizar.



*Fig. 5 Curva de pérdida del modelo.*

Después, se realizó una fase de prueba con la base de datos de pacientes de SURA con los 1079 sujetos iniciales, para evaluar el rendimiento del modelo en un conjunto de datos distinto.

Para garantizar la calidad de las segmentaciones de los órganos, se realizó una inspección visual de los resultados a cargo de un médico radiólogo especializado en imágenes corporales. Esta profesional utilizó un formulario para clasificar las segmentaciones de los órganos como "buena", "regular" o "mala", y proporcionó observaciones adicionales según su criterio, asegurando de este modo la precisión de las segmentaciones obtenidas en el estudio.

Tras la validación de las segmentaciones, los estudios de los sujetos seleccionados se archivaron en el repositorio central de procesamiento XNAT. Se estableció un flujo de trabajo (Fig.1) que abarcó la anonimización de los datos,

---

eliminando cualquier información privada que pudiera identificar al paciente, como el nombre y la cédula. Además, se convirtieron los archivos DICOM al formato NifTI para facilitar el procesamiento subsiguiente.

En este proceso, se aplicó el modelo de segmentación de órganos previamente entrenado y validado a cada estudio dentro del flujo de trabajo. Una vez obtenidas las segmentaciones, se empleó la función `fsldata` de FSL para extraer automáticamente los volúmenes de los riñones, hígado, bazo, esófago, estómago, vesícula biliar y páncreas. Estos volúmenes se calcularon mediante el recuento de voxeles en cada segmentación, teniendo en cuenta la resolución espacial de la imagen, es decir, el tamaño del voxel en cada dimensión.

- **Segmentación**

Después de este procedimiento, se avanzó al proceso de segmentación, el cual se inicia con la importación de las diversas imágenes de los estudios a XNAT desde el PACS. Esta acción facilita la conversión del formato DICOM, utilizado para capturar las imágenes del PACS, al formato NifTI. Tal conversión ofrece ventajas significativas, ya que estandariza y simplifica la gestión de los datos. Numerosos programas y herramientas de análisis de imágenes médicas están específicamente diseñados para trabajar con archivos NifTI. Además, los archivos NifTI son más fáciles de manejar y visualizar en comparación con los archivos DICOM, que pueden consistir en múltiples archivos y contener metadatos complejos. Por lo general, los archivos NifTI son archivos únicos que contienen toda la información necesaria para la imagen.

La segmentación en MONAI implica el proceso de identificar y delimitar estructuras o regiones específicas en imágenes médicas, tales como el abdomen en imágenes de tomografía computarizada (TC) o resonancia magnética (RM). Este proceso puede ser manual, donde un operador o especialista define las áreas de interés en cada corte de la imagen. Por otro lado, puede ser automático, como en nuestro caso, donde algoritmos computacionales identifican y delimitan las estructuras

---

utilizando propiedades de intensidad de píxeles, texturas u otras características presentes en la imagen.

- **Extracción de volúmenes**

La extracción de volúmenes de órganos en XNAT y MONAI implica un proceso de segmentación previo, seguido por el cálculo del volumen basado en la segmentación obtenida. Por lo tanto, es crucial la implementación de MONAI, una biblioteca de aprendizaje profundo para aplicaciones médicas. En MONAI, la extracción de volúmenes de órganos se suele realizar mediante modelos de segmentación, como el un modelo multiorgánico UNETR, basado en redes neuronales convolucionales. Este modelo, después de ser entrenado con datos etiquetados (segmentados), se puede utilizar para segmentar estructuras de interés en nuevas imágenes médicas y calcular sus volúmenes correspondientes basados en la segmentación obtenida.

La extracción de volúmenes debe ir acompañada de la normalización de los volúmenes de los órganos. En este caso, se optó por normalizarse en relación con el Índice de Masa Corporal (IMC), debido a la accesibilidad dentro de la Teradata de SURA a este valor. Esto implica ajustar los volúmenes de acuerdo con el IMC del paciente. Dicha normalización es fundamental para facilitar una comparación efectiva entre pacientes. Al realizar análisis estadísticos, como correlaciones u otras pruebas, se busca explorar la relación entre los volúmenes de los órganos y el IMC. Esta exploración estadística pretende integrar la variabilidad inherente en la morfología de diferentes sujetos, considerando que el tamaño de los órganos puede estar correlacionado con características morfológicas individuales, como la altura y el peso.

En el proyecto, el enfoque principal se centró en las tomografías abdominales contrastadas, dado que este tipo de imágenes implica la administración de un agente de contraste que mejora la visibilidad de estructuras anatómicas y tejidos. Esta mejora en la visualización permite una diferenciación más precisa entre órganos y tejidos, lo que puede resultar en una segmentación más precisa por parte del algoritmo de inteligencia artificial. Además, facilita la interpretación de la segmentación por parte del

---

médico radiólogo, lo que puede contribuir a la detección de anomalías que pueden no ser evidentes en el reporte médico inicial.

- **Curva normativa**

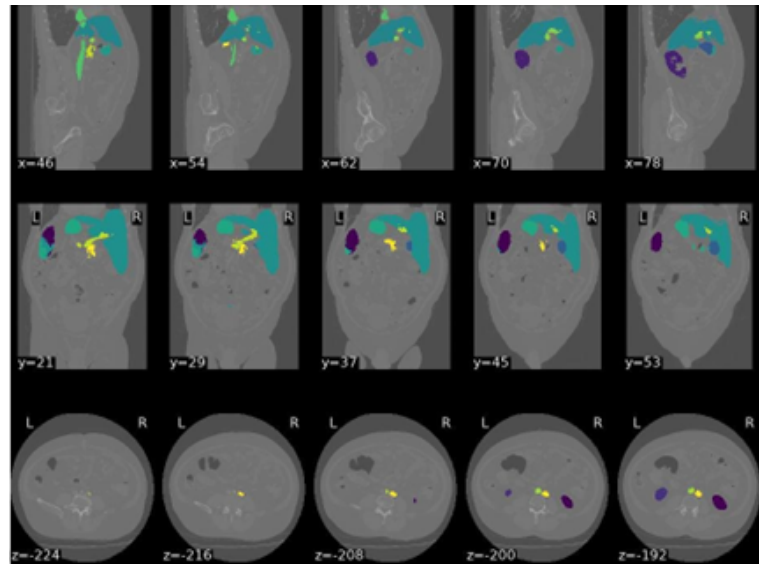
Estas curvas se basan típicamente en datos recopilados de una muestra representativa de la población, lo que permite establecer patrones esperados para diferentes grupos de edad, género u otros criterios relevantes [39].

En este proceso de control de calidad posterior al análisis del médico radiólogo se determinó que algunos sujetos se debían descartar del modelo normativo, por lo que a partir de aquellos sujetos que se validaron como útiles para engrosar nuestro modelo se crearon algunos modelos normativos para tener así un primer acercamiento y hacer una análisis de qué otras variables podrían afectar al modelo, donde se pudo comprender más a fondo que cuando se dispone de pocos datos, la generación de curvas normativas puede ser más desafiante y menos precisa. En tales casos, es posible que no se pueda representar una curva suave o detallada debido a la limitada cantidad de información.

- ***Validación de la segmentación realizada previamente del proyecto MONAI a nivel local para la definición y extracción de los volúmenes de: riñones, hígado, bazo, páncreas, estómago, esófago, adrenales entre otros órganos.***

Se desarrolló una metodología para analizar la segmentación de forma automática en la plataforma de procesamiento XNAT, llamado Quality check que son un conjunto de imágenes con cortes específicos de la segmentación. Esta iniciativa surgió ante la falta de un método efectivo para visualizar rápidamente los resultados de la segmentación de los diferentes sujetos del proyecto implementado en XNAT. Como se observa en la figura 6, cada eje (coronal, sagital y axial) tenía 5 cortes, pero no mostraban las estructuras precisas además de que los colores de los etiquetados de cada órgano variaba entre las diferentes imágenes.





*Fig. 6 Quality check inicial en el proyecto de volumetría de abdomen.*

Para afrontar este desafío y capitalizar la colaboración de los médicos radiólogos en el proyecto, iniciamos solicitando su orientación sobre qué cortes deberíamos integrar en nuestro proceso de control de calidad. Basándonos en criterios médicos, decidimos llevar a cabo un control de calidad con 15 cortes axiales específicamente seleccionados, los cuales son comúnmente utilizados por los profesionales de la salud para analizar las diversas estructuras abdominales. Cada uno de estos cortes fue elegido meticulosamente para garantizar la relevancia y utilidad para los médicos radiólogos, y su selección fue sometida en todo momento a la aprobación de los profesionales médicos. Este enfoque nos permitió identificar los cortes que proporcionaban la información más significativa, sirviendo como la base para que los médicos radiólogos pudieran analizar detalladamente la segmentación.

En este proyecto, nos valimos de la biblioteca Nilearn, una herramienta de Python ampliamente reconocida por su utilidad en el análisis y visualización de datos de imágenes. Nilearn se enfoca en la manipulación, análisis y representación de datos provenientes de técnicas de adquisición de imágenes como la tomografía computarizada (TC), y se ajusta perfectamente a la convención radiológica. Fue

esencial para nuestro proyecto operar dentro de esta convención, ya que facilita significativamente la interpretación de imágenes para los profesionales de la salud. Optamos por utilizar la función `plot_roi` de Nilearn debido a sus ventajas distintivas en comparación con otras funciones disponibles en la biblioteca entre ellas el manejo de las imágenes en convención radiológica.

Luego, utilizando colores específicos para cada uno de los 10 órganos involucrados en nuestro modelo normativo (ver Tabla 1), facilitamos la visualización y evitamos sesgos y confusiones. Para esta tarea, implementamos la biblioteca Nilearn, la cual resultó de gran utilidad al simplificar significativamente la representación de los múltiples cortes, como se observa en la figura 7.

Además, se aseguró de definir los cortes de la tomografía mediante proporciones (Ecuación 11) para lo que se llevó un estudio sobre las proporciones que más nos aportarán información de los resultados de la segmentación de la mayoría de los órganos, lo que permitió estandarizar el algoritmo para todos los sujetos del proyecto, dado que las imágenes presentes en el proyecto no tenían las mismas dimensiones. También se permitió identificar la nomenclatura del sujeto al que pertenece el control de calidad, lo que facilitó identificar problemas específicos en la segmentación de un sujeto para estudiar su comportamiento y analizar posibles patrones dentro de la segmentación.

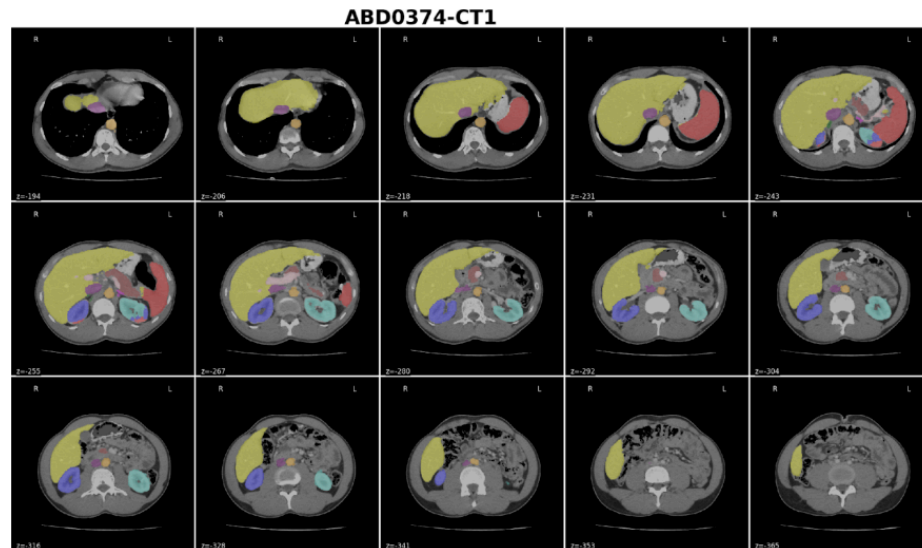
$$(valor + (paso * z\_dim) * X)$$

*Ecuación 12. Cálculo de los diferentes cortes para el Quality Check*

Donde el valor es igual a la coordenada inferior en el eje Z, paso es el factor de escala entre el sistema coordenado y la matriz de transformación, Z\_dim es la dimensión en el eje z de la imagen y X es un porcentaje específico de la imagen.

Como último paso, implementamos este Control de Calidad en Xnat para automatizar el proceso cada vez que se agrega un nuevo sujeto al proyecto. Además,

creamos un archivo PDF para facilitar la visualización rápida de los resultados de las segmentaciones de todos los sujetos por parte de los médicos radiólogos.



*Fig. 7 Quality check implementado para validación.*

Es importante destacar que todas las imágenes no comparten las mismas dimensiones, lo que requirió un ajuste basado en la matriz de transformación de cada imagen (Ecuación 11). Este ajuste se centró en el componente de traslación o desplazamiento en el eje Z (profundidad), teniendo en cuenta el step o paso. Esto nos permitió analizar tanto el sistema de coordenadas de la imagen como los cortes que están a diferentes escalas, facilitando así la manipulación de los diferentes cortes por proporciones.

Con base en este desarrollo implementado, se validó de nuevo la segmentación. Dado que la primera validación se realizó con un número limitado de sujetos, este segundo proceso nos permitió abarcar a una gran parte de la población en un período de tiempo corto. Un médico radiólogo, el profesional más adecuado para esta labor, realizó la tarea específica de validación, entregándonos la validación de unos 20 sujetos, sin embargo, por petición del médico radiólogo al encontrar varias

---

inconsistencias solicitó al equipo acceder a la segmentación completa para corroborar de manera más detalladas posibles inconsistencias en la segmentación.

Para esto se seleccionaron múltiples sujetos aleatoriamente, abarcando los rangos de edad, para evitar sesgos en el estudio y así lograr una visualización completa de la segmentación, en dicho proceso resultó crucial desarrollar una guía de la herramienta de visualización de imágenes (MRicron) para el médico radiólogo. Esto garantizó que pudiera utilizar correctamente la herramienta y validar la segmentación con éxito. La iniciativa resalta la importancia de contar con un equipo multidisciplinario dentro del proyecto.

Además, se llevó a cabo una prueba pareada por edad en la cual se seleccionaron aleatoriamente sujetos sanos y se los emparejó con sujetos patológicos, tanto para esplenomegalia como para hepatomegalia. El propósito fue identificar cualquier patrón que pudiera indicar, como era de esperarse, mayores volúmenes en el hígado o el bazo obtenidos mediante el algoritmo de segmentación en los sujetos patológicos. Esta comparación permitió una evaluación más precisa de las diferencias en los volúmenes extraídos por el algoritmo de segmentación de estos órganos entre los grupos de sujetos sanos y aquellos con afecciones hepáticas o esplénicas, lo que nos permitió evaluar el desempeño del algoritmo tanto en su segmentación como en su extracción de volúmenes.

A partir del Quality Check desarrollado, se procedió a inspeccionar visualmente los resultados de la segmentación en una muestra representativa de la población patológica, compuesta por 74 sujetos entre sujetos con hepatomegalia y esplenomegalia.

**3. Filtración de los datos, para la construcción de un modelo normativo robusto.**

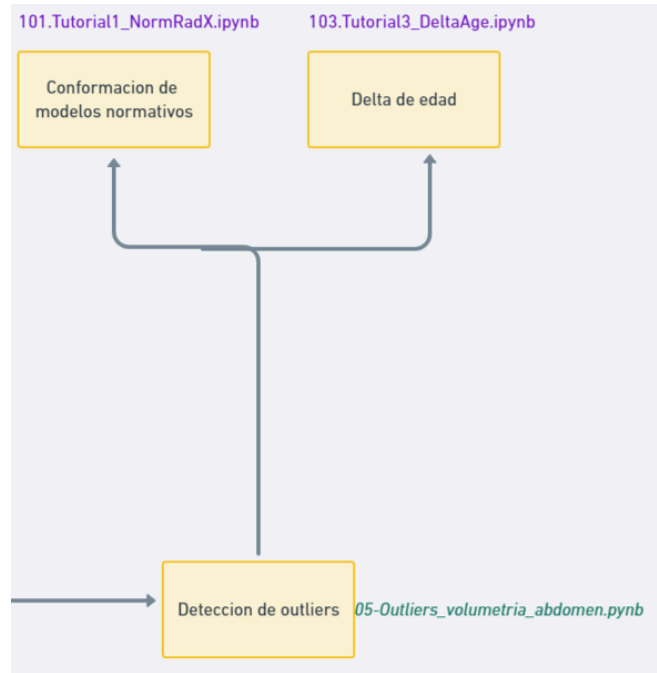


Fig. 8 Flujo de trabajo implementado para filtrar los datos y la construcción de un modelo normativo robusto.

- **Revisar y aplicar técnicas que permitan detectar valores atípicos para su posterior tratamiento o eliminación en caso de que afecte significativamente la estadística de nuestro modelo normativo.**

A pesar de los sobresaltos encontrados durante el proceso de validación de la segmentación se procedió a depurar los datos y continuar con el flujo de trabajo para implementar en el futuro, tras la mejora de la segmentación definida. Para esto entonces, se aplicó un procedimiento a los volúmenes obtenidos de Xnat para identificar outliers o valores atípicos. Es importante tener en cuenta las diferencias morfológicas entre la población masculina y femenina, por lo que se desarrollo un modelo normativo específico para cada género.

---

Se creó un algoritmo de detección automática de valores atípicos en Python a través de una función. Esta función permite elegir la técnica a utilizar por el algoritmo, así como otros parámetros de entrada, como el umbral a emplear en la técnica, el conjunto de datos con los volúmenes normalizados con el Índice de Masa Corporal (IMC), los nombres de las columnas del dataframe que serán evaluadas para detectar los valores atípicos, y un parámetro "verbose". Este último, utilizado para describir la cantidad de detalles o información proporcionada en la salida del programa, proceso u operación. Cuando este parámetro se establece como verdadero teniendo en cuenta que se trata de una variable booleana, la función muestra una serie de gráficos que permiten analizar el comportamiento de los valores atípicos tanto en la distribución por estructura (ver figura 20) como en la cantidad de cada uno de los sujetos que componen la población de sujetos sanos (ver figura 22).

En el marco teórico desarrollado previamente, se presentaron diversas técnicas de detección de outliers que se incluyeron para ser utilizadas en la función, considerando su implementación futura según las consideraciones que se tomen. Es importante presentar las diferentes técnicas de detección de outliers disponibles en la función, con el fin de ofrecer una visión completa de las opciones disponibles y permitir una evaluación adecuada de cuál sería la más apropiada para el contexto específico del estudio.

La salida del algoritmo consistía en dos matrices, independientemente del valor de la variable booleana "verbose". En ambos casos, se realizaba una evaluación de los datos suministrados, lo que generaba dos dataframes como resultado. Uno de los dataframes contenía el valor de todos los volúmenes, junto con una columna adicional que especificaba el número de valores atípicos detectados, que se utilizó para crear un archivo CSV con los datos y alimentar el modelo normativo. La otra matriz estaba compuesta por unos y ceros, donde los unos representaban la presencia de valores atípicos en esa instancia específica. Esta salida permitía evaluar los órganos que presentaban valores atípicos para cada uno de los sujetos incluidos en el proyecto.

Para realizar una primera prueba de el modelo normativo, se utilizó la técnica de detección de outliers conocida como Desviación Absoluta de la Mediana (MAD), la cual

---

se calcula como la mediana de las desviaciones absolutas de cada dato respecto a la mediana de todo el conjunto de datos (Ecuación 8).

Cabe destacar que, aunque inicialmente se planeaba incluir el estómago, el esófago y la vesícula en el modelo normativo, estos órganos fueron excluidos a raíz de la consideración de los médicos radiólogos. Respecto al estómago, la imposibilidad de asegurar el estado de ayuno de los pacientes al momento de la tomografía aumentaba el riesgo de variaciones no deseadas en este órgano. Por otro lado, el esófago es altamente móvil y anatómicamente variable entre individuos, lo que dificulta la creación de un modelo normativo preciso. La variabilidad en la posición y tamaño de este órgano puede ser significativa entre diferentes personas, lo que afectaría su exactitud por lo que se decidió abordar solo los diez órganos restantes en los modelos normativos.

Se desarrolló así un algoritmo automatizado para abordar la adecuación de los datos, el cual permite detectar outliers de manera automática. Este algoritmo evalúa cada uno de los diez órganos y descarta del desarrollo del modelo normativo aquellos sujetos que presenten outliers en tres o más órganos. De esta manera, se identifican posibles errores de segmentación mediante el número de outliers, lo que facilitará la detección de posibles errores en la segmentación y corroboraría a su vez la depuración adecuada de los datos.

Además de esto, se llevó a cabo una evaluación del porcentaje de outliers por órgano en la misma población anterior, tanto masculina como femenina, de este modo se pretende luego de la futura validación de la nueva segmentación encontrar aquellos órganos que pudieran resultar más sensibles a errores de segmentación, de modo que le permita al equipo de trabajo identificar falencias en el algoritmo con una mayor certeza.

Así mismo se implementó un algoritmo para que se haga un conteo de manera automática de los sujetos con menos de 3 outliers (que son los que se incluirán en el modelo normativo final) por rango de edad cuidando de que no se tengan sesgos en la creación del modelo normativo y se abarque el rango de edades planteadas inicialmente por el equipo tanto en población masculina como femenina.

---

- ***Creación del modelo normativo poblacional SURA.***

Aunque es importante destacar que no se han desarrollado modelos normativos definitivos, hemos avanzado en esa dirección basándonos en la población obtenida después de la validación de los reportes (ver Fig. 10) y en el algoritmo desarrollado para la detección automática de valores atípicos. Aunque aún no son concluyentes, estos modelos representan un primer paso significativo hacia una comprensión más sólida de los modelos normativos. En esencia, estos modelos son estructuras teóricas destinadas a establecer estándares o normas idealizadas en contextos específicos. Su aplicación abarca diversas disciplinas, ya que se utilizan para definir estándares de referencia, patrones deseados o comportamientos ideales, actuando como puntos de comparación o medidas.

Dado que los modelos normativos desarrollados no ofrecen conclusiones definitivas, decidimos concentrarnos en la creación de modelos para cinco órganos específicos: ambos riñones, el bazo, el hígado y la arteria aorta abdominal. Nuestro objetivo era comprender mejor cómo se comportan las curvas normativas en estos órganos. Para esta fase, seleccionamos exclusivamente sujetos con menos de tres valores atípicos entre los diez órganos analizados.

Dividimos nuestros datos de muestra en un conjunto de entrenamiento del 80% y un conjunto de prueba del 20%, utilizando validación cruzada k-fold con 5 pliegues. En cada iteración de este proceso, los datos de entrenamiento se reorganizan y se dividen nuevamente en subconjuntos de entrenamiento y prueba. Luego, ajustamos nuestro modelo utilizando el subconjunto de entrenamiento y evaluamos su rendimiento utilizando el subconjunto de prueba. Las evaluaciones de rendimiento de cada pliegue se promedian para obtener una evaluación. Finalmente, ajustamos nuestro modelo utilizando todo el conjunto de entrenamiento y lo aplicamos al conjunto de prueba para su evaluación definitiva. Posteriormente, comparamos el rendimiento de nuestro modelo utilizando cuatro métodos de regresión diferentes: regresión Cuantil, regresión Centil, procesos gaussianos (GP) y suavizado de diagrama de dispersión estimado localmente (LOESS).



---

Las métricas empleadas para comparar el rendimiento de los modelos fueron el Error Cuadrático Medio (RMSE) y el Error Cuadrático Medio Estandarizado (SMSE). Estas métricas permiten al algoritmo identificar automáticamente cuál de las cuatro regresiones tenía un menor error. Cuando el error es menor, indica una mejor capacidad predictiva y mayor precisión en las predicciones del modelo. En esencia, un error más bajo sugiere que las predicciones están más cercanas a los valores reales en promedio, lo que refleja un mejor ajuste del modelo a los datos y una mayor confiabilidad en sus predicciones. En resumen, un error más bajo es un indicador positivo de la calidad del modelo, indicando que este captura de manera más efectiva la relación entre las variables en el conjunto de datos que es lo que finalmente se busca.

- ***Implementación del modelo en sujetos sanos y patológicos para determinar su funcionalidad.***

Ante el desarrollo del modelo normativo y la comparación de las diferentes curvas generadas para cada uno de los órganos, se decidió realizar una comparación entre sujetos sanos y aquellos con patologías, con el fin de examinar minuciosamente posibles disparidades entre ambas poblaciones. Para este análisis, se redujo la cantidad de órganos estudiados a dos: el hígado y el bazo, ya que se contaba con un mayor número de sujetos con patologías relacionadas con estos órganos. Específicamente, se seleccionó la hepatomegalia como patología a comparar para el hígado, mientras que para el bazo se optó por la esplenomegalia.

Para llevar a cabo esta comparación, se empleó la prueba de delta de edad, este indicador puede utilizarse para estimar si un individuo presenta un envejecimiento acelerado o retardado en relación con su edad cronológica real. El algoritmo analiza los volúmenes de órganos específicos, como el hígado y el bazo en nuestro caso, y los compara con una base de datos que contiene volúmenes de órganos de individuos de distintas edades.

Se llevaron entonces a cabo dos pruebas, considerando la regresión cuantil como la mejor opción después de analizar resultados anteriores. En una de ellas, se

---

construyó la curva utilizando los volúmenes del órgano afectado por la patología, el hígado para evaluar sujetos con hepatomegalia y el bazo para evaluar sujetos con esplenomegalia. En la segunda prueba, la curva se formó utilizando los volúmenes de los diez órganos inicialmente considerados para la construcción de los modelos normativos.

Finalmente, se llevó a cabo un análisis comparativo emparejado por edad y sexo entre sujetos sanos y aquellos con patologías para examinar el comportamiento de ambas muestras (ver figuras 14 y 15). Durante este análisis, se identificaron algunas irregularidades que proporcionan más evidencia sobre los errores en la segmentación, reforzando la necesidad urgente de mejorar el modelo de segmentación utilizado. Además, al revisar los informes generados, se observó una relación notable entre los sujetos con hepatomegalia y esplenomegalia. Se encontró que la coexistencia de hepatomegalia y esplenomegalia en un mismo individuo era más frecuente de lo anticipado, lo que destaca la importancia de investigar más a fondo estas condiciones y su posible asociación en futuros estudios.

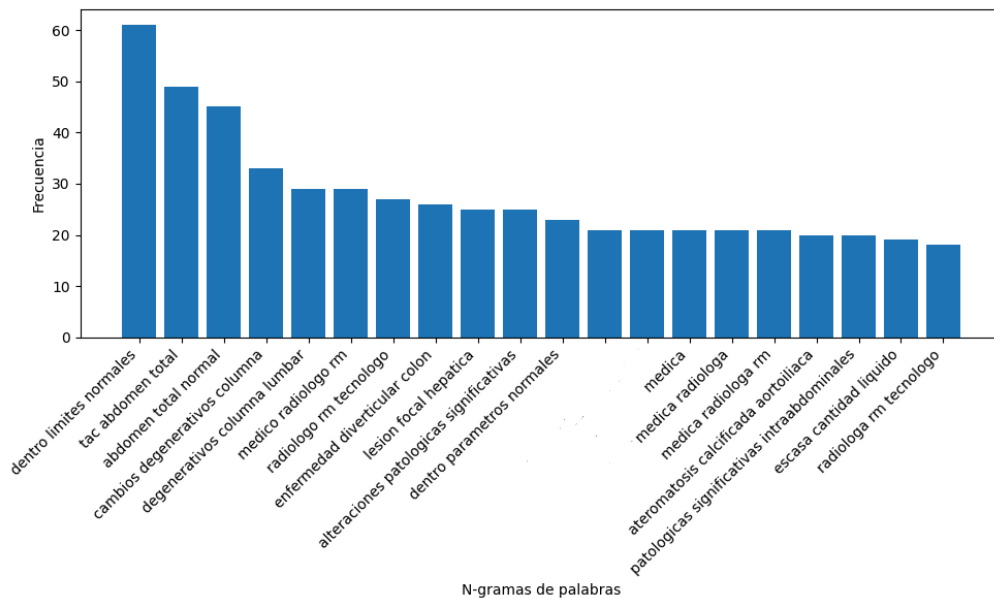
---

## **VII. RESULTADOS**

### ***1. Identificar y seleccionar una muestra de sujetos sanos***

Al de encontrar múltiples errores en la recolección de los datos fue necesario someter los informes médicos todos los individuos a una validación por parte de profesionales de la salud para garantizar la integridad y la precisión de los datos del proyecto, así los primeros sujetos implementados (1079) fueron llevados a un proceso de validación de sus reportes médicos para establecer cuáles sujetos eran idóneos para la conformación de la población sana que se había estado buscando desde un inicio.

Con la intención de ampliar la población de nuestro proyecto, decidimos aplicar N-gramas directamente a las conclusiones de los informes de aquellos sujetos que realmente habían sido validados como sanos por profesionales de la salud, como se detalla en la Figura 9. Sin embargo, encontramos que los N-gramas resultaron ser menos informativos de lo esperado. Por lo tanto, optamos por llevar a cabo una revisión rápida de varios informes médicos, nuevamente con la asistencia de profesionales de la salud. Esta estrategia combinada fue esencial para mejorar los criterios de inclusión previamente establecidos. Como resultado, logramos recuperar 1,561 nuevos estudios, lo que aumentó el número total a 2,640 sujetos sanos (sin excluir aquellos que no fueron clasificados como sanos), validados por médicos radiólogos. Además, se incorporaron 10 pacientes con cirrosis, 34 sujetos con hepatomegalia y 32 con esplenomegalia en la plataforma de procesamiento XNAT.

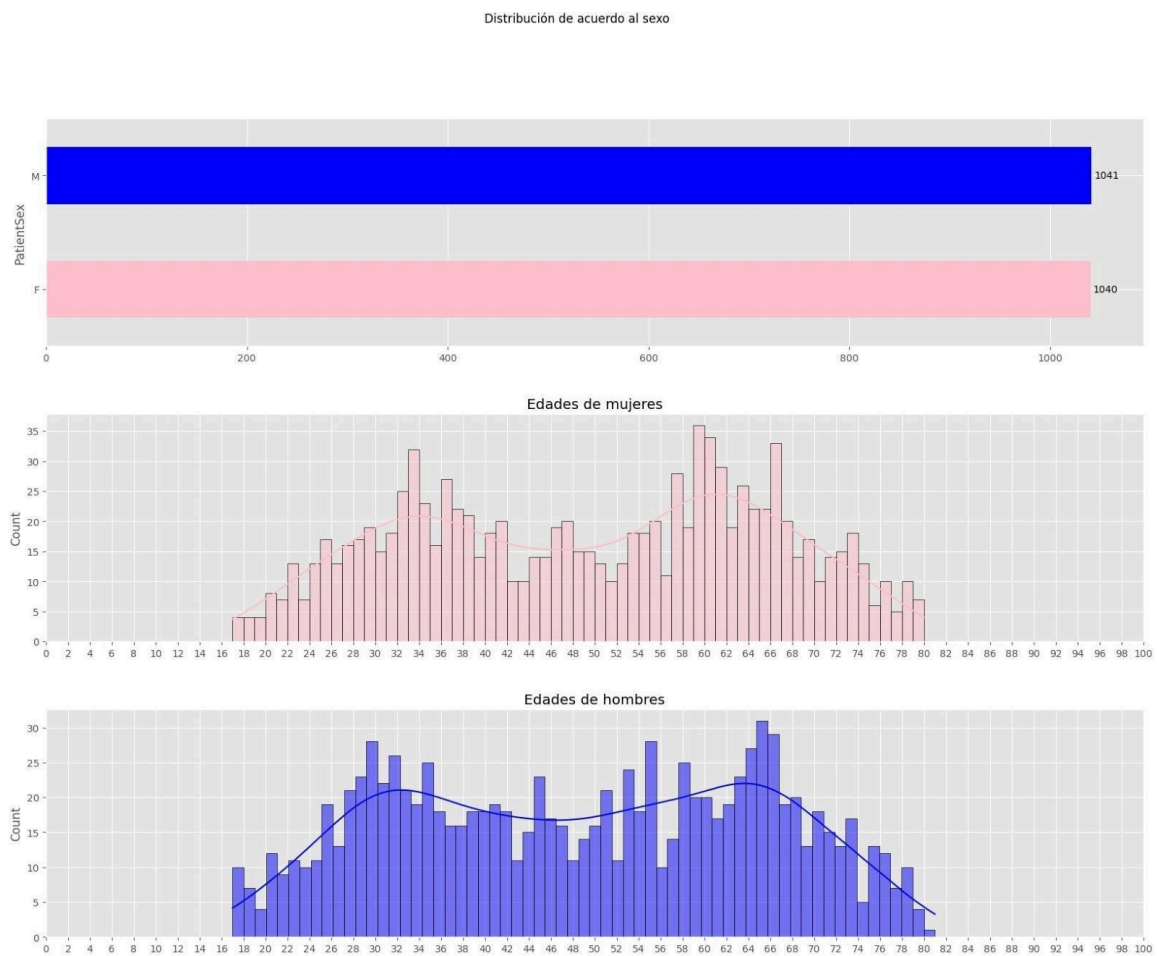


*Fig. 9 N-grama de las conclusiones de los estudios de la población validada sana.*

Durante este período, se llevó a cabo una reclasificación de los sujetos afectados por patologías de interés, sin embargo, durante este proceso, se observó que la mayoría de los sujetos no pudieron ser reclasificados, lo que pone de relieve la complejidad de esta tarea. Es importante destacar que esta tarea de reclasificación fue realizada por un médico radiólogo diferente al que realizó la validación inicial, por lo que esto nos condujo a la necesidad de replantearnos el formato de validación de los reportes con el fin de perder la menor cantidad de información posible dado lo complejo que puede resultar la alimentación de nuestro proyecto en cuestión de tiempo.

Después de analizar la reclasificación de las patologías de interés, identificamos tres que resultaron ser comunes en comparación con otras: hepatomegalia, esplenomegalia y aterosclerosis. Esto nos llevó a realizar consultas específicas a los sujetos afectados por estas patologías, incluyendo directamente el nombre de las mismas en las conclusiones del informe estructurado. Esta estrategia de consulta nos proporcionó la certeza de que estábamos tratando con casos específicos de estas patologías, lo que simplificó el proceso de identificación y recopilación de datos relevantes para nuestro posterior análisis, obteniendo así una muestra representativa de la población, como se muestra en la Tabla 2.

Finalmente, hemos logrado con éxito enriquecer nuestro proyecto en XNAT con tomografías abdominales contrastadas de sujetos genuinamente sanos, obteniendo una población total de 2081 individuos, tanto masculinos como femeninos, con prácticamente la misma cantidad de muestras al observar que el número de sujetos sanos en ambos grupos es similar, lo que indica una distribución equitativa entre géneros en la muestra. Además, hemos cubierto el rango completo de edades establecido anteriormente, como se muestra en la Figura 10. Este logro garantiza que nuestro modelo normativo no presente sesgos debido a la falta de datos en rangos de edad específicos, lo que resalta nuevamente la riqueza de datos de la compañía. Con este resultado, queda demostrado que el modelo normativo es viable con los datos actuales y, además, es posible aumentar la muestra en el futuro.



*Fig. 10 Distribución de la población y edades de los sujetos sanos por género.*

---

**2. *Desarrollar un proceso de control de calidad de la segmentación.***

Al revisar algunas segmentaciones, se detectaron errores en ciertos sujetos, revelando patrones no deseados como la asignación incorrecta de etiquetas correspondientes al hígado en partes de las piernas y la identificación errónea de algunas estructuras por parte del algoritmo, como se ilustra en la Figura 13. El equipo de trabajo ya había previsto estas situaciones. Como se dijo, alcanzar una segmentación perfecta no es factible, por lo que fue crucial realizar un proceso de validación.

El formato adecuado para la validación, tanto de los informes médicos para garantizar la inclusión de sujetos genuinamente sanos como para la validación de la segmentación, se desarrolló con un enfoque ingenieril. Esto no solo aseguró la validación de la información, sino que también proporcionó datos adicionales y valiosos para permitir un análisis más detallado de los datos recopilados. Se diseñó un formato de validación para la segmentación que permitió evaluar la calidad de la segmentación de las 10 estructuras como buena, mala o regular, e incluyó un espacio específico para observaciones. Esto facilitó un análisis exhaustivo de los resultados proporcionados por el médico.

Gracias al respaldo de los médicos radiólogos hacia el Quality Check desarrollado en colaboración con ellos en este proyecto (Fig. 7), se logró mejorar de manera significativa la evaluación rápida de la segmentación. Como resultado, se procedió a iniciar la validación de la segmentación basándose en dicho control de calidad, identificando al páncreas como el órgano que presentó mayores dificultades de segmentación (Fig. 11).

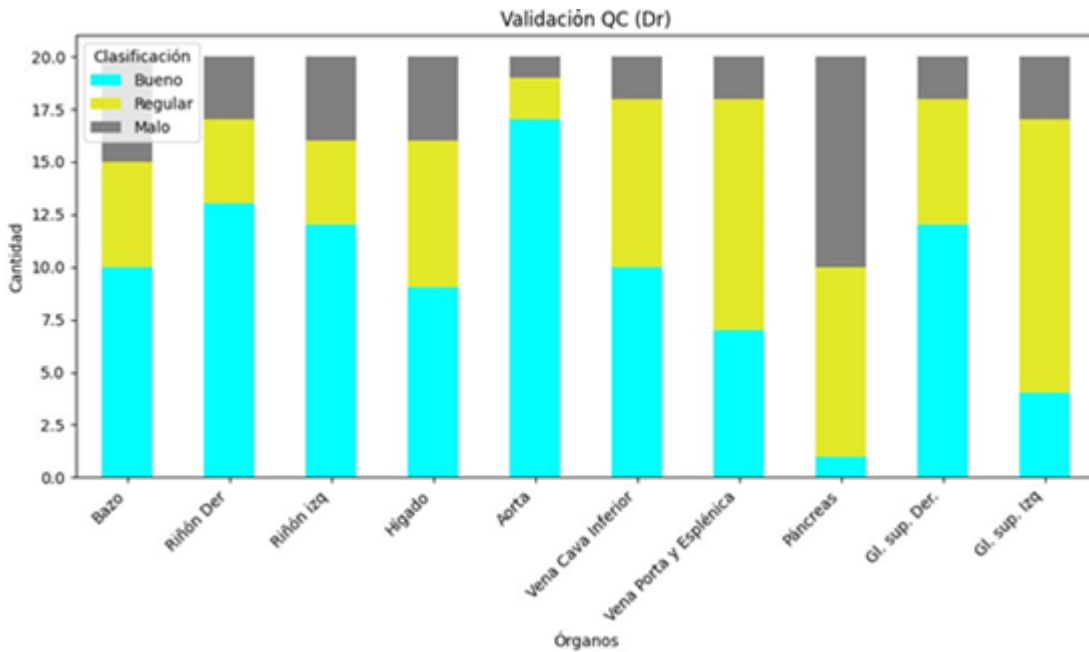


Fig. 11 Resultados de la validación del Quality Check

Después de esta validación y gracias al formato desarrollado, el cual permitió tener en cuenta las observaciones, fue posible notar que el médico proporcionó una retroalimentación crucial. Esta reveló una limitación que el grupo de investigación no había considerado: algunas estructuras, especialmente las pequeñas como las glándulas suprarrenales, no eran claramente visibles con el enfoque utilizado en los múltiples cortes, lo que dejaba margen para valoraciones subjetivas en la segmentación.

Al enfrentarnos a estas limitaciones, se hizo necesario también visualizar la segmentación completa en 3D, lo que nos llevó a desarrollar una guía detallada para capacitar al personal médico en el manejo adecuado del visor utilizado en este proyecto, en este caso, MRicron. Sin embargo, al analizar los resultados (Fig. 12), observamos que la segmentación del hígado presentaba un desempeño insatisfactorio, ya que la mayoría de las segmentaciones fueron calificadas como regulares por el médico radiólogo.

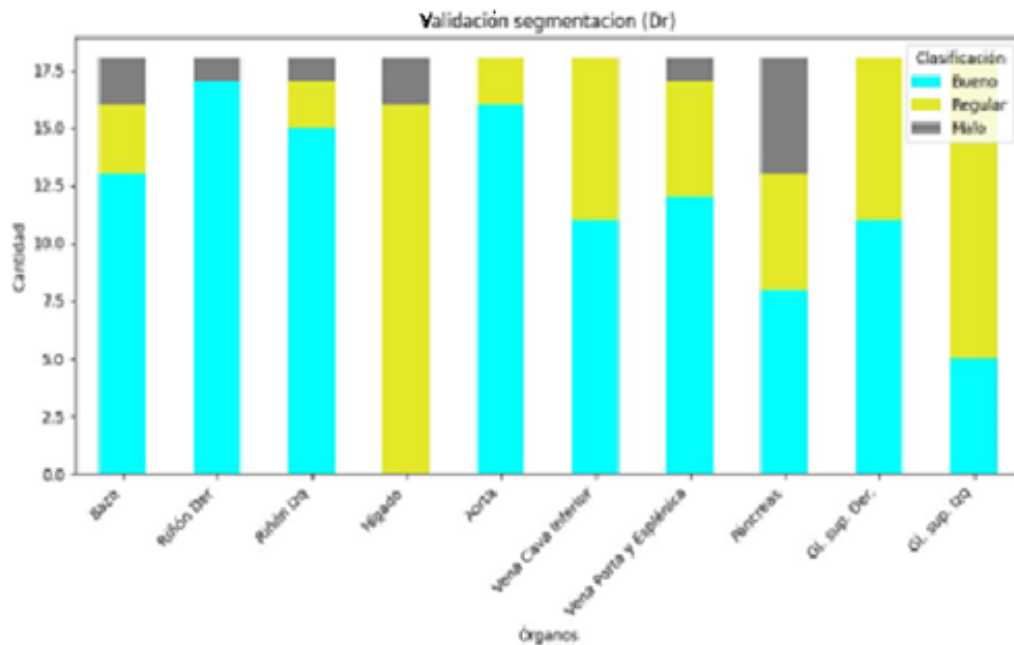
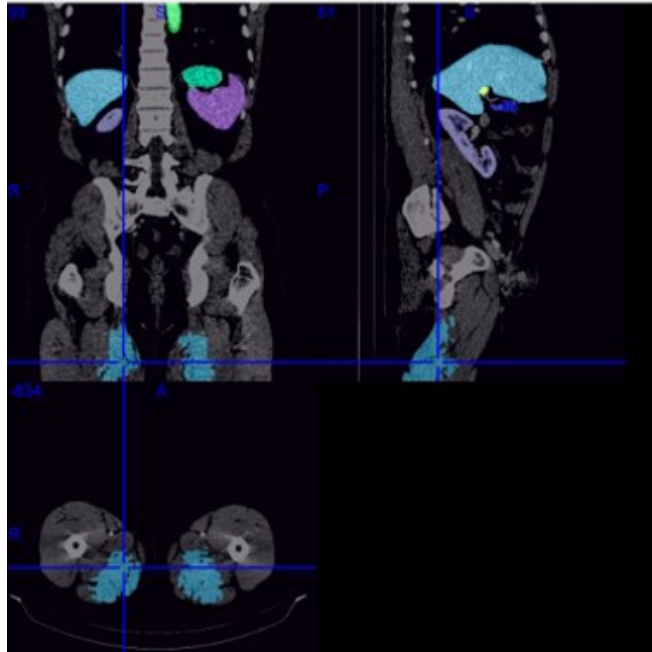


Fig. 12 Resultados de la validación de la segmentación en el visor Mricron.

Al examinar más de cerca la segmentación, se identificaron varios errores en los contornos, especialmente en algunas estructuras. Se observó que, en algunos casos, los contornos de los órganos se extendían más allá de sus límites reales, lo que resultaba en una segmentación excesiva. Por otro lado, en otras ocasiones, la segmentación no lograba capturar todo el órgano, dejando áreas sin incluir. Esta inconsistencia en la delimitación de los contornos de los órganos plantea desafíos significativos en la precisión y la fiabilidad de la segmentación automática de órganos abdominales mediante inteligencia artificial.





*Fig. 13 Sujeto perteneciente a la población de sanos con problemas de segmentación*

Es destacable por otro lado, que, durante la validación realizada del QC (Figura 11) y la validación completa de la segmentación (Figura 12), la arteria aorta se ha revelado como la estructura abdominal con los resultados más consistentes, predominantemente clasificada como segmentaciones exitosas.

En el caso de los sujetos con hepatomegalia, como se puede observar en la Figura 14, se identificaron solo unos pocos casos en los que se apreciaban claramente mayores volúmenes en los sujetos afectados. Sin embargo, es importante destacar que en la población femenina se observaron más picos de volúmenes elevados en comparación con la población masculina.

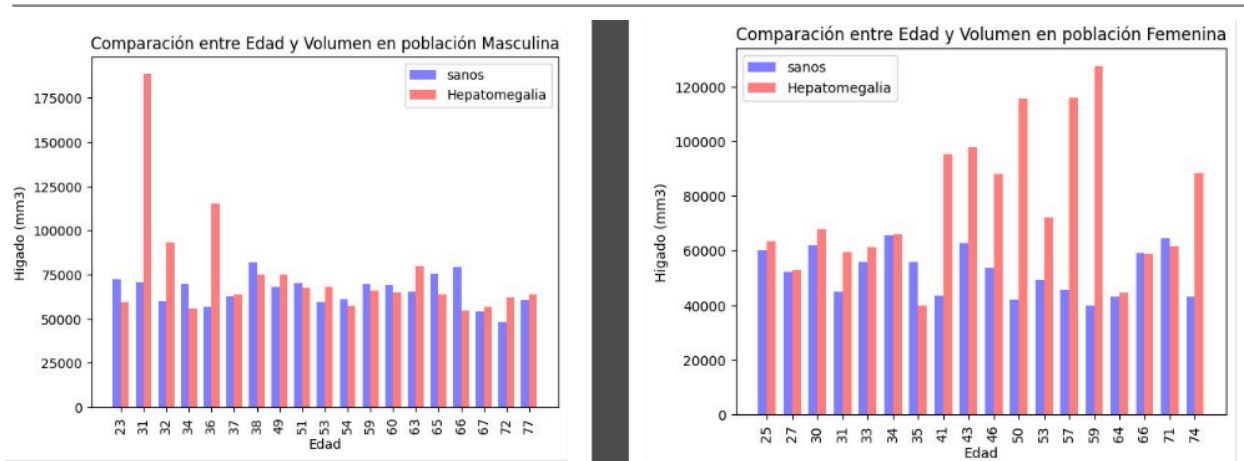


Fig 14. Prueba pareada para comparar algunos volúmenes entre sujetos sanos y con hepatomegalia

Al analizar los sujetos con esplenomegalia, se observó un patrón similar, como se muestra en la figura xyz. Esto llevó al equipo de investigación a considerar la evaluación de las segmentaciones solo en la población patológica. Es posible que los resultados anteriores se deban a errores de segmentación también en esta población, ya que es importante destacar que anteriormente en este mismo proyecto se llevó a cabo un análisis de la cirrosis hepática, la cual implica una reducción en el volumen del hígado. Sin embargo, los resultados obtenidos en la segmentación de los sujetos con esta patología no fueron óptimos en evaluaciones previas.

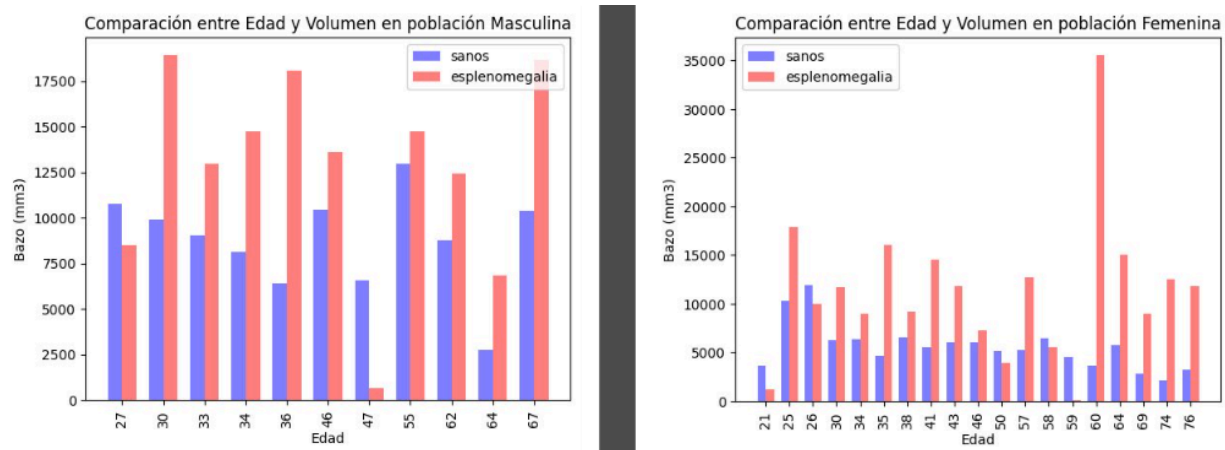
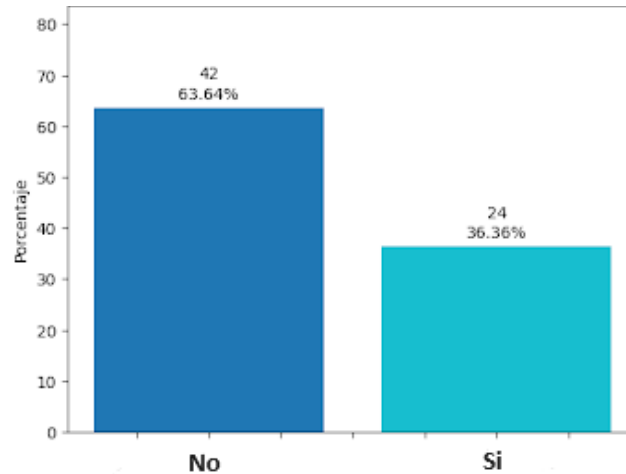


Fig 15. Prueba pareada para comparar algunos volúmenes entre sujetos sanos y con esplenomegalia

A pesar de que ya se había decidido mejorar la segmentación por parte del grupo de investigación, se encontró que alrededor del 36% de la población patológica presentaba errores en la segmentación tal como se observa en la figura 16. Este porcentaje es significativo, especialmente considerando que el modelo desarrollado tiene como objetivo identificar estas patologías como una herramienta de ayuda diagnóstica inicial. Sin una segmentación precisa en esta población, resulta prácticamente imposible darle utilidad a esta herramienta, lo que refuerza, desde un punto de vista ingenieril, la decisión médica previa de mejorar el modelo de segmentación.



*Fig.16 Análisis del porcentaje de sujetos patológicos con problemas de segmentación*

Al examinar los problemas de segmentación en la población patológica, identificamos patrones recurrentes. Por ejemplo, en sujetos con esplenomegalia, que implica un agrandamiento del bazo, el algoritmo tendía a identificar incorrectamente el bazo como si fuera el hígado. Este error resultaba en dificultades adicionales de segmentación en los riñones, ya que el algoritmo, al identificar el hígado en el lado izquierdo del abdomen, asumió que el riñón izquierdo era el riñón derecho, lo que ocasiona problemas de lateralidad en el abdomen, como se ilustra en la figura 17.

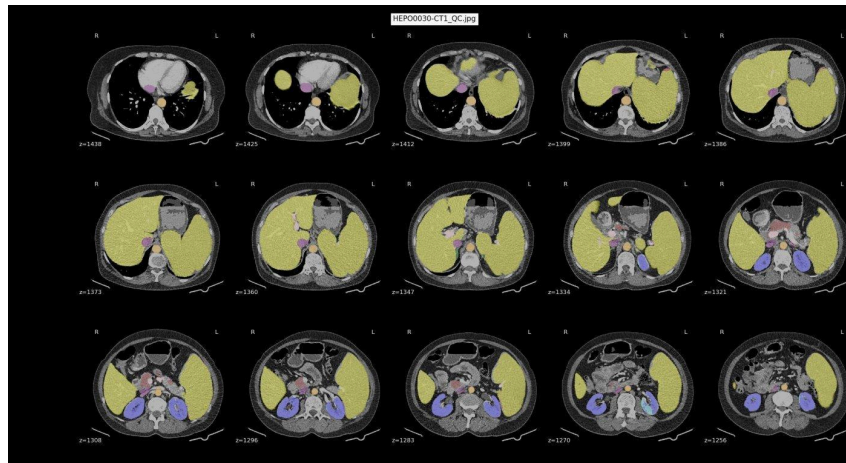


Fig. 17 QC de sujeto patológico con problemas en su segmentación

La exposición anterior se ve respaldada por un comportamiento inusual que también se observó de manera recurrente entre los sujetos con afecciones patológicas. Cuando la identificación errónea del bazo como hígado era parcial, la confusión en la lateralidad de los riñones también era parcial, como se puede apreciar en la figura 18.

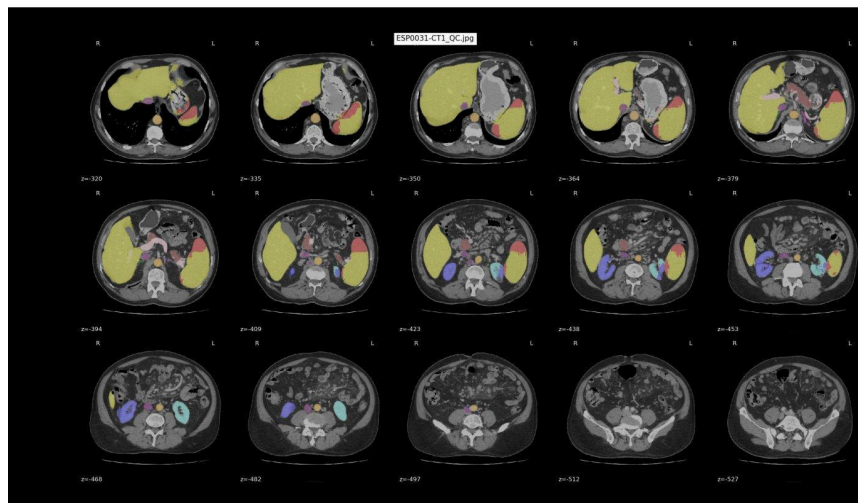
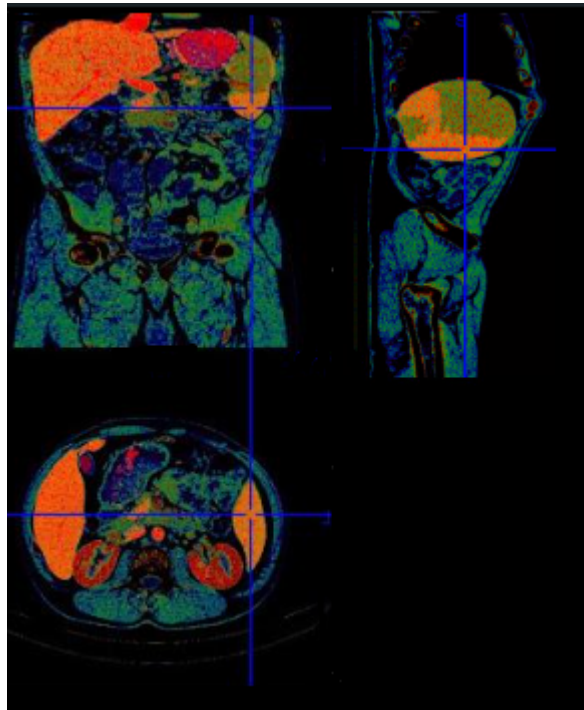


Fig. 18 QC de sujeto patológico con problema parcial en su segmentación

Después, al visualizar los datos en el visor 3D (Ver figura 19), se confirmaron las observaciones realizadas desde el Quality Check en estos sujetos patológicos. Esto permitió corroborar de manera más clara los problemas de segmentación identificados previamente.



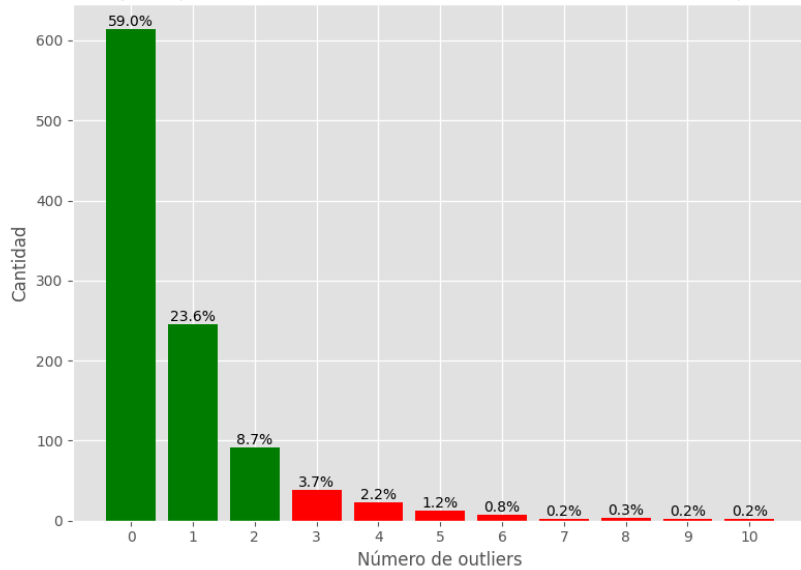
*Fig.19 Visualización en el visor MRicron de segmentación perteneciente a la población patológica*

### **3. Filtración de los datos para la construcción de un modelo normativo robusto.**

Los sujetos que presentaban menos de 3 valores atípicos, según el algoritmo automático de detección de outliers desarrollado en este estudio, fueron inicialmente considerados como ideales para conformar el modelo normativo, tanto para hombres como para mujeres. En particular, se empleó una técnica conocida como Desviación Absoluta de la Mediana (MAD) con un umbral establecido en 2. Luego de depurar los datos, se seleccionó el 91.3% de las segmentaciones correspondientes a la población masculina, equivalente a 951 muestras de sujetos sanos, y el 93.9% de la población femenina, lo que representó 977 muestras. Es relevante destacar que la mayoría de los individuos en ambas muestras no mostraron valores atípicos en ninguna de las estructuras evaluadas. Aunque se observaron algunas diferencias entre las dos

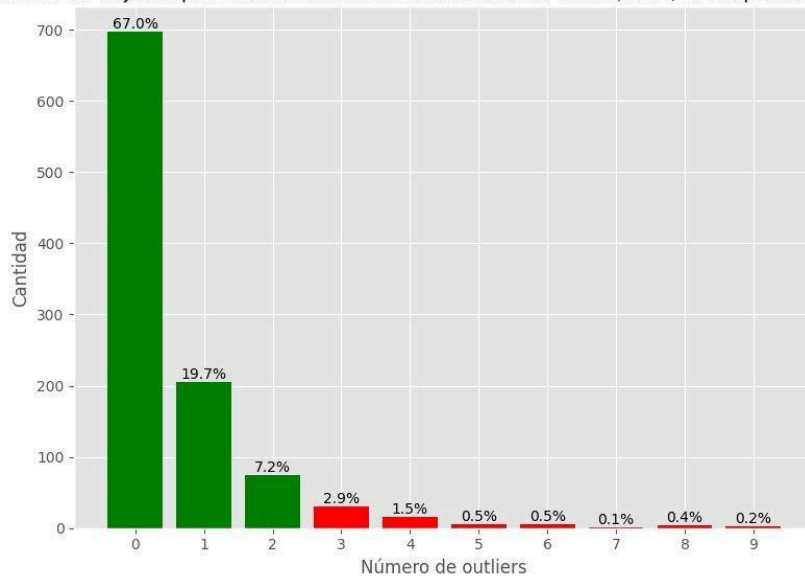
poblaciones, en general, la distribución de valores atípicos fue similar, ya que la mayoría de los sujetos en ambas poblaciones carecían de estos valores, como se puede apreciar en las figuras 20 y 21.

Distribución de sujetos por número de outliers utilizando Z-score(MAD) en la población Masculina



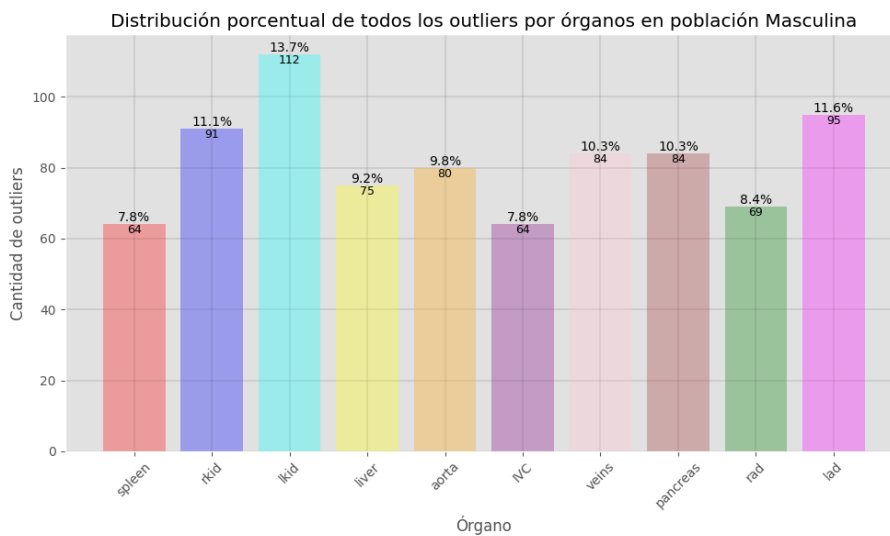
*Fig. 20 Distribución la población masculina según el número de outliers*

Distribución de sujetos por número de outliers utilizando Z-score(MAD) en la población Femenina

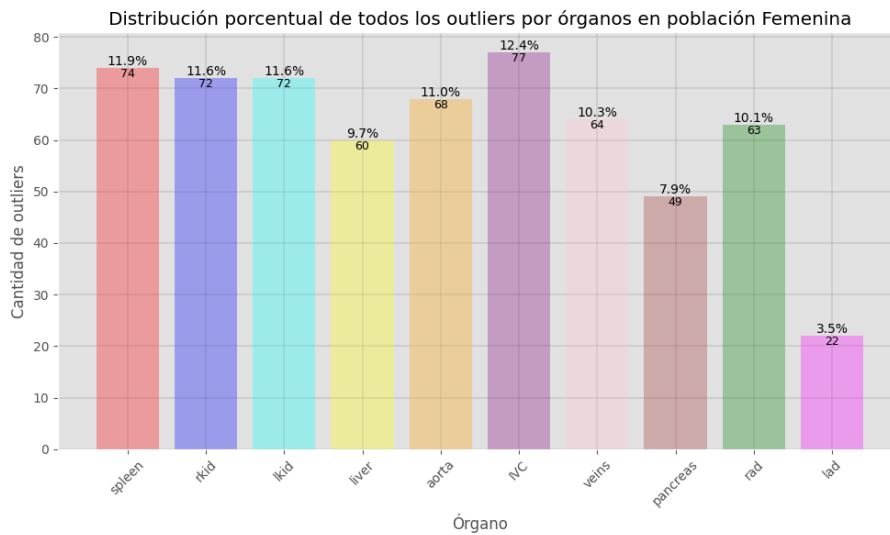


*Fig. 21 Distribución la población femenina según el número de outliers*

Además, se observó que en la población masculina, los riñones son uno de los órganos con mayor presencia de outliers, junto con la glándula suprarrenal izquierda. Es relevante señalar que el riñón izquierdo muestra la mayor cantidad de valores atípicos entre todas las otras estructuras, mientras que el bazo y la vena cava inferior exhiben la menor cantidad de valores atípicos (Fig. 22). Por otro lado, en la población femenina, aunque se registró una menor cantidad de outliers en general, resulta paradójico que los órganos que más presentan estos valores atípicos sean los mismos que mostraron menos en la población masculina, como el bazo y la vena cava inferior. Además, la glándula suprarrenal izquierda nuevamente, al contrario de lo observado en la población masculina, exhibió una baja presencia de outliers de manera destacable en comparación con el número de otros órganos(Fig. 23).



*Fig. 22 Distribución porcentual de outliers por órgano en población masculina*



*Fig. 23 Distribución porcentual de outliers por órgano en población femenina*

Además, como se mencionó anteriormente, el objetivo principal era mantener un equilibrio entre los distintos rangos de edad. Por lo tanto, se llevó a cabo un estudio detallado de la cantidad de sujetos dentro de cada rango de edad que presentaban menos de 3 outliers, ya que estos se utilizaron para construir un modelo normativo. Se encontró que hubo un equilibrio relativo en los rangos de edad de los sujetos utilizados en la implementación del modelo normativo, tanto en la población masculina (ver Tabla 3) como en la femenina (ver Tabla 4).

Rango de edad	0 outlier	1 outlier	2 outliers	Total
18-38 años	213	77	24	314
39-59 años	227	81	28	336
60-80 años	174	88	39	301
Total	614	246	91	951

*Tabla 3. Distribución de sujetos con los que se implementó el modelo normativo por rango de edad en la población masculina.*



<b>Rango de edad</b>	<b>0 outlier</b>	<b>1 outlier</b>	<b>2 outliers</b>	<b>Total</b>
18-38 años	230	69	22	321
39-59 años	249	61	23	333
60-80 años	218	75	30	323
Total	697	205	75	977

*Tabla 4. Distribución de sujetos con los que se implementó el modelo normativo por rango de edad en la población femenina.*

Antes de establecer la curva normativa, se llevó a cabo una evaluación exhaustiva de la distribución de los volúmenes en milímetros cúbicos, centrándose específicamente en el hígado (ver Figura 24 y Figura 25) y el bazo (ver Figura 26 y Figura 27), los únicos órganos para los cuales se desarrollaron las curvas normativas. El objetivo era determinar si los valores identificados como outliers por el algoritmo de detección automática estaban realmente significativamente separados del grueso de los datos, de acuerdo con el umbral predefinido en el algoritmo de detección automática de outliers desarrollado. Esta evaluación se realizó utilizando gráficos de boxplot en ambos géneros, lo que permitió una visualización clara de la dispersión de los datos y la identificación de posibles valores atípicos.

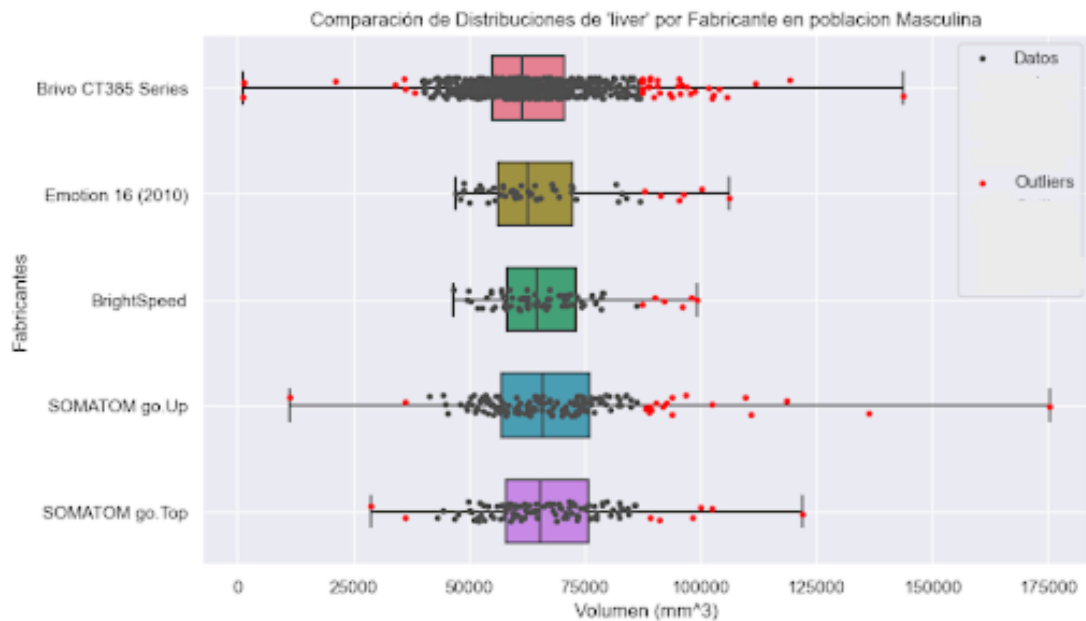


Fig.24 Distribución de los volúmenes de hígado en la población masculina.

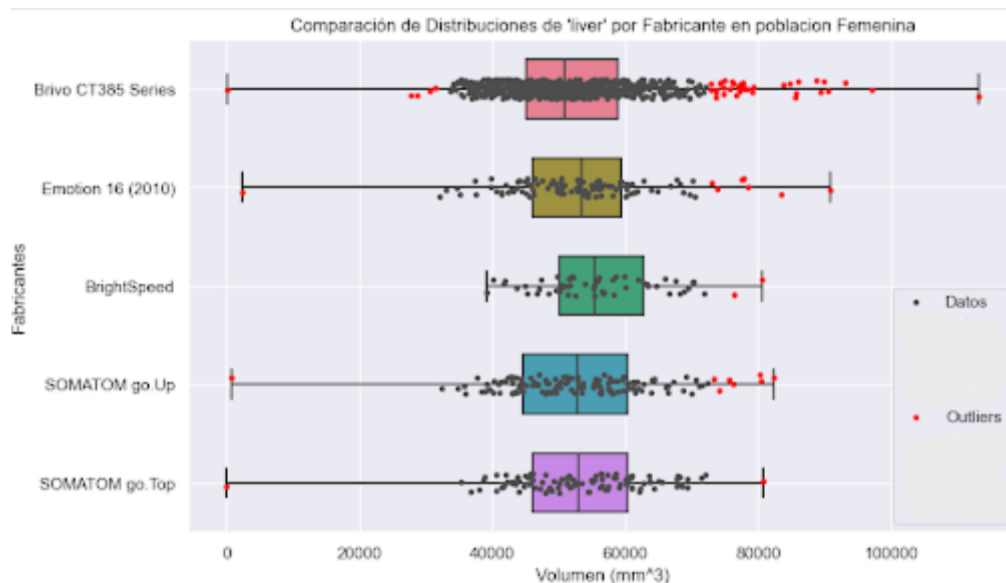


Fig.25 Distribución de los volúmenes de hígado en la población femenina.

Al visualizar las gráficas anteriores, se observa que, en términos generales, los volúmenes hepáticos en la población masculina son mayores, con una mayor variabilidad en general, así como una mediana y un rango intercuartílico (IQR) más altos. Además, se observa que el tomógrafo Brivo CT385 Series de GE Healthcare es el

equipo más utilizado en los estudios de pacientes sanos, tanto en hombres como en mujeres, lo que podría explicar la mayor cantidad de valores atípicos en los estudios realizados con estos equipos. Por otro lado, se encontró que el tomógrafo BrightSpeed, también de GE Healthcare, presentó menor variabilidad en los volúmenes hepáticos en ambos géneros.

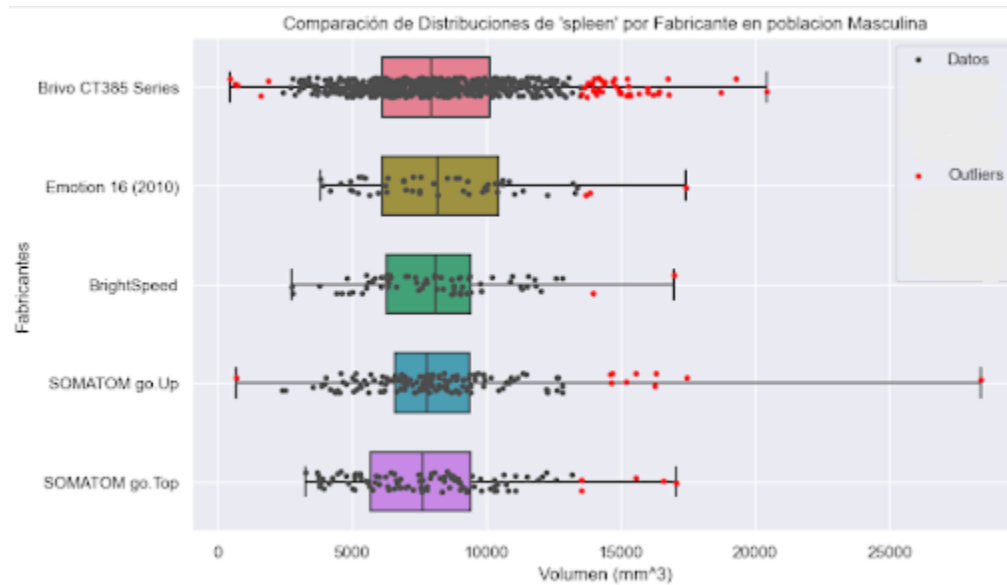


Fig.26 Distribución de los volúmenes de bazo en la población masculina.

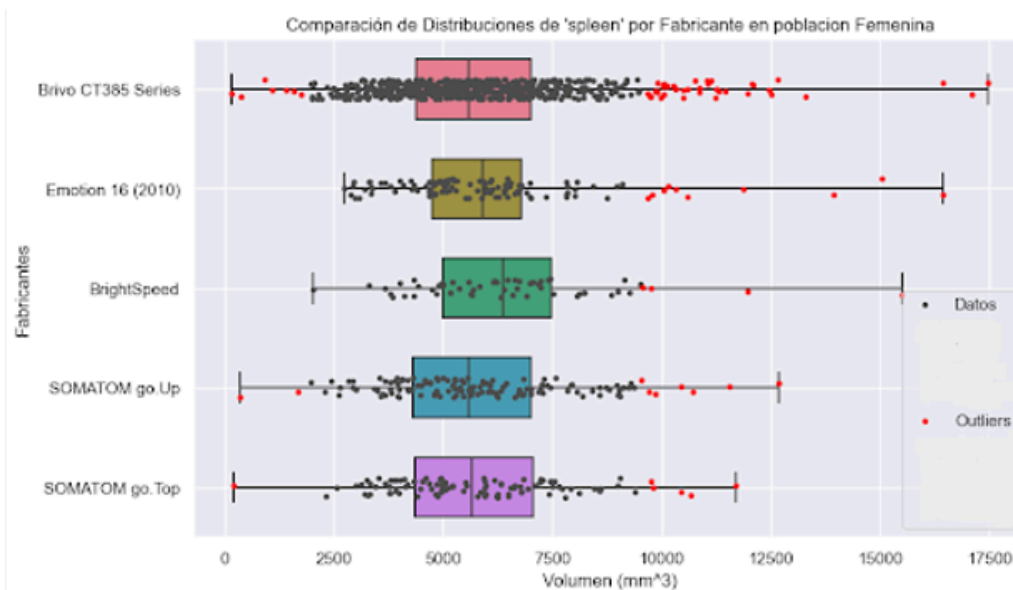
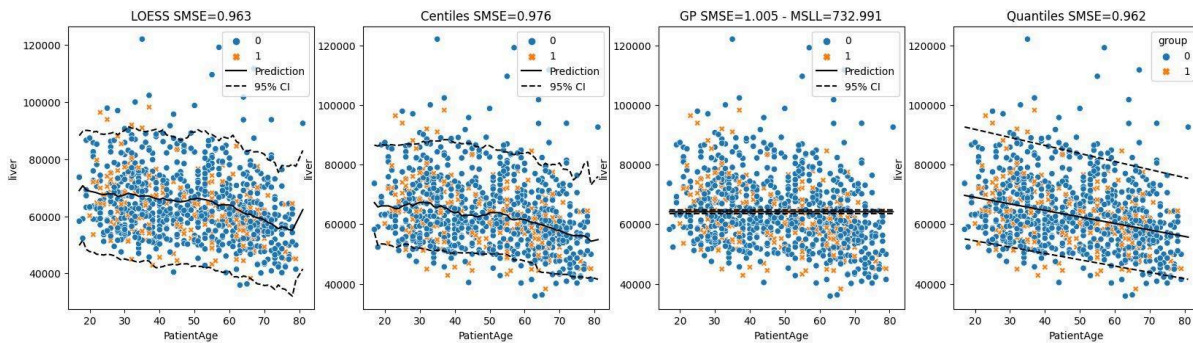


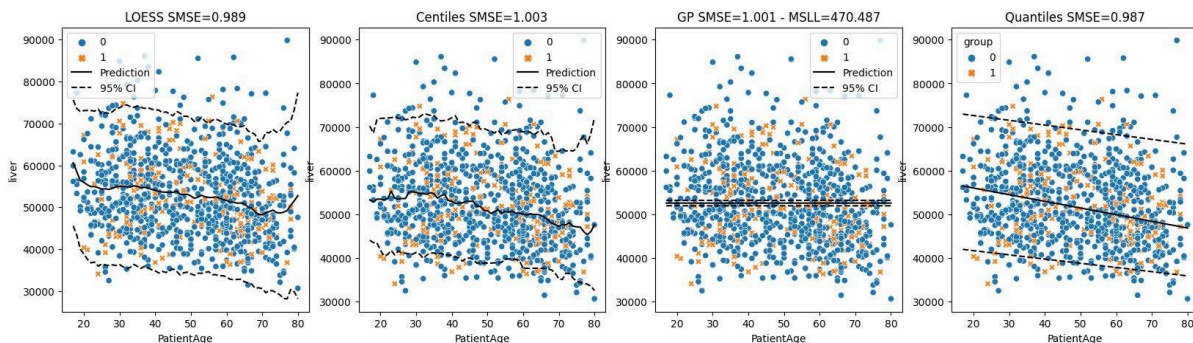
Fig. 27 Distribución de los volúmenes de bazo en la población femenina.

Al analizar los boxplot de los volúmenes del bazo, se evidencia, de manera similar a lo encontrado en el hígado, que en general las medianas de los volúmenes del bazo son menores en la población femenina que en la masculina. Asimismo, se observa una mayor variabilidad en los volúmenes del bazo en comparación con los del hígado, siendo esta disparidad más notable en la población femenina al examinar la distribución de los volúmenes obtenidos con el tomógrafo BrightSpeed.

Después de esta etapa, como se mencionó anteriormente, se dividieron los datos en conjuntos de entrenamiento y prueba para aplicar la validación cruzada y elaborar las curvas de regresión correspondientes tanto para la población masculina (véase Figura 28) como para la población femenina (véase Figura 29) al evaluar el modelo normativo del hígado. Tras un análisis exhaustivo, se optó finalmente por la regresión de cuantiles debido a que demostró tener un menor error en ambas poblaciones.

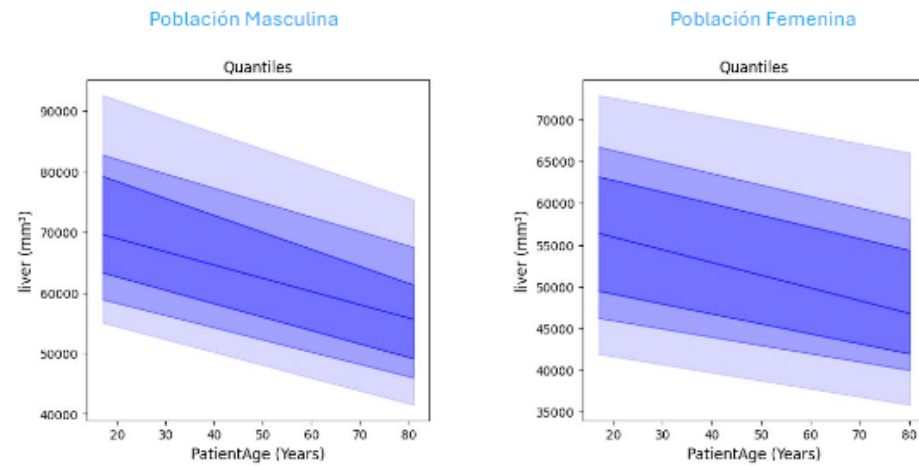


*Fig.28 Resultados comparativos de las regresiones con los volúmenes de hígado en población masculina*



*Fig.29 Resultados comparativos de las regresiones con los volúmenes de hígado en población femenina*

Se reitera la tendencia de que los volúmenes del hígado tienden a ser mayores en la población masculina. Además, se observa una disminución gradual en los volúmenes hepáticos con el aumento de la edad en ambos sexos, lo cual concuerda con hallazgos previos [5]. Este patrón se ilustra claramente en la Figura 30.



*Fig.30 Curvas normativas generadas para el hígado en ambas poblaciones*

De manera análoga, se llevó a cabo un análisis comparativo de los resultados de las regresiones para el bazo tanto en la población masculina (ver Figura 31) como en la población femenina (ver Figura 32). En consonancia con lo observado en el hígado, nuevamente en la población masculina, la regresión por cuantiles demostró tener los mejores resultados al exhibir un error menor. Sin embargo, en la población femenina, el modelo de regresión loess obtuvo el mejor rendimiento según la métrica de error cuadrático medio estandarizado. A pesar de esta discrepancia, tras evaluar los resultados de manera integral y considerando pruebas anteriores, se decidió construir la curva normativa basándose en la regresión por cuantiles, que en general fue la que mejor desempeño mostró a lo largo del proyecto.

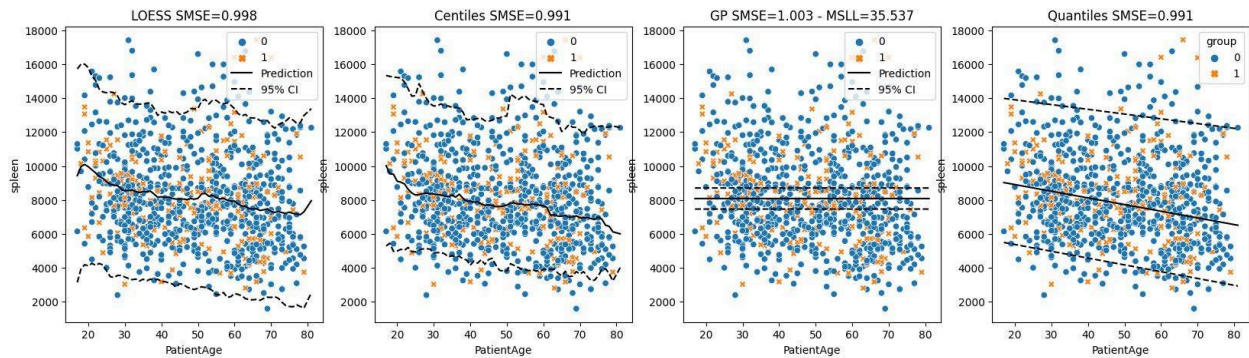


Fig.31 Resultados comparativos de las regresiones con los volúmenes de bazo en población masculina

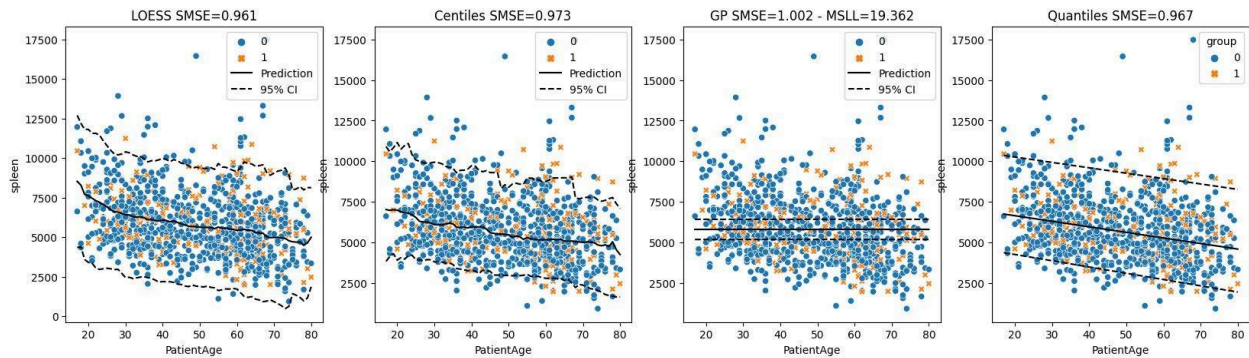


Fig.32 Resultados comparativos de las regresiones con los volúmenes de bazo en población femenina

Al revisar nuevamente las curvas normativas del bazo, es importante destacar la tendencia general de que los volúmenes de todos los órganos abdominales incluidos inicialmente en el modelo normativo tienden a ser mayores en la población masculina. Además, de manera similar a lo observado en el hígado, se aprecia una disminución gradual en los volúmenes esplénicos con el aumento de la edad en ambos sexos, un patrón que coincide con investigaciones previas [5]. Esto se evidencia claramente en la Figura 33.



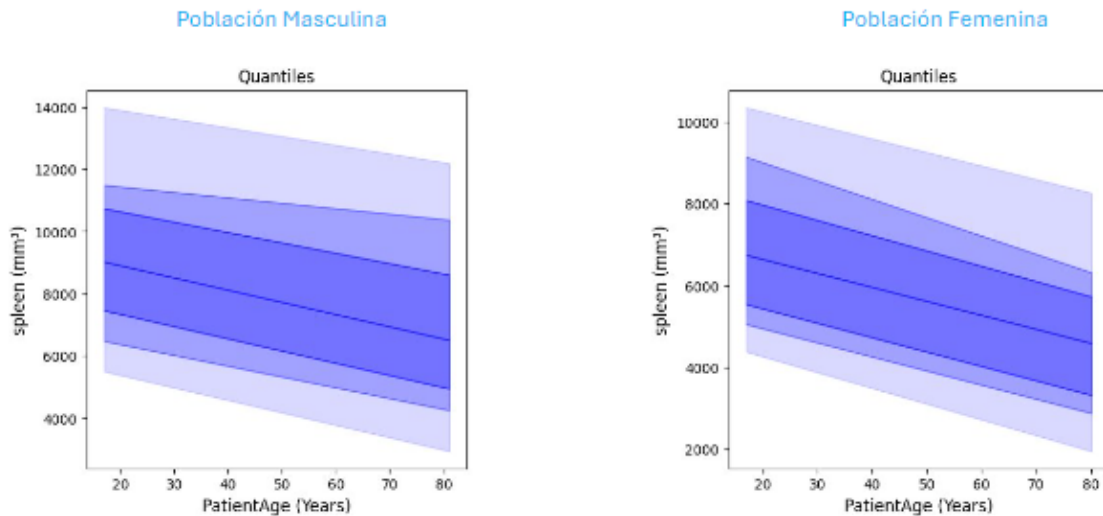


Fig.33 Curvas normativas generadas para el bazo en ambas poblaciones

Finalmente y con la intención de hacer una apreciación inicial de diferencias entre la población sana y la población con hepatomegalia, se implementó una prueba de delta de edad, construyendo la curva a partir de la regresión por cuantiles de los volúmenes de sujetos sanos con el fin de hacer una clasificación de los sujetos patológicos, en unas primeras pruebas se realizó la prueba de delta de edad teniendo solo en cuenta los volúmenes de hígado para cotejarlos con los sujetos con hepatomegalia tal como se observa en la figura 34

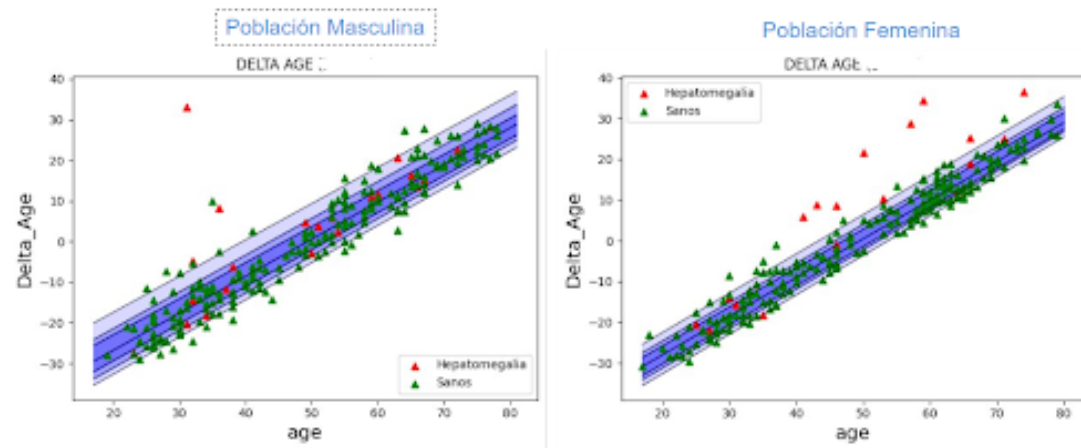


Fig.34 Prueba de delta de edad considerando el volumen del hígado para la predicción de la edad .

En estos resultados, al analizar la población femenina, se observa que una proporción considerable de pacientes con hepatomegalia y una proporción pequeña de sujetos sanos presentan valores atípicos en la prueba de delta de edad (ubicados fuera del rango entre el cuantil 5 y 95). Este hallazgo es notable, especialmente considerando que, durante el control de calidad de la segmentación, los resultados no fueron satisfactorios. Sin embargo, por otro lado, en la población masculina, se observan resultados algo diferentes, ya que la mayoría de los sujetos con hepatomegalia se encuentran dentro del rango de valores normales (entre el cuantil 5 y 95).

Así mismo se realizó la prueba de delta de edad a partir de los volúmenes obtenidos del bazo en la población sana para cotejarlos con sujetos con la condición patológica de esplenomegalia tal como se muestra en la figura 35



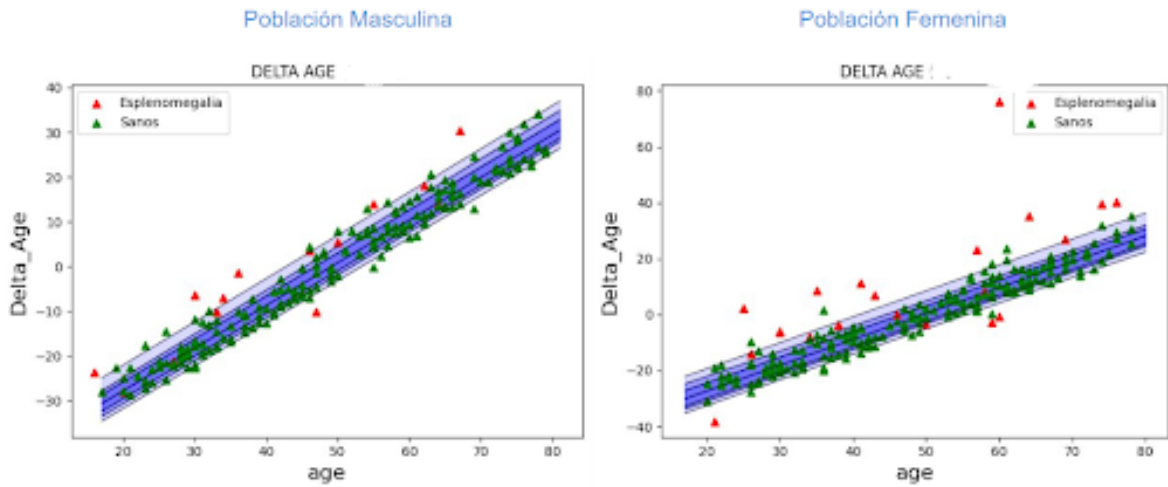


Fig.35 Prueba de delta de edad considerando el volumen del bazo para la predicción de la edad

En estos resultados, al analizar nuevamente la población femenina, se observa que una proporción considerable de pacientes con esplenomegalia exhiben valores atípicos en la prueba de delta de edad, situándose fuera del rango entre el cuantil 5 y 95. Por otro lado, en la población masculina, se observan resultados algo distintos en comparación con los obtenidos para la hepatomegalia. En este caso, una gran parte de los sujetos con esplenomegalia también se sitúan fuera del rango de valores normales (entre el cuantil 5 y 95); sin embargo, estos valores no tienden a alejarse tanto de la curva como sucede en la población femenina. Además, tanto en la población masculina como en la femenina, al analizar la predicción del delta de edad, los sujetos sanos de prueba no se alejan tanto de los valores normales como sucedió en algunos casos con los sanos al realizar la prueba análoga en el hígado.

Después, se llevó a cabo otra prueba de clasificación mediante el delta de edad, pero en esta ocasión se consideraron los volúmenes de los diez órganos inicialmente establecidos para el modelo normativo. En el caso de la comparación entre los sujetos sanos y los sujetos con hepatomegalia, se observó, como se muestra en la figura 36, que en ambas poblaciones existe una menor diferencia entre la edad predicha y la edad real del sujeto. Esto sugiere que es más conveniente considerar el volumen de los diez órganos al predecir la edad del individuo. Sin embargo, al realizar la clasificación dentro

de los valores normales, los sujetos patológicos no lograron diferenciarse tan claramente como quizás sucedió en la curva realizada a partir de los volúmenes de un solo órgano.

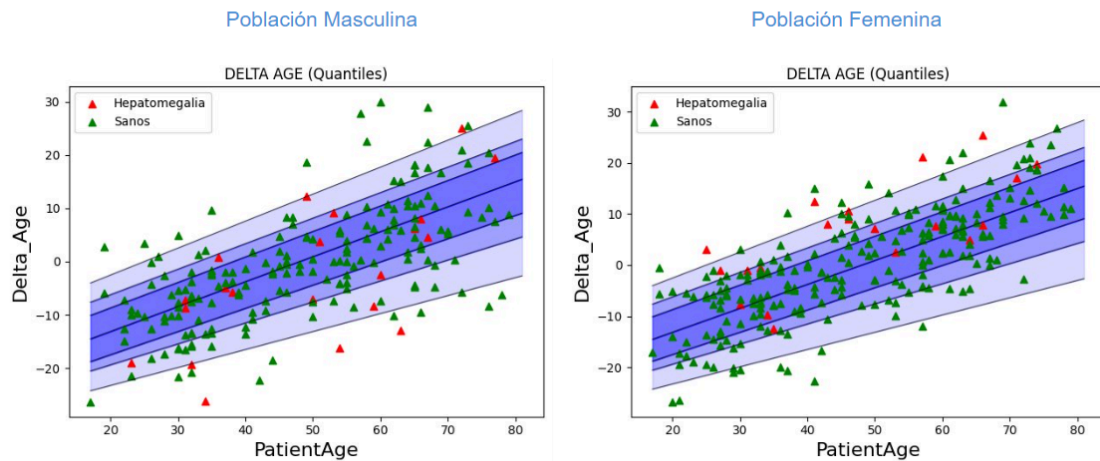


Fig.36 Prueba de delta de edad considerando los diez órganos para la predicción de la edad en sujetos sanos y con hepatomegalia

De manera similar ocurrió con los sujetos con esplenomegalia, como se ilustra en la figura 37. En ambas poblaciones, la clasificación de los sujetos patológicos como valores atípicos por fuera de la curva no es tan precisa. No obstante, se observa que el delta de edad, es decir, la diferencia entre la edad predicha y la edad real, es menor en comparación con la clasificación realizada anteriormente.

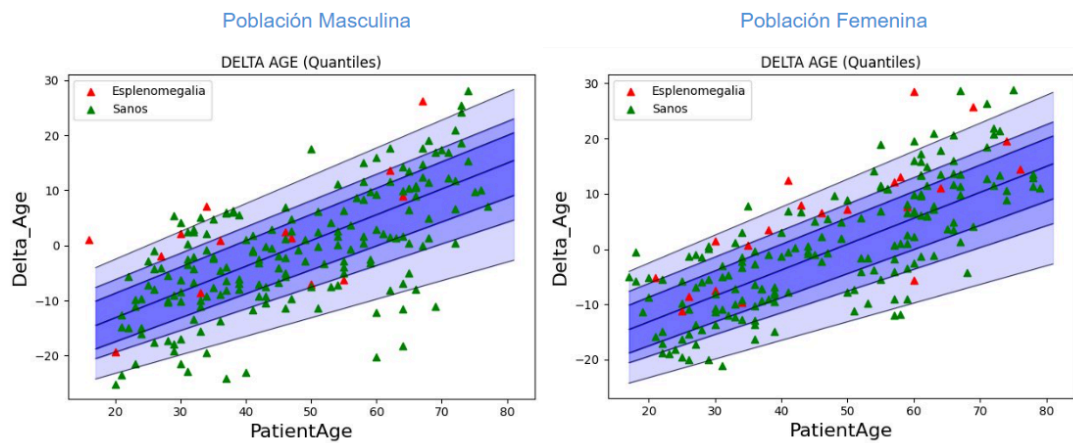


Fig.37 Prueba de delta de edad considerando los diez órganos para la predicción de la edad en sujetos sanos y con esplenomegalia.

---

## **VIII. DISCUSIÓN**

A lo largo del proyecto, la estrecha colaboración con profesionales médicos desempeñó un papel crucial al aportar un valor significativo a los resultados obtenidos. Esta interacción no solo enriqueció nuestros descubrimientos, sino que también complementa los conocimientos adquiridos en el ámbito de la bioingeniería. La sinergia entre el equipo médico y el equipo de bioingeniería, que posee un amplio conocimiento en el área de imágenes médicas, fortaleció nuestra comprensión de la disciplina y amplió nuestro horizonte de conocimientos. Este vínculo estrecho facilitó la implementación de un riguroso control de calidad tanto en los datos recopilados como en el proceso de segmentación, asegurando la robustez y la fiabilidad de nuestros hallazgos.

Durante el transcurso del proyecto, al realizar el análisis exploratorio de los datos, se detectaron errores en la recopilación inicial que habían pasado desapercibidos. Se encontraron sujetos catalogados como sanos que, de hecho, presentaban patologías o habían sido sometidos a procedimientos quirúrgicos como nefrectomía y pancreatoclectomía, como se ilustra en la Figura 2. Esta omisión de órganos específicos en nuestro modelo normativo debía ser abordada, ya que podría tener un impacto significativo en el análisis. La ausencia de datos de volumen de estos órganos en los sujetos afectados podría introducir sesgos en la creación de modelos normativos específicos para cada órgano.

Al comparar la validación de la primera consulta generada en Elasticsearch-findsca con la segunda (Figura 38), se destaca la eficacia de utilizar informes estructurados para realizar consultas en estas herramientas. Nos concentramos específicamente en las conclusiones de estos informes médicos, lo que simplifica el proceso de selección basado en los criterios de inclusión y exclusión previamente establecidos. Este enfoque no solo enriqueció el proceso de consulta de datos para el proyecto, sino que también ha tenido un impacto positivo en la empresa. Desde esta metodología, se ha comenzado a utilizar para la consulta de datos en otros estudios, lo que ha contribuido a mejorar la eficiencia en la búsqueda de información.

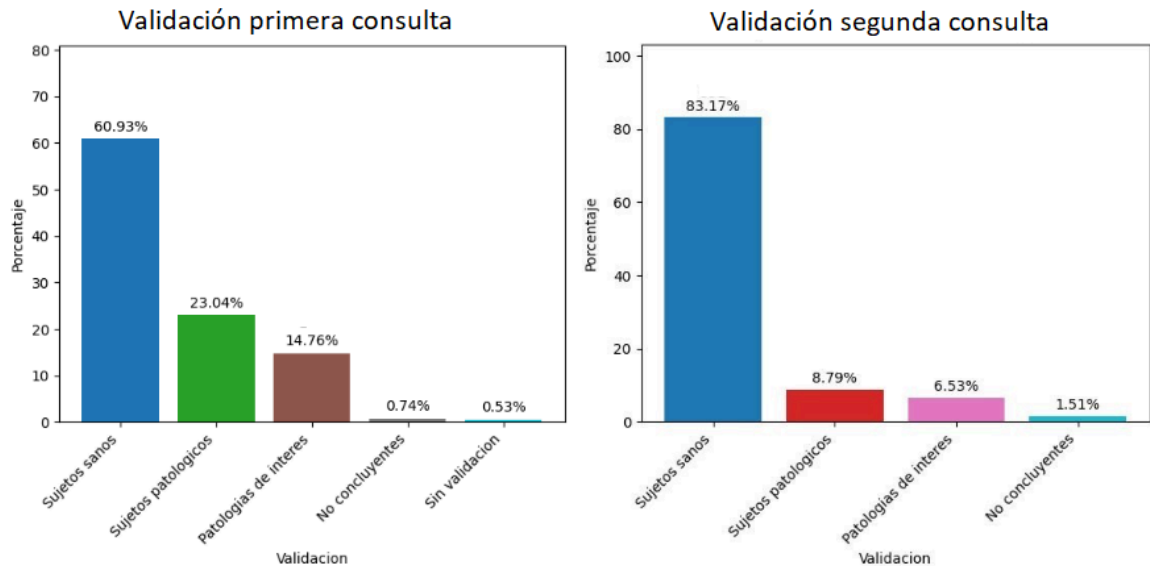


Fig. 38 Comparación porcentual de los resultados de la validación de sujetos catalogados como sanos en ambas consultas.

Por otro lado, al llevar a cabo la reclasificación de los sujetos patológicos, nos enfrentamos a ciertas dificultades que merecen atención. Una de ellas es la variabilidad en la forma de registrar las patologías, como por ejemplo, la discrepancia entre el término "aterosclerosis" y "ateromatosis aórtica calcificada" en diferentes informes médicos. A pesar de estas diferencias, es crucial destacar que, desde el punto de vista médico, se refieren a la misma enfermedad. Este aspecto subraya la importancia de establecer criterios claros y uniformes para la clasificación de enfermedades en los informes médicos, lo que facilita la interpretación y comparación de datos en estudios futuros. Además, considerando el interés de la compañía en implementar herramientas de inteligencia artificial, comprender el funcionamiento de estos algoritmos es fundamental para aprovechar al máximo su potencial y ofrecer soluciones efectivas a los usuarios.

Debido a la naturaleza subjetiva del diagnóstico de la aterosclerosis aórtica y a la necesidad de una presencia significativa de placas para que esta afecte al análisis volumétrico de la aorta, según los criterios médicos establecidos, se decidió no

---

profundizar en esta patología en esta etapa del proyecto. Como resultado, los casos de aterosclerosis aórtica no fueron incorporados en la plataforma de procesamiento Xnat. No obstante, se identificaron individuos con esta afección para su posible inclusión en etapas posteriores del proyecto, una vez que se disponga de bases médicas más sólidas y criterios de inclusión más claros.

Al abordar el tema de la segmentación a través del Quality Check, especialmente en relación con la alta cantidad de segmentaciones del páncreas evaluadas por el médico radiólogo como regulares o deficientes, es fundamental resaltar las principales razones detrás de estas dificultades. En primer lugar, es frecuente encontrar imágenes con escaso contraste y contornos poco definidos en la región del páncreas. Esto dificulta la tarea de delimitar con precisión esta estructura durante el proceso de segmentación. Además, se observan considerables variaciones en la escala de grises y la textura de las imágenes, lo que complica aún más la identificación automática del páncreas por parte del algoritmo de inteligencia artificial.

Investigaciones previas han destacado una amplia variabilidad en la ubicación, forma y tamaño del páncreas entre diferentes individuos. Estas variaciones anatómicas añaden una capa adicional de complejidad a la tarea de segmentación, ya que el algoritmo debe ser capaz de ajustarse a estas diferencias y reconocer la estructura del páncreas de manera precisa y consistente en todas las imágenes [58].

Al estudiar la segmentación a través del visor en 3D, se identificaron varios patrones preocupantes. Entre ellos, se observaron confusiones en los contornos del hígado y la identificación incorrecta de estructuras que se superponen con el hígado, como las venas suprahepáticas. Además, los contornos del hígado no estaban claramente definidos en la mayoría de los casos. También se detectaron situaciones donde partes de las extremidades inferiores eran erróneamente identificadas como parte del hígado por el algoritmo.

Es fundamental destacar que este modelo de segmentación está destinado a ser implementado en los servicios de la compañía, lo que subraya la necesidad de mantener altos estándares de precisión en estos casos. Estos hallazgos fueron decisivos para que el equipo de investigación optó por centrarse en mejorar la

---

segmentación en futuros trabajos. Esto se debe a la relevancia del hígado en nuestro modelo normativo y la importancia de garantizar su correcta identificación y delimitación en las imágenes médicas.

En los hallazgos al examinar las distintas segmentaciones de la población patológica, nos encontramos con patrones interesantes de lateralidad. Esto sugiere especialmente que el algoritmo se basa en la posición relativa de los órganos en el abdomen para determinar su identidad. Por lo tanto, es crucial tener precaución y comprender la variabilidad anatómica que puede existir entre los sujetos, ya que esto puede influir significativamente en la precisión de la segmentación automática de los órganos abdominales. Estos resultados refuerzan la idea de la necesidad de mejoras en la segmentación para garantizar una identificación más precisa y consistente de los órganos en futuros proyectos y aplicaciones clínicas.

Durante la implementación del flujo de trabajo que incorpora el algoritmo para la detección automática de outliers junto con los modelos normativos, se optó por utilizar el método del Z-score modificado. Este enfoque, que se basa en la mediana en lugar de la media, fue elegido debido a las limitaciones del Z-score tradicional cuando hay presencia de outliers en la muestra. El Z-score tradicional, que se apoya en la media y la desviación estándar, puede experimentar dificultades para detectar outliers menos extremos debido a la presencia de outliers más extremos, un fenómeno conocido como enmascaramiento [40].

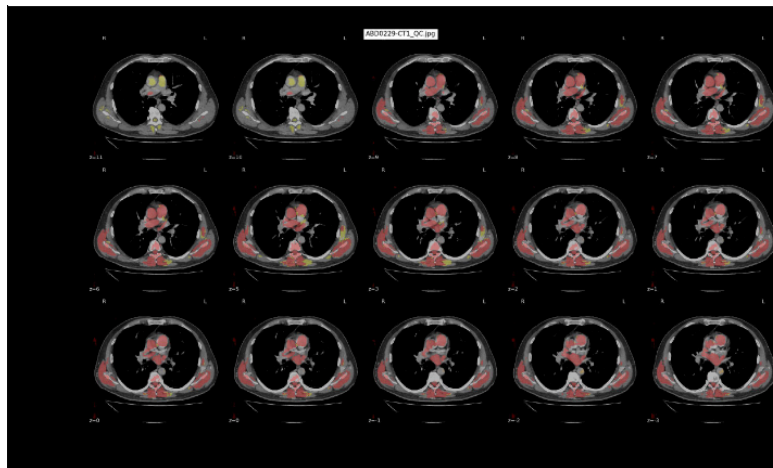
El método del Z-score modificado ofrece una mayor robustez frente a la presencia de outliers, lo que lo hace adecuado para su implementación en casos donde los datos no siguen una distribución normal, como las distribuciones t de Student y Poisson. Esta elección estratégica mejora la capacidad del algoritmo para identificar de manera precisa los outliers en conjuntos de datos que pueden contener valores atípicos [40].

Además, la elección de la Mediana Absoluta de las Desviaciones (MAD) en el contexto de la segmentación de volúmenes de órganos presenta diversas ventajas. Esta técnica exhibe una mayor robustez frente a valores extremos, lo que la hace menos sensible a estos en comparación con otras métricas [59].

---

La importancia de esta cualidad se hace evidente en nuestro proyecto de segmentación abdominal, donde incluso ante una posible aprobación por parte de profesionales de la salud, es crucial tener en cuenta que la perfección en la segmentación es inalcanzable, como se ha destacado anteriormente. La presencia de valores anómalos podría distorsionar la interpretación de los resultados, especialmente si no siguen una distribución normal. Sin embargo, las diferencias absolutas entre los valores y la mediana minimizan el impacto de los valores extremos, que podrían afectar significativamente otros métodos basados en la media y la desviación estándar. Esta capacidad para detectar valores atípicos proporciona una mayor confiabilidad sin necesidad de hacer suposiciones sobre la distribución de los datos. La eficacia de este enfoque ha sido respaldada por investigaciones previas similares, especialmente en análisis multivariantes, lo que confiere una mayor validez al modelo normativo [60].

Es crucial resaltar que los resultados obtenidos mediante la aplicación del algoritmo de detección automática de outliers, tanto en términos del recuento de outliers en ambas poblaciones como en la distribución de outliers por órgano, si bien informativos, no son completamente concluyentes. Más bien, han servido para identificar segmentaciones con mayores inconsistencias. Las diferencias anatómicas entre individuos pueden dar lugar a variaciones en la forma y el tamaño de los órganos, lo que, a su vez, puede afectar la detección de outliers. Además, las características de las imágenes médicas, como la calidad de la imagen, el contraste y la resolución, pueden influir en la precisión de la segmentación y, por ende, en la identificación de valores atípicos. Por ejemplo, en la Figura 39 se observa una notable cantidad de valores atípicos debido a errores en la adquisición de imágenes durante el estudio. Por lo tanto, es esencial tener en cuenta que ciertos algoritmos de segmentación pueden presentar limitaciones en la detección de ciertos tipos de estructuras o en condiciones específicas.



*Fig.39 Sujeto identificado con outliers en 10 órganos*

Al examinar los resultados de las pruebas de delta de edad, se cuestiona la fiabilidad de cualquier clasificación realizada por el algoritmo en los sujetos patológicos. Como se ha demostrado a lo largo del estudio, unos resultados deficientes en la segmentación pueden generar sesgos significativos. Por ejemplo, se observó que en algunos casos el algoritmo confunde el bazo con el hígado, tratándolos como una única estructura, lo que resulta en un aumento considerable en los volúmenes de esa área específica. Este error tiene un impacto directo en los resultados, especialmente considerando que la predicción depende en gran medida del volumen de las estructuras analizadas.

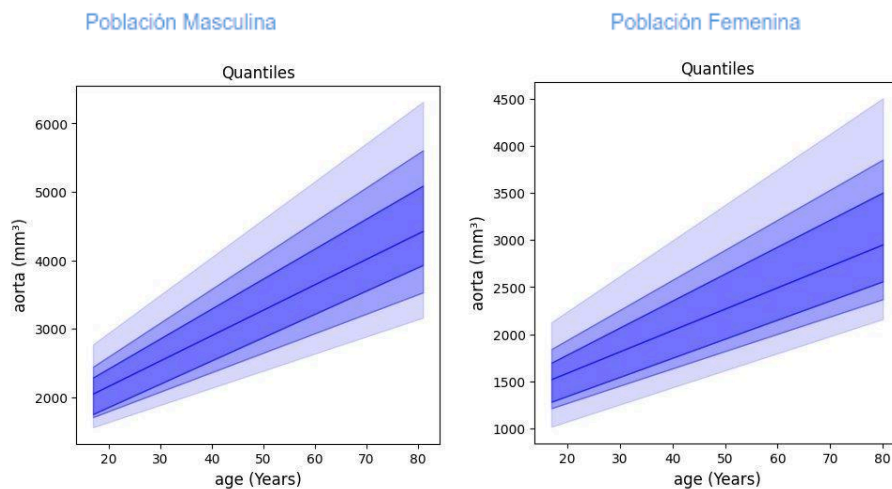
Aunque la segmentación de la arteria aorta arrojó resultados muy satisfactorios en ambos procesos de validación: tanto en el Quality Check como en la visualización completa de la segmentación. Este hallazgo concuerda con investigaciones previas que resaltan la facilidad de identificación de la arteria aorta mediante tomografía computarizada [61]. Además, nuestros resultados se alinean con estudios anteriores que han utilizado redes neuronales UNET para la segmentación de la arteria aorta, aunque con la distinción de que nuestro enfoque se centró en la segmentación de múltiples órganos [62].

Es notable que, aunque la arteria aorta no fue considerada en el modelo de delta de edad debido a consideraciones médicas, sin embargo una revisión de la literatura



reveló un comportamiento peculiar en esta estructura en relación con el envejecimiento. Se encontró que, a diferencia de otras estructuras no vasculares, el diámetro de la arteria aorta tiende a aumentar con la edad en la población sana [63]. Este hallazgo se alinea con estudios previos que indican un incremento significativo en el diámetro del istmo aórtico con el envejecimiento, además de los resultados de la curva normativa de esta estructura como se ve en la figura 40.

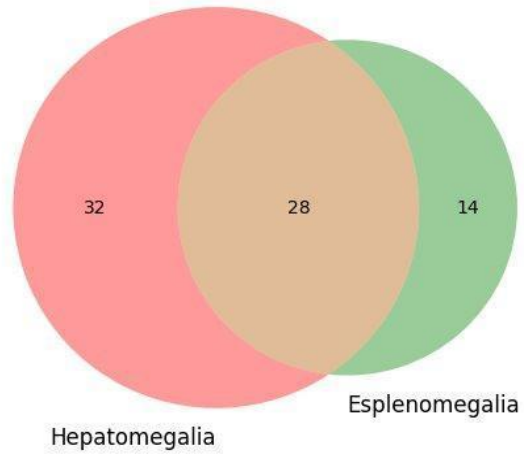
El istmo aórtico, situado aproximadamente a nivel de las undécima y duodécima vértebras torácicas, reviste una importancia anatómica crucial. Representa el punto donde la aorta realiza un descenso desde el arco aórtico hacia la región abdominal, marcando así la transición entre la porción torácica y la abdominal de este vaso sanguíneo [61]. Este hallazgo subraya la necesidad de considerar no sólo las estructuras específicas incluidas en el modelo de delta de edad, sino también aquellas con características particulares de cambio asociadas con el envejecimiento. Esta consideración puede mejorar el pronóstico del delta de edad al utilizar el volumen de varios órganos.



*Fig.40 Curva normativa de la arteria aorta*

Por otro lado, resulta notable destacar que al examinar los informes médicos de los sujetos patológicos, se encontró que aproximadamente el 38% de la muestra total presenta hepatoesplenomegalia. Esta coexistencia de agrandamiento tanto del hígado como del bazo es relativamente común en la población estudiada. Este hallazgo sugiere

la importancia de estudiar la relación y posibles implicaciones clínicas de la presencia simultánea de ambas condiciones médicas.



*Fig.41 Relación encontrada entre la población con esplenomegalia y hepatomegalia*

## **IX. CONCLUSIONES**

- El descubrimiento de sujetos patológicos que previamente se habían considerado sanos en el proyecto subraya la importancia del análisis exploratorio de datos. Sin este análisis meticuloso, habría sido difícil detectar tales errores. La depuración de los datos realizada facilitará trabajos futuros con tomografías de abdomen contrastadas, ya que, a partir de valoraciones médicas, se tiene la certeza de que la población de sujetos considerados sanos realmente lo son, evitando sesgos. Con un conjunto de datos adecuado para el reentrenamiento del modelo, generado a partir de la validación médica, será posible abordar la segmentación de las imágenes con mayor precisión y confianza.

- Es esencial mejorar el proceso de segmentación, ya que, durante la validación realizada por el equipo médico, se identificaron errores significativos. Estos errores abarcaron desde la inclusión incorrecta de las venas suprahepáticas en la etiqueta de las venas porta y esplénica hasta inconsistencias persistentes en la delineación de órganos como el páncreas y el hígado. La optimización de la precisión en la segmentación se vuelve crucial para garantizar resultados más fiables y coherentes en el análisis de imágenes médicas.

- Es crucial señalar que la curva de pérdida no garantiza exclusivamente la precisión del modelo. En ocasiones, la pérdida puede disminuir, pero el modelo sigue generando segmentaciones incorrectas. Esto puede atribuirse a diversos factores, como la complejidad de los datos, la naturaleza específica del problema de segmentación o incluso posibles errores en el código del modelo. En nuestro caso particular, la cantidad limitada de tomografías utilizadas para el entrenamiento del modelo podría contribuir a este fenómeno, donde, a pesar de una aparente mejora según la curva de pérdida (fig. 5), persisten errores notables al visualizar las segmentaciones.

- La implementación del control de calidad (QC) en XNAT ha desempeñado un papel fundamental en la preservación de la coherencia en los resultados a lo largo del tiempo y entre diferentes conjuntos de datos, tanto en sujetos sanos como en aquellos con patologías. Este aspecto ha sido crucial para asegurar la robustez del modelo en un corto período de tiempo y su capacidad para generalizar de manera efectiva, lo que representa un logro significativo en términos de eficiencia al validar la segmentación. Como resultado, se ha convertido en una herramienta invaluable para los profesionales de la salud, brindando una base sólida para análisis clínicos precisos y confiables.

- Se observó que los problemas de segmentación pueden ser más notorios en la población patológica, lo que nos motiva a llevar a cabo un análisis más detallado de estas inconsistencias. Esto nos permite anticipar posibles problemas en la segmentación no solo de tomografías de sujetos patológicos, sino también de sujetos sanos. Este enfoque continuo nos capacita para mejorar constantemente la calidad del modelo desde una perspectiva ingenieril más profunda.

- La validación realizada por expertos en la materia se convierte en un recurso invaluable en el proceso de control de calidad. Este procedimiento no solo garantiza la precisión anatómica de las segmentaciones de acuerdo con los estándares médicos o científicos, sino que también adquiere una relevancia crítica en proyectos, especialmente en entornos regulados o investigaciones médicas. Esto asegura el cumplimiento de estándares y regulaciones específicas, añadiendo valor tanto al proyecto de investigación como a la completa integridad y validez de los resultados obtenidos. Sin embargo, es importante destacar que el tiempo de los profesionales de la salud dedicado a la investigación es limitado debido a sus responsabilidades en la prestación de servicios. A lo largo del proyecto, nos enfrentamos a este desafío y buscamos manejar eficientemente esta restricción.

- 
- Se diseñó el flujo de trabajo con la capacidad de realizar un control de calidad automático en las imágenes y filtrar los datos mediante un algoritmo desarrollado para la detección automática de valores atípicos. Aunque no se lograron construir modelos normativos definitivos, se avanzó significativamente en cómo abordar la depuración de los datos para la construcción de un modelo normativo robusto en el futuro, particularmente tras las mejoras futuras esperadas en las segmentaciones obtenidas por el modelo.

---

## **X. RECOMENDACIONES**

La recomendación clave para mejorar los resultados de la segmentación y lograr una mayor consistencia es aumentar significativamente el conjunto de datos de entrenamiento. Se sugiere entrenar el mismo modelo o, posiblemente, explorar otros modelos con una cantidad mucho mayor de imágenes. Idealmente, se deberían utilizar imágenes de tomografía abdominal provenientes del repositorio SURA, que cuenta con una amplia base de datos de imágenes médicas. Esta fuente de datos ofrece una oportunidad invaluable que aún no se ha aprovechado completamente y podría ser fundamental para mejorar la precisión y la robustez de los modelos de segmentación utilizados.

---

## REFERENCIAS

- [1] Litjens, Geert, et al. "A survey on deep learning in medical image analysis." *Medical image analysis* 42 (2017): 60-88.
- [2] O. Giannou and O. Giannou, "Automated distinction of neoplastic from healthy liver parenchyma based on machine learning," *Neural Comput Appl*, vol. 34, no. 22, pp. 19629–19638, 2022, doi: 10.1007/s00521-022-07599-2.
- [3] Expósito Gallardo, M. D. C., & Ávila Ávila, R. (2008). *Aplicaciones de la inteligencia artificial en la Medicina: perspectivas y problemas*. *Acimed*, 17(5), 0-0.
- [4] Kline, T. L., Korfiatis, P., Edwards, M. E., Warner, J. D., Irazabal, M. V., King, B. F., ... & Erickson, B. J. (2016). Automatic total kidney volume measurement on follow-up magnetic resonance images to facilitate monitoring of autosomal dominant polycystic kidney disease progression. *Nephrology Dialysis Transplantation*, 31(2), 241-248.
- [5] D. M. Felipe Salech, D. L. Rafael Jara, and D. A. Luis Michea, "CAMBIOS FISIOLÓGICOS ASOCIADOS AL ENVEJECIMIENTO PHYSIOLOGICAL CHANGES ASSOCIATED WITH NORMAL AGING," 2012.
- [6] Ramírez Giraldo, J. C., Arboleda Clavijo, C., & McCollough, C. H. (2008). *Tomografía computarizada por rayos X: fundamentos y actualidad*. *Revista Ingeniería Biomédica*, 2(4), 54-66.
- [7] Ceruto Marrero, G., Ramírez Pérez, J. F., Castro Limones, A. N., & Pérez Delgado, B. (2018). *Componente informático para el cálculo de volumen en imágenes de tomografía axial computarizada*. *Cuba: III Convención Internacional de Salud Pública, Cuba Salud*.
- [8] Dubey, Y., Mange, P., Barapatre, Y., Sable, B., Palsodkar, P., & Umate, R. (2023). *Unlocking Precision Medicine for Prognosis of Chronic Kidney Disease Using Machine Learning*. *Diagnostics*, 13(19), 3151.

- 
- [9] C. Chen et al., "Deep Learning for Cardiac Image Segmentation: A Review," *Front Cardiovasc Med*, vol. 7, no. March 2020, doi: 10.3389/fcvm.2020.00025.
- [10] Pava Meriño, K. M. (2023). *Implementación de un modelo normativo de volumetría corporal en pacientes SURA mediante algoritmos de segmentación automática utilizando imágenes de tomografía computarizada.*
- [11] Schiappacasse, G., Maldonado, I., Varela, C., Müller, K., & Ortiz, J. (2014). Síndrome del ligamento arcuato medio. *Revista chilena de radiología*, 20(4), 149-155.
- [12] Kim, T. M., Choj, S. J., Ko, J. Y., Kim, S., Jeong, C. W., Cho, J. Y., ... & Kim, Y. G. (2023). Fully automatic volume measurement of the adrenal gland on CT using deep learning to classify adrenal hyperplasia. *European Radiology*, 33(6), 4292-4302
- [13] V. de Padua V. Alves et al., "Computed tomography-based measurements of normative liver and spleen volumes in children," *Pediatr Radiol*, pp. 378–386, 2022, doi: 10.1007/s00247-022-05551-z.
- [14] Ramírez, M. E. (2018). Cáncer de mama. *Revista Médica Sinergia*, 2(01), 8-12.
- [15] Verdugo, M. A. (2004). Tomografía computada multicorte. *Revista chilena de cirugía*, 56, 185-190
- [16] Alonso, R. C., Vázquez, A. Á., Vázquez, C. A., Vasallo, I. T., Alfonso, A. F., Rodríguez, M. R., & de Vega, V. M. (2023). Tomografía computarizada con energía espectral en la diferenciación de los adenomas y metástasis suprarrenales: utilidad de los mapas de descomposición de materiales y de imágenes monocromáticas. *Radiología*, 65(5), 402-413.
- [17] Sellar-Pérez, G., Herrera-Gutiérrez, M. E., Lebrón-Gallardo, M., & Quesada-García, G. (2009). Planteamientos generales para el mantenimiento del donante de órganos. *Medicina intensiva*, 33(5), 235-242.
- [18] J. Bozek, L. Griffanti, S. Lau, and M. Jenkinson, "NeuroImage Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation



- 
- criteria,” Neuroimage, vol. 268, no. January, p. 119864, 2023, doi: 10.1016/j.neuroimage.2023.119864.*
- [19] Schrier, R. W., Arroyo, V., Bernardi, M., Epstein, M., Henriksen, J. H., & Rodés, J. (1988). *Peripheral arterial vasodilation hypothesis: a proposal for the initiation of renal sodium and water retention in cirrhosis. Hepatology, 8(5), 1151-1157.*
- [20] M. Takahashi et al., “Automated volume measurement of abdominal adipose tissue from entire abdominal cavity in Dixon MR images using deep learning,” *Radiol Phys Technol, vol. 16, no. 1, pp. 28–38, 2022, doi: 10.1007/s12194-022-00687-x.*
- [21] Rotella, J. M., Urpi, J., Heredia, M., & Brahim, C. (2009). *El hombro: una nueva visión de su evolución. Revista de la Asociación Argentina de Ortopedia y Traumatología, 74(4), 390-407.*
- [22] Álvarez, M. C., González, M., Massigoge, A., Kaufmann, C., & Gutiérrez, M. (2010). *La densidad mineral ósea y la variabilidad ontogénica en guanaco (Lama guanicoe). Implicancias para la construcción de marcos de referencia en Zooarqueología. Zooarqueología a principios del siglo XXI: Aportes teóricos, metodológicos y casos de estudio, 95-106.*
- [23] Gómez Vargas, E., Obregón Neira, N., & Rocha Arango, D. F. (2013). *Métodos de segmentación de nubes en imágenes satelitales. Tecnura, 17(36), 96-110.*
- [24] Chumbimuni Contreras, R. C., López, C., & Bosnjak, A. (2008). *Segmentación Utilizando Contornos Activos (Snakes) de Imágenes de Resonancia Magnética del Cerebro. In IV Latin American Congress on Biomedical Engineering 2007, Bioengineering Solutions for Latin America Health: September 24th–28th, 2007 Margarita Island, Venezuela (pp. 313-316). Springer Berlin Heidelberg.*
- [25] Caselles, V., & Frangi, A. (2006). *La segmentación de imágenes. El método de los contornos activos geométricos. Encuentros multidisciplinares, 8(23), 14-26.*
- [26] Casado Coscollá, Á. (2020). *Segmentación de tejidos cerebrales a partir de imágenes de resonancia magnética usando técnicas de Deep Learning.*
- [27] Guzmán-Loreto, G. R., Pérez-Diosdado, A. Z., Durán-Reyes, E., Torres-Calderon, A., González-Guzman, L., Preciado-Trujillo, H., & Guzman-Cabrera, R. (2023).

- 
- Aplicación de la inteligencia artificial en la clasificación automática de imágenes médicas. JÓVENES EN LA CIENCIA, 21, 1-6.*
- [28] S. Timón, M. Rincón, and R. Martínez-Tomás, “Extending XNAT platform with an incremental semantic framework,” *Front Neuroinform*, vol. 11, no. August, pp. 1–13, 2017, doi: 10.3389/fninf.2017.00057.
- [29] “What’s new in 0.6 — MONAI 1.1.0 Documentation.” [https://docs.monai.io/en/stable/whatsnew\\_0\\_6.html](https://docs.monai.io/en/stable/whatsnew_0_6.html) (accessed Mar. 26, 2023).
- [30] “tutorials/auto3dseg at main · Project-MONAI/tutorials · GitHub.” <https://github.com/Project-MONAI/tutorials/tree/main/auto3dseg> (accessed Mar. 26, 2023).
- [31] A. Termine, C. Fabrizio, C. Caltagirone, and L. Petrosini, “A Reproducible Deep-Learning-Based Computer-Aided Diagnosis Tool for Frontotemporal Dementia Using MONAI and Clinica Frameworks,” *Life*, vol. 12, no. 7, 2022, doi: 10.3390/life12070947.
- [32] “Deep Learning for Medical Image Analysis” de Dinggang Shen y otros.
- [33] “Machine Learning and Medical Imaging” de Guorong Wu y otros.
- [34] Gomes, R., Pham, T., He, N., Kamrowski, C., Wildenberg, J. Analysis of Swin-UNet vision transformer for Inferior Vena Cava filter segmentation from CT scans (2023) *Artificial Intelligence in the Life Sciences*, 4, art. no. 100084.
- [35] Zhou, L., Li, Z. Automatic multi-label temporal bone computed tomography segmentation with deep learning (2023) *International Journal of Medical Robotics and Computer Assisted Surgery*, 19 (5), art. no. e2536.
- [36] Nordio, G., Easmin, R., Giacomel, A., Dipasquale, O., Martins, D., Williams, S., Turkheimer, F., Howes, O., Veronese, M., Jauhar, S., Rogdaki, M., McCutcheon, R., Kaar, S., Vano, L., Rutigliano, G., Angelescu, I., Borgan, F., D’Ambrosio, E., Dahoun, T., Kim, E., Kim, S., Bloomfield, M., Egerton, A., Demjaha, A., Bonoldi, I., Nosarti, C., Maccabe, J., McGuire, P., Matthews, J., Talbot, P.S., and the FDOPA PET imaging working group An automatic analysis framework for FDOPA PET neuroimaging (2023) *Journal of Cerebral Blood Flow and Metabolism*, 43 (8), pp. 1285-1300.

- 
- [37] Doran, S.J., D'Arcy, J., Collins, D.J., Andriantsimiavona, R., Orton, M., Koh, D.-M., Leach, M.O. *Informatics in radiology: Development of a research PACS for analysis of functional imaging data in clinical research and clinical trials* (2012) *Radiographics*, 32 (7), pp. 2135-2150. Cited 18 times
- [38] Chu, P.W., Yu, S., Wang, Y., Seibert, J.A., Cervantes, L.F., Kasraie, N., Chu, C.A., Smith-Bindman, R. *Reference phantom selection in pediatric computed tomography using data from a large, multicenter registry* (2022) *Pediatric Radiology*, 52 (3), pp. 445-452. Cited 8 times.
- [39] Gallo Gómez, J. P. (2023). *Desarrollo de un modelo normativo de volumetría cerebral en población SURA.*
- [40] Pelea, L. P. (2019). *Valores atípicos en los datos, ¿ cómo identificarlos y manejarlos?. Revista del Jardín Botánico Nacional*, 40, 99-107
- [41] Salazar, S. C., Lopez, D. L., Escobar, S. R., Uribe, A. D. U., Alvarez, J. I. G., Munera, E. N. M., ... & Perez, C. L. H. *Gestion Eficiente de la Afluencia en el Metro de Medellín: Un Enfoque Predictivo ARIMA.*
- [42] Benjamini, Y., & Braun, H. (2002). *John W. Tukey's contributions to multiple comparisons. Annals of Statistics*, 1576-1594.
- [43] DE CONDICIONES, S. E. O. (2018). *REVISIÓN TÉCNICA Y NOTAS ACLARATORIAS*
- [44] Shrestha, P., Park, Y., Kwon, H., & Kim, C. G. (2017). *Error outlier with weighted Median Absolute Deviation threshold algorithm and FBG sensor based impact localization on composite wing structure. Composite Structures*, 180, 412-419.
- [45] B. Zhan, E. Song, H. Liu, Z. Gong, G. Ma, and C. C. Hung, "CFNet: A medical image segmentation method using the multi-view attention mechanism and adaptive fusion strategy," *Biomed Signal Process Control*, vol. 79, no. P1, p. 104112, 2023, doi: 10.1016/j.bspc.2022.104112.
- [46] J. K. Kuusisto et al., "Revisiting left atrial volumetry by magnetic resonance imaging: the role of atrial shape and 3D angle between left ventricular and left atrial axis," *BMC Med Imaging*, vol. 21, no. 1, pp. 1–12, 2021, doi: 10.1186/s12880-021-00701-5.

- 
- [47] Peláez, I. M. (2016). *Modelos de regresión: lineal simple y regresión logística*. *Revista Seden*, 14, 195-214.
- [48] Soto Garcés, R. E. (2023). *Generalidades de la regresión cuantil y aplicaciones en R*.
- [49] Servy, E., García, M. D. C. E., & Paccapelo, V. (2006). *Regresión no paramétrica: una aplicación*
- [50] Cobo Carrillo, M. (2022). *Procesos gaussianos con aplicaciones en inversiones*.
- [51] Romero Matarin, C. (2023). *Hacia un futuro energético predecible: aplicando Machine Learning para prever los precios de la electricidad*.
- [52] Pérez-Planells, L., Delegido, J., Rivera-Caicedo, J. P., & Verrelst, J. (2015). *Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos*. *Revista de teledetección*, (44), 55-65.
- [53] Chacón Chamorro, M. V. *Estudio de la reducción del sobreajuste en arquitecturas de redes neuronales residuales ResNet en un escenario de clasificación de patrones (Doctoral dissertation, Universidad Nacional de Colombia)*.
- [54] Bostan, P. A., Heuvelink, G. B., & Akyurek, S. Z. (2012). *Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey*. *International Journal of Applied Earth Observation and Geoinformation*, 19, 115-126.
- [55] A. H. Nvidia et al., "UNETR: Transformers for 3D Medical Image Segmentation." [Online]. Available: <https://monai.io/research/unetr>
- [56] Harsh, P., Chakraborty, R., Tripathi, S., & Sharma, K. (2021, June). *Attention U-Net Architecture for Dental Image Segmentation*. In *2021 International Conference on Intelligent Technologies (CONIT)* (pp. 1-5). IEEE
- [57] <https://www.synapse.org/#!/Synapse:syn3193805/wiki/217789>, "Multi-Atlas Labeling Beyond the Cranial Vault - Workshop and Challenge."
- [58] Zheng, H., Chen, Y., Yue, X., Ma, C., Liu, X., Yang, P., & Lu, J. (2020). *Deep pancreas segmentation with uncertain regions of shadowed sets*. *Magnetic Resonance Imaging*, 68, 45-52.

- 
- [59] Elamir, E. (2012, July). *Mean absolute deviation about median as a tool of explanatory data analysis. In Proceedings of the World Congress on Engineering (Vol. 1).*
- [60] Tschuchnig, M. E., & Gadermayr, M. (2022, March). *Anomaly detection in medical imaging-a mini review. In Data Science–Analytics and Applications: Proceedings of the 4th International Data Science Conference–iDSC2021 (pp. 33-38). Wiesbaden: Springer Fachmedien Wiesbaden. monocromáticas. Radiología, 65(5), 402-413.*
- [61] Hartley, M. C., Langan III, E. M., Cull, D. L., Taylor, S. M., Carsten III, C. G., & Blackhurst, D. W. (2009, September). *Evaluación del diámetro de la aorta torácica descendente proximal con la edad: implicaciones para la implantación de una endoprótesis de aorta torácica. In Anales de Cirugía Vasculat (Vol. 23, No. 5, pp. 705-710). Elsevier Doyma*
- [62] García Gómez, A. (2019). *Herramienta automática para la segmentación de la arteria aorta a partir de imágenes de tomografía computarizada.*
- [63] Ocampo, J. M., & Gutiérrez, J. (2005). *Envejecimiento del sistema cardiovascular. Revista colombiana de cardiología, 12(2), 53-6*

