

Construcción de un prototipo para la segmentación de grasa parafaríngea en imágenes de resonancia magnética

Oscar Ivan Duarte Torres

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE INGENIERÍA

MAESTRÍA EN BIOINGENIERÍA, MAESTRÍA EN INGENIERÍA DE SISTEMAS Y
COMPUTACIÓN

BOGOTÁ D.C.

2023

Construcción de un prototipo para la segmentación de grasa parafaríngea en imágenes de resonancia magnética

Autor:

Oscar Ivan Duarte Torres

MEMORIA DEL TRABAJO DE GRADO REALIZADO PARA CUMPLIR UNO DE
LOS REQUISITOS PARA OPTAR A LOS TITULOS DE
MAGISTER EN BIOINGENIERIA
MAGÍSTER EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

Directores

Cesar Julio Bustacara Medina

Liliana Margarita Otero Mendoza

Asesor

Luis Felipe Uriza Carrasco

Comité de Evaluación del Trabajo de Grado

Leonardo Florez Valencia

Patricia Hidalgo Martinez

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE INGENIERÍA

MAESTRÍA EN BIOINGENIERÍA, MAESTRÍA EN INGENIERÍA DE SISTEMAS Y
COMPUTACIÓN

BOGOTÁ D.C.

ENERO 12 DE 2023

PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN BIOINGENIERÍA, MAESTRÍA EN INGENIERÍA DE SISTEMAS Y
COMPUTACIÓN

Rector Magnífico

Jorge Humberto Peláez, S.J.

Decano Facultad de Ingeniería

Ingeniero Lope Hugo Barrero Solano

Director Maestría en Bioingeniería

Ingeniero Daniel Ricardo Suarez Venegas

Director Maestría en Ingeniería de Sistemas y Computación

Ingeniera Angela Carrillo Ramos

Director Departamento de Ingeniería de Sistemas

Ingeniero Cesar Julio Bustacara Medina

Artículo 23 de la Resolución No. 1 de junio de 1946

“La Universidad no se hace responsable de los conceptos emitidos por sus alumnos en sus proyectos de grado. Sólo velará porque no se publique nada contrario al dogma y la moral católica y porque no contengan ataques o polémicas puramente personales. Antes bien, que se vean en ellos el anhelo de buscar la verdad y la Justicia”

AGRADECIMIENTOS

Agradezco a mi familia por todo el apoyo brindado para el cumplimiento de mis metas y por el impulso para alcanzar mis objetivos. Le agradezco a mis directores de trabajo de grado el profesor Cesar Bustacara y la profesora Liliana Otero por la dedicación, paciencia, acompañamiento y consejos de los cuales aprendí y fueron muy importantes para la finalización del proyecto. Al doctor Luis Felipe Uriza por la asesoría y el aporte de las imágenes de resonancia magnética, también a la doctora María Daniela Alvarado y al doctor Daniel Sebastián Chaves por el trabajo en la segmentación manual de las imágenes de resonancia magnética. Agradecerle a cada uno de mis docentes ya que de sus conocimientos aprendí lo necesario para llegar a esta etapa, y para finalizar quiero dar un agradecimiento especial a la Pontificia Universidad Javeriana por ser mi casa de estudios para esta etapa académica y brindarme las herramientas para alcanzar mis objetivos.

Contenido	
1	Introducción 11
2	Objetivos 13
2.1	Objetivo general 13
2.2	Objetivos específicos 13
3	Metodología 14
3.1	Análisis de la investigación 14
3.2	Diseño 15
3.3	Evaluación 15
4	Antecedentes 16
4.1	Segmentación en imágenes de resonancia magnética 16
4.2	Segmentación en imágenes de tomografía computarizada 17
4.3	Estudios relacionados 18
4.4	Métrica de calidad de la segmentación 19
4.5	Red neuronal convolucional 19
4.6	Red neuronal U-net 20
4.7	Red U-net 3D 21
4.8	Funciones de activación 21
4.8.1	Función sigmoide 22
4.8.2	Función lineal unitaria rectificadora (ReLU) 22
4.9	Análisis de investigaciones relacionadas 23
5	Diseño de la solución 26
5.1	Análisis de la solución 26
5.1.1	Requerimientos del sistema 26
5.1.2	Caracterización del proceso de segmentación 28
5.1.3	Población y Muestra 29
5.2	Diseño del prototipo 29
5.2.1	Arquitectura del prototipo 30
6	Implementación 35
6.1.1	Caracterización de las imágenes 35
6.1.2	Selección de las imágenes 35
6.1.3	Segmentación manual de imágenes 36
6.1.4	Preprocesamiento de imágenes 36

6.1.5	Prototipo	38
7	Pruebas y resultados	43
7.1	Entrenamiento y Segmentación	43
8	Discusión	46
9	Conclusiones y trabajos futuros	48
10	Referencias	49
11	Anexos	54
11.1	Anexo 1. Aval de comité de ética	54
11.2	Anexo 2. Documento de Especificación de Requerimientos de Software	54
11.3	Anexo 3. Documento de Arquitectura de Software	54
11.4	Anexo 4. Repositorio del prototipo	54

Índice de Figuras

Figura 1 Modelo general de la metodología.....	14
Figura 2 Arquitectura U-net	20
Figura 3 Arquitectura red U-net 3D	21
Figura 4 Ubicación grasa parafaríngea.....	25
Figura 5 Diagrama del Sistema de Segmentación de Grasa Parafaríngea.....	26
Figura 6 Proceso de segmentación de grasa parafaríngea	28
Figura 7 <i>Casos de uso del sistema</i>	30
Figura 8 Patrón de arquitectura en capas.....	31
Figura 9 Vista lógica del prototipo.....	32
Figura 10 Vista de implementación del prototipo	33
Figura 11 Arquitectura U-net 3D implementada.....	34
Figura 12 Transformaciones preprocesamiento	37
Figura 13 Gráfica de dispersión de intensidades máximas	38
Figura 14 Vista inicial del prototipo.....	39
Figura 15 Vistas intermedias segmentación	39
Figura 16 Vista final segmentación.....	39
Figura 17 Confirmación de inicio de entrenamiento.....	40
Figura 18 Vista de selección de carpetas entrenamiento.....	40
Figura 19 Vista final entrenamiento	40
Figura 20 Perdida de DICE en el entrenamiento y validación con ReLu	43
Figura 21 Perdida de DICE en el entrenamiento y validación con Sigmoide.....	44
Figura 22 Volumen segmentación de cada imagen.....	46

Índice de Tablas

Tabla 1 Comparación de estudios de segmentación de grasa del cuerpo y estructuras de la vía aérea superior.....	23
Tabla 2 Requerimientos funcionales	27
Tabla 3 Características seleccionadas para el modelo de calidad.....	27
Tabla 4 Requerimientos no funcionales	28
Tabla 5 Escenarios de atributos de calidad.....	31
Tabla 6 Caracterización de imágenes consolidada	35
Tabla 7 Comparación de métricas para la validación, entrenamiento y prueba	44
Tabla 8 Comparación volúmenes calculados	44
Tabla 9 Imágenes segmentadas y volumen	45
Tabla 10 Comparación resultados estudios	47
Tabla 11 Comparación de características de los estudios.....	47

Resumen

El propósito del estudio realizado y presentado en este documento, es la construcción de un prototipo para la detección y segmentación del tejido graso en la región parafaríngea y cuyo impacto podrá verse en su potencial uso para encontrar relaciones entre la presencia de grasa parafaríngea y la existencia y severidad de la AOS. El prototipo hace uso de una red convolucional U-net 3D para detectar y segmentar la grasa en imágenes de resonancia magnética en tres dimensiones. El estudio se realizó siguiendo las fases análisis, diseño y evaluación, integrando los ciclos de rigor, relevancia y diseño, pertenecientes a la ciencia del diseño. En la fase de análisis se determinaron las diferentes técnicas que se pueden utilizar para detectar y segmentar grasa en la región parafaríngea. En la fase de diseño se realizó la selección de las imágenes a usar y se desarrolló el prototipo mediante una metodología de construcción de software basada en el modelo en espiral. En la fase de evaluación se realizó la segmentación de la grasa parafaríngea usando el prototipo implementado. Para esta fase la métrica de exactitud que se utilizó es el coeficiente Dice, que está basada en relaciones de voxeles, el cuál fue de 0.8086 para la red U-net 3D con función de activación ReLU y de 0.6367 con la misma red, pero usando la función de activación sigmoide.

1 Introducción

Las vías aéreas superiores comprenden una parte del sistema respiratorio que va desde los labios y las fosas nasales hasta la tráquea. En el contexto de los estudios llevados a cabo mediante el uso de imágenes médicas en las vías aéreas, se encuentran estudios relacionados con trastornos del sueño como la apnea obstructiva del sueño. La apnea obstructiva del sueño (AOS), es un trastorno respiratorio del sueño, en el cual ocurre estrechamiento parcial o completo, o colapso, de las vías respiratorias faríngeas durante el sueño, que produce una interrupción repetida del flujo de aire, desaturación de oxígeno e interrupción del sueño (Carter & Watenpugh, 2008).

La apnea obstructiva del sueño es un problema importante de salud pública (Motamedi, McClary, & Amedee, 2009) porque es un factor de riesgo para enfermedades cardiovasculares como la hipertensión arterial, falla cardíaca, arritmia y enfermedad coronaria (Kato, Adachi, Koshino, & Somers, 2009). La prevalencia de AOS es alta y se puede incrementar debido a la epidemia actual de obesidad. La AOS tiene efectos económicos, aumenta los accidentes de tránsito, y disminuye la calidad del trabajo (Morsy et al., 2019).

Los estudios mediante métodos no invasivos, como las imágenes médicas se encaminan a realizar la detección de obstrucciones y caracterización de las vías aéreas superiores con métodos automáticos o semiautomáticos de segmentación de imágenes, para apoyar los procesos de diagnóstico y/o tratamiento de la Apnea Obstructiva del Sueño (Ivanovska, Dober, Laqua, Hegenscheid, & Völzke, 2013; Javed, Kim, Khoo, Ward, & Nayak, 2016; MuhammadLaiq Ur Rahman et al., 2015; Shahid et al., 2017). Estos estudios se realizan usando imágenes de Tomografías Computarizadas, Resonancia Magnética, ultrasonido y recientemente con Tomografías Computacionales de haz cónico o CBCT (Cone Beam Computed Tomography) por su sigla en inglés.

Hay varios estudios que sugieren la existencia de una relación entre la AOS y los tejidos grasos en las vías aéreas superiores, en los cuales se investiga si la acumulación de grasa en los tejidos que rodean la vía aérea superior retropalatina contribuyen a la aparición de la AOS. Estos estudios fueron realizados con datos obtenidos a partir de medidas antropométricas (Carter & Watenpugh, 2008; Ruiz et al., 2016), y de imágenes de resonancia magnética (Dhaliwal et al., 2018; Shahid et al., 2017).

A partir de la relación aparente mencionada anteriormente, ha surgido el interés en el estudio de la acumulación de grasa en la lengua y en la región retropalatina de pacientes con AOS, encontrando que la grasa depositada en los tejidos que rodean la vía aérea superior contribuyen a la colapsabilidad de la vía aérea en personas con y sin AOS (Kim et al., 2014; Li et al., 2012; Shahid et al., 2017; Wang et al., 2020).

Los estudios de las vías aéreas superiores a nivel general presentan usos variados, orientados principalmente en encontrar relaciones entre la anatomía de las diferentes estructuras en la región como la lengua, la faringe, el volumen de los tejidos grasos, entre otros, y la existencia de la AOS. Por esta razón surge el interés en investigar si es posible establecer la existencia de una relación entre la acumulación de grasa en la región parafaríngea y en la lengua con la severidad de la AOS en pacientes adultos.

La búsqueda de la respuesta a la pregunta anterior se encuentra bajo investigación por parte de un equipo interdisciplinario de odontólogos, radiólogos e ingenieros de la Pontificia Universidad Javeriana y el Hospital Universitario San Ignacio. En la búsqueda de aportar a la investigación, en el presente estudio se planteó la siguiente pregunta: ¿Cómo detectar y segmentar de forma automática el tejido graso en la región parafaríngea a partir de imágenes de Resonancia Magnética? Con los resultados de la segmentación automática, se construyó un prototipo funcional que puede ayudar en investigaciones futuras a identificar la relación entre la acumulación de grasa en la región parafaríngea con las obstrucciones de la vía aérea en individuos con AOS.

2 Objetivos

En esta sección se dan a conocer los objetivos que permitirán responder la pregunta: ¿Cómo detectar y segmentar de forma automática el tejido graso en la región parafaríngea a partir de imágenes de Resonancia Magnética?

2.1 Objetivo general

Desarrollar un prototipo funcional que permita la detección y segmentación del tejido graso en la región parafaríngea en imágenes de resonancia magnética

2.2 Objetivos específicos

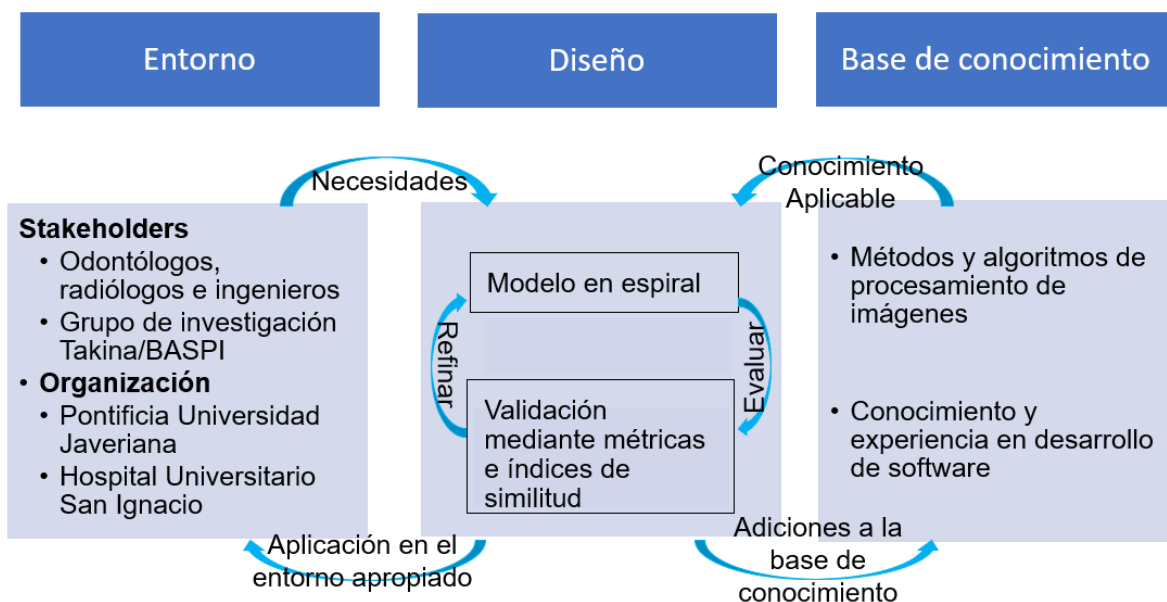
1. Identificar las diferentes técnicas de detección y segmentación de grasa en la vía aérea superior, con el fin de seleccionar la más adecuada para su implementación.
2. Diseñar e implementar un pipeline para la detección y la segmentación de la grasa parafaríngea en imágenes de resonancia magnética
3. Validar el prototipo usando métricas de calidad basadas en relaciones de voxeles con respecto a la segmentación manual.

3 Metodología

Para el desarrollo y consecución de los objetivos propuestos en este Trabajo de Grado, se utilizó una metodología basada en la ciencia del diseño. Esta metodología se basa en tres ciclos principales: la identificación del entorno, el diseño/investigación y la obtención de la base de conocimiento. Los ciclos principales se pueden dividir en fases según lo planteado por Ken Peffers en (Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, & Samir Chatterjee, 2007)

En la siguiente figura se muestra el modelo general de la ciencia del diseño adaptado de (Hevner et al., 2004) para la presente propuesta:

Figura 1 Modelo general de la metodología



Las fases establecidas en la metodología son:

1. Análisis de la investigación
2. Diseño del artefacto
3. Evaluación del artefacto

3.1 Análisis de la investigación

La fase de análisis de la investigación se realizó a partir de investigaciones bibliográficas y de casos de estudio para conocer las técnicas y algoritmos de detección y segmentación de grasa en la vía aérea superior y en otras regiones del cuerpo como el abdomen, corazón y piernas. Este proceso se realizó mediante una búsqueda rigurosa en diferentes bases de datos.

Para el desarrollo de esta fase se realizaron las siguientes actividades:

- A. Identificar y seleccionar las técnicas que permitan realizar la segmentación del tejido graso en la región parafaríngea.
- B. Generar el estado del arte

3.2 Diseño

La fase de diseño se realiza con base en una metodología de construcción de software, la cual utilizará un modelo en espiral definido por Barry Boehm en (Boehm, 1986). Del modelo genérico se omiten las actividades de mantenimiento, dado que se construirá un prototipo funcional del sistema, también se unifican las fases de definición de objetivos y análisis de riesgos.

Para el desarrollo de esta fase se realizaron las siguientes etapas y sus correspondientes actividades:

- A. Etapa de definición de objetivos, alternativas, restricciones y riesgos
 - a. Obtener los requerimientos del sistema de detección y segmentación de grasa
 - b. Realizar el análisis de los requerimientos.
 - c. Caracterizar el proceso de detección y segmentación.
- B. Etapa de desarrollo y pruebas
 - a. Diseñar la arquitectura del prototipo de detección y segmentación.
 - b. Codificar el prototipo.
 - c. Documentar el prototipo
 - d. Realizar pruebas funcionales del prototipo.

3.3 Evaluación

En esta fase se evaluaron las imágenes segmentadas resultantes del método desarrollado comparándolas con las segmentaciones realizadas de forma manual. Para lo cual se realizaron las siguientes actividades, definidas a partir de otros estudios revisados en la literatura:

- A. Comparar el método propuesto con la segmentación manual mediante el coeficiente DICE
- B. Generar análisis de resultados.
- C. Generar memoria del trabajo de grado

4 Antecedentes

En esta sección se hace una revisión de literatura dando a conocer estudios y definiciones relacionadas con el desarrollo del proyecto, finalizando con un análisis del estado del arte en el cual se enumeran las técnicas utilizadas, algunas características y los resultados obtenidos en cada uno de los estudios, para verificar principalmente cuales son las técnicas más usadas y con mejores resultados.

El proceso de dividir una imagen en regiones significativas se denomina segmentación de la imagen. En el caso de las vías aéreas la segmentación es el proceso de extracción de la estructura de interés mediante la identificación de aquellos voxels de la imagen que pertenecen a los espacios de las vías respiratorias en función de su intensidad, o de los patrones de intensidad en una región de interés (Kabaliuk et al., 2017). Las fuentes de imágenes médicas que se tendrán en cuenta en la presente revisión son la resonancia magnética o MR y la tomografía computarizada o CT. En la presente revisión se tuvieron en cuenta estudios de los últimos 10 años y realizados sobre la vía aérea superior.

4.1 Segmentación en imágenes de resonancia magnética

Los datos de la MR ofrecen el mejor contraste de tejido/aire, pero los artefactos de movimiento pueden estar presentes debido al largo tiempo de adquisición de la imagen. En las vías respiratorias superiores, esto afecta particularmente a la lengua y al paladar blando (Kabaliuk et al., 2017). A pesar de la afectación mencionada anteriormente, diferentes autores han realizado estudios utilizando este tipo de imágenes en estudios relacionados con la AOS. Tatyana Ivanovska es una de las investigadoras que más estudios ha realizado junto con otros autores en cuanto a segmentación de imágenes de resonancia magnética. Estos estudios han sido realizados principalmente en la zona de la faringe, y han sido enfocados al desarrollo de métodos semiautomáticos (Ivanovska, Dober, Laqua, Hegenscheid, & Völzke, 2013) y automáticos (MuhammadLaiq Ur Rahman et al., 2015; Shahid et al., 2017) de segmentación con análisis de imágenes basado en objetos, los cuales fueron comparados con métodos como conjunto de nivel, crecimiento de regiones, multiotsu, watershed, y c-medias, siendo los métodos desarrollados por los autores mejores en cuanto al coeficiente de similitud Dice.

Otro estudio que utilizó un método de segmentación de umbral, se enfocó en probar que el exceso de grasa se deposita en la lengua de pacientes obesos con apnea de sueño, encontrando una posible relación entre la obesidad y la AOS (Kim et al., 2014). Cabe resaltar que las investigaciones realizadas con segmentación de imágenes de MR han estado enfocadas en obtener datos acerca de la grasa en la región faríngea y en la lengua, buscando la relación entre la AOS y un mayor volumen de grasa en las regiones mencionadas.

Ivanovska et al. (2021) proponen la utilización de redes convolucionales con arquitectura u-net 2D y 3D para realizar la segmentación automática de grasa parafaríngea en imágenes de MR de cabeza, columna. Este estudio, continua con la investigación que busca determinar la influencia de la grasa parafaríngea en el desarrollo del síndrome de apnea obstructiva del sueño.

Bommineni et al. (2022) proponen la utilización de redes neuronales convolucionales, con arquitectura u-net modificada, para la segmentación automática y precisa de estructuras y tejidos de las vías aéreas superiores considerados como factores de riesgo para la aparición de la AOS. Los resultados obtenidos por los investigadores son prometedores y dejan ver el potencial del método propuesto en estudios clínicos a gran escala para reemplazar la segmentación manual.

4.2 Segmentación en imágenes de tomografía computarizada

La CT convencional es ampliamente usada como una fuente de modelos anatómicos en 3D debido a su buena resolución espacial y de contraste y, en comparación con la MR, tiene artefactos de movimiento limitados debido a la corta duración del escaneo, pero tiene una pobre diferenciación de tejidos blandos. La CBCT es ampliamente usada para varias aplicaciones en odontología, y ha crecido en popularidad debido a su bajo costo y baja dosis de radiación comparada con la CT convencional (Kabaliuk et al., 2017).

Las imágenes de tomografía computacional de haz cónico han sido usadas en diferentes estudios de las vías aéreas superiores, en los cuales su metodología varía pasando por el uso de programas para segmentar hasta el desarrollo de nuevos métodos de segmentación. En la investigación realizada por El & Palomo. (2010), los autores realizan una comparación de la segmentación generada por tres programas Dolphin3D, InVivoDental, OrthoSegment, pero sin identificar las técnicas de segmentación usadas y obteniendo una baja precisión en las segmentaciones.

Bui, Ong, & Foong. (2015) proponen un método que utiliza la segmentación de conjunto de nivel para delinear la cavidad nasal y los senos paranasales, en conjunto con la implementación un contorno activo localizado anisotrópico. El método automático propuesto fue comparado con la segmentación manual, y otros métodos de segmentación como el contorno activo local o global basado en regiones, obteniendo una alta precisión e indicando que puede ser mejorado.

Neelapu et al.. (2017) proponen un método de inicialización de contorno automático para las vías aéreas faríngea y sinonasal, el cual es probado en conjunto con cinco métodos de segmentación con el objetivo de evaluar su robustez. Los métodos de segmentación usados fueron el conjunto de nivel de Chan-Vese, el conjunto de nivel localizado de Chan-Vese, el

conjunto de nivel de distancia Bhattacharya, growcut y el método de campo disperso, siendo este último el recomendado por los autores.

En el estudio realizado por Kabaliuk et al. (2017), realizan una comparación objetiva de métodos de segmentación en dos subregiones de la vía aérea superior, la cavidad nasal y la región faríngea. Los métodos usados fueron el umbral global, la segmentación multi paso de conjunto de nivel, evolución de serpiente en donde se usaron métodos de inicialización por agrupamiento, inicialización por clasificación e inicialización por atracción de borde. El método de evolución de serpiente con inicialización por clasificación fue el que mostró los mejores resultados en la cavidad nasal, mientras que en la región faríngea fue relativamente insensible al método de segmentación usado.

En la revisión sistemática realizada por Alsufyani, Flores-Mir, & Major (2012) se encuentran más usos de la segmentación de imágenes de CBCT pero en varios casos no se identifican los métodos utilizados o las investigaciones se encuentran enfocadas en los estudios clínicos, razón por la cual no mencionan los algoritmos usados.

4.3 Estudios relacionados

En esta sección se presentan otros estudios relacionados con la segmentación de grasa en diferentes regiones del cuerpo sobre imágenes de CT y MR principalmente. Los métodos de segmentación utilizados en estos estudios varían desde la segmentación con crecimiento de regiones y c medias difusas hasta redes neuronales convolucionales en regiones como el abdomen, las piernas y el corazón. En (Maddalo et al., 2017) utilizaron un software llamado Fatcalc para la segmentación no supervisada de grasa abdominal en MR y realizaron un análisis de las diferencias volumétricas a nivel de cortes y del volumen completo. En (Zlokolica et al., 2017) utilizaron c medias difusas para la segmentación de grasa en el epicardio superando en índice DICE a métodos como snake, crecimiento de regiones y umbralización local adaptativa morfológica (LAMT), siendo este último un método desarrollado por los mismos autores.

En diferentes estudios como los realizados por (Amer, Nassar, Bendahan, Greenspan, & Ben-Eliezer, 2019), (Masoudi et al., 2020), (Dabiri et al., 2020), (He et al., 2021) y (MacLean et al., 2021) utilizaron redes neuronales convoluciones basadas en la arquitectura U-net para la segmentación de grasa en las piernas, abdomen y corazón en CT y MR, obteniendo un rendimiento superior o cercano a 0.9 en cuanto al índice DICE y Jaccard.

Otros métodos utilizados para segmentación de grasa son el crecimiento de regiones en (Aggarwal & Panwar Jogi, 2020), conectividad difusa en (Irmakci et al., 2019) y extracción de características con filtro de Gabor en (Kazemi et al., 2019), en los cuales los autores obtuvieron un coeficiente DICE superior al 0.8.

4.4 Métrica de calidad de la segmentación

El coeficiente de similitud Dice (DICE), también llamado el índice de superposición, es la métrica más usada para la validación de segmentación de volúmenes médicos (Taha & Hanbury, 2015). El coeficiente DICE mide la superposición espacial entre dos segmentaciones, definido como $DSC(A,B) = 2(A \cap B)/(A+B)$ donde \cap es la intersección y A y B son regiones (Zou et al., 2004). La fórmula para calcular el índice DICE se define así:

$$DICE = \frac{2 * (A \cap B)}{A + B} = \frac{2TP}{2TP + FP + FN}$$

Siendo TP los verdaderos positivos, FP los falsos positivos y FN los falsos negativos. A partir de la fórmula anterior se obtiene un resultado entre 0 y 1, siendo 0 una superposición nula y 1 una total superposición entre las dos segmentaciones.

4.5 Red neuronal convolucional

Las redes neuronales convolucionales o CNN (Convolutional Neural Network) por su sigla en inglés, “son un tipo de redes neuronales especializadas para el procesamiento de datos que tienen una topología en forma de cuadrícula o malla, como las series de tiempo en intervalos regulares en una dimensión o las imágenes que pueden ser vistas como mallas de píxeles en dos dimensiones” (Goodfellow, Bengio, & Courville, 2016) y los volúmenes que pueden ser vistos como mallas en tres dimensiones de voxels. Las redes convolucionales son redes neuronales que usan una operación llamada convolución en al menos una de sus capas.

$$s(t) = (x * w)(t)$$

En la operación de convolución el primer argumento es conocido como entrada (x) y es usualmente una matriz multidimensional de datos y el segundo argumento es conocido como kernel (w) y es usualmente una matriz multidimensional de parámetros, la salida es conocida normalmente como mapa de características. Las capas de las redes convolucionales están típicamente compuestas por tres etapas: la etapa de convolución donde se realizan convoluciones en paralelo para producir un conjunto de activaciones lineales, la segunda etapa o etapa de detección donde cada activación lineal se pasa por una función de activación no lineal como la función de activación lineal rectificadora o ReLU, y en la tercera etapa se usa una función de agrupación o pooling para modificar la salida de la capa (Goodfellow et al., 2016).

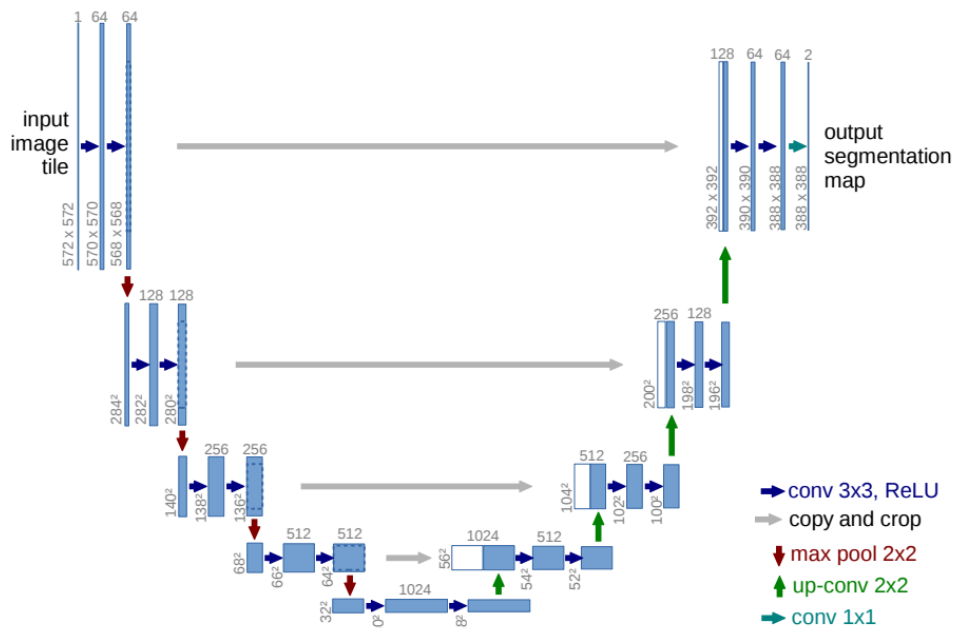
Un ejemplo de agrupación es la operación max pooling donde a partir de un vecindario rectangular se toman los valores máximos como salida.

4.6 Red neuronal U-net

Es una red neuronal convolucional que consta de un camino de contracción en el lado izquierdo y uno de expansión en el lado derecho como se puede observar en la

Figura 2. Esta red neuronal no requiere de una gran conjunto de datos para ser entrenada ya que según lo planteado por (Ronneberger, Fischer, & Brox, 2015) se pueden emplear técnicas de aumento de datos aplicando deformaciones en el conjunto de imágenes existente.

Figura 2 Arquitectura U-net



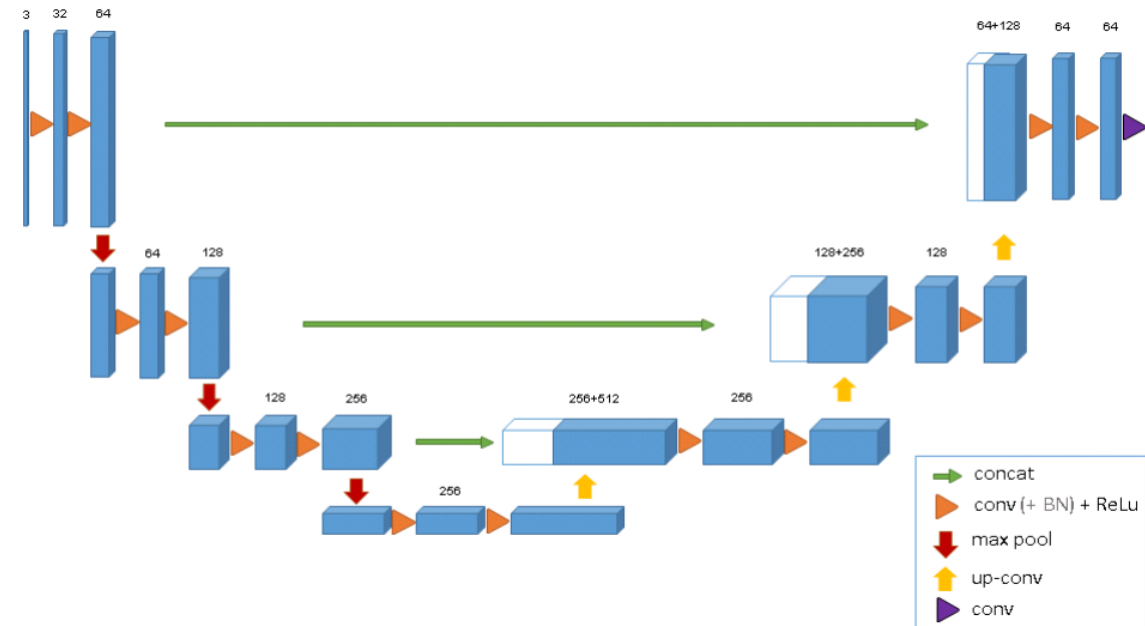
Nota. Adaptado de U-Net: Convolutional Networks for Biomedical Image Segmentation, por Ronneberger et al., 2015

Esta red neuronal en su camino de contracción consiste en la repetida aplicación de convoluciones 3x3, seguida de Unidades de Rectificación Lineal (ReLU) y max-pooling 2x2 con paso 2 para downsampling que duplica el número de canales de salida del bloque. En el camino de expansión cada bloque consiste en un sobre muestreo o upsampling seguida por una convolución ascendente de 2x2 que reduce a la mitad el número de canales, una concatenación con el mapa de características correspondiente en el camino de contracción y dos convoluciones 3x3 cada una seguida por una ReLU. La capa final tiene una convolución 1x1 que mapea cada uno de los 64 vectores de características al número deseado de clases.

4.7 Red U-net 3D

La red U-net 3D se basa en la arquitectura U-net estándar descrita anteriormente pero implementada para 3 dimensiones y con algunos cambios en los bloques de contracción y análisis, también llamados caminos de análisis y síntesis por (Çiçek, Abdulkadir, Lienkamp, Brox, & Ronneberger, 2016).

Figura 3 Arquitectura red U-net 3D



Nota. Adaptado de 3D U-net: Learning dense volumetric segmentation from sparse annotation, por Çiçek et al., 2016

En la Figura 3 se puede observar la arquitectura de la red U-net 3D implementada por los autores. En esta red existen concatenaciones entre capas de la misma resolución que permiten mantener los detalles de alta resolución. Otro cambio en esta arquitectura con respecto a la estándar es la adición de la normalización en lote (BN o Batch Normalization) antes de cada ReLU, lo que permite la que permite obtener mejores tiempos tanto en entrenamiento como en validación.

4.8 Funciones de activación

Es una función que define la salida de un nodo o neurona dada una entrada o conjunto de entradas. Esta función es usada en cada neurona de la red para ayudarla a aprender los patrones complejos de los datos, permitiendo que esta pueda hacer predicciones, además introduce no linealidad a la red neuronal. Las funciones de activación comparten una serie de propiedades como son la no linealidad, son diferenciables y continuas (Gustineli, 2022).

Las funciones de activación presentan una serie de desafíos como lo son el desvanecimiento del gradiente y el problema de las neuronas muertas. El desvanecimiento del gradiente ocurre cuando los valores del gradiente durante el entrenamiento se encuentran cercanos a cero a medida que la retropropagación va más profunda en la red haciendo que los pesos se saturan y no se actualicen apropiadamente lo que hace que el decrecimiento de la pérdida se detenga, evitando así que se realice el entrenamiento de una forma apropiada (Gustineli, 2022).

El problema de las neuronas muertas ocurre cuando los valores de la salida de la función de activación son cercanos a cero, evitando así que la neurona se active y que pueda contribuir a la salida de la red. Lo anterior hace que la suma ponderada de una parte de la red sea cero y que una parte de la entrada sea cero, afectando el rendimiento de la red (Gustineli, 2022).

4.8.1 Función sigmoide

La función sigmoide, también llamada función sigmoide logística, es una de las funciones más conocidas y usadas para redes neuronales feedforward. La función sigmoide es definida como:

$$f(x) = \frac{1}{1 + e^{-x}}$$

La función sigmoide se encuentra acotada entre 0 y 1, lo cual hace que para grandes cambios en las entradas ocurran pequeños cambios en las salidas, produciendo así un cambio muy pequeño en el gradiente lo cual hace que la función sea propensa a sufrir del problema de desvanecimiento de gradiente (Gustineli, 2022).

4.8.2 Función lineal unitaria rectificadora (ReLU)

La función ReLU es una de las funciones de activación más usadas en la actualidad, fue introducida a inicios de los 2000 y reintroducida de forma exitosa en el 2010 (Venkatesan & Li, 2017). La función es definida de la siguiente forma:

$$f(x) = \max(0, x)$$

Esta función es continua, no se encuentra acotada y no es exponencial como la función sigmoide, lo que hace que tenga un bajo costo computacional. La función ReLU tiene mejor rendimiento y generalización comparada con la función sigmoide. Debido a que los valores negativos que entran a la función salen como cero, esta función sufre del problema de las neuronas muertas, ya que las neuronas con pesos negativos no aportan a la salida final (Gustineli, 2022).

4.9 Análisis de investigaciones relacionadas

Los métodos de segmentación automática requieren de un punto de referencia para la evaluación de los resultados, y este punto de referencia es la segmentación manual realizada por expertos. La segmentación manual es un proceso que requiere de mucho tiempo, razón por la cual existe una demanda de métodos automáticos o semiautomáticos con poca participación del usuario (Alsufyani et al., 2012). Los estudios de segmentación de imágenes de MR de la vía aérea superior se encuentran enfocados en la región faríngea y se encuentran muy relacionados con la AOS, mostrando resultados prometedores pero que requieren de una mayor cantidad de datos y estudios para ser validados.

En cuanto a la segmentación de imágenes de tomografía de haz cónico, los estudios son más variados y están tendiendo a la completa automatización de los algoritmos semiautomáticos, pero dependiendo todavía de la participación del usuario. Los estudios están sujetos a la calidad de las imágenes médicas según Kabaliuk et al. (2017), razón por la cual se puede decir que aún no existe una solución global a la segmentación de imágenes médicas. Los algoritmos actuales también requieren ser probados con una mayor cantidad de datos para que los estudios sean estadísticamente significativos (Kabaliuk et al., 2017) y poder probar su fiabilidad en casos clínicos reales (Neelapu et al., 2017).

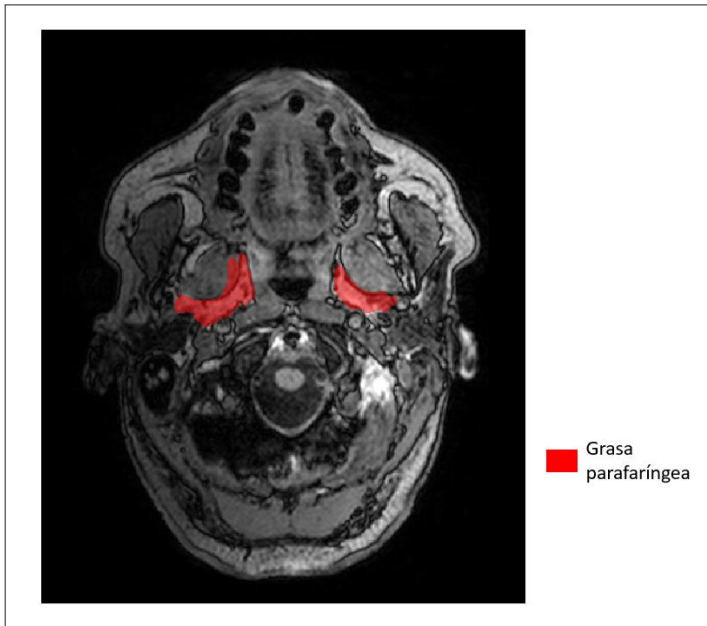
A continuación, se presenta un análisis de los diferentes métodos utilizados principalmente para segmentar tejido graso en varias regiones del cuerpo humano como la vía aérea superior, el abdomen, las piernas y el corazón.

Tabla 1 Comparación de estudios de segmentación de grasa del cuerpo y estructuras de la vía aérea superior

Año	Técnica	Tipo de imagen	Región del estudio	Resultados (DICE)	Referencia
2013	Multiotsu - Análisis de componentes conectados 2D y 3D	MRI	Vía aérea superior	0.8577	(Ivanovska, Dober, Laqua, Hegenscheid, & Völzke, 2013)
2015	Análisis basado en objetos 3D	MRI	Vía aérea superior	0.916	(MuhammadLaiq Ur Rahman et al., 2015)
2016	Crecimiento de regiones 3D	MRI	Vía aérea superior	0.84-0.94	(Javed, Kim, Khoo, Ward, & Nayak, 2016)
2017	Umbralización global	CBCT	Vía aérea superior	0.85	(Kabaliuk et al., 2017)
2017	Conjunto de nivel multipaso	CBCT	Vía aérea superior	0.89	(Kabaliuk et al., 2017)
2017	Umbralización	CBCT	Vía aérea superior	0.9	(Kabaliuk et al., 2017)
2017	Clustering	CBCT	Vía aérea superior	0.96	(Kabaliuk et al., 2017)
2017	Clasificación	CBCT	Vía aérea superior	0.92	(Kabaliuk et al., 2017)
2017	C Medias Difusas	CT	Corazón	0.54-0.82	(Zlokolica et al., 2017)
2017	LAMT	CT	Corazón	0.07-05	(Zlokolica et al., 2017)
2017	Crecimiento de regiones	CT	Corazón	0.15-0.39	(Zlokolica et al., 2017)
2017	Snake	CT	Corazón	0.15-0.38	(Zlokolica et al., 2017)

2017	Análisis de componentes automático	MRI	Vía aérea superior	0.779	(Shahid et al., 2017)
2017	Crecimiento de regiones	MRI	Vía aérea superior	0.358	(Shahid et al., 2017)
2017	C Medias Difusas	MRI	Vía aérea superior	0.434	(Shahid et al., 2017)
2017	Multitotsu	MRI	Vía aérea superior	0.398	(Shahid et al., 2017)
2017	Watersheds	MRI	Vía aérea superior	0.286	(Shahid et al., 2017)
2017	Conjunto de nivel	MRI	Vía aérea superior	0.406	(Shahid et al., 2017)
2019	Extracción de características + filtro de Gabor	CT	Corazón	0.98	(Kazemi et al., 2019)
2019	Conectividad difusa	MRI	Piernas	0.806-0.878	(Irmakci et al., 2019)
2019	Deep convolutional autoencoder	MRI	Piernas	0.911-0.93	(Amer, Nassar, Bendahan, Greenspan, & Ben-Eliezer, 2019)
2020	Crecimiento de regiones modificado	MRI-PDF	Abdomen	0.8747-0.9094	(Aggarwal & Panwar Jogi, 2020)
2020	U-net+FCN based	CT	Abdomen	0.7521-0.9894	(Dabiri et al., 2020)
2020	U-net	MRI	Abdomen	0.9746	(Masoudi et al., 2020)
2021	U-net	CT	Abdomen	0.98	(MacLean et al., 2021)
2021	DeepLab+MobileNetV2	CT	Abdomen	0.975	(MacLean et al., 2021)
2021	Deeplab+Xception	CT	Abdomen	0.979	(MacLean et al., 2021)
2021	U-net+Deep Supervision	CCTA	Corazón	0.97	(He et al., 2021)
2021	3D U-net	CCTA	Corazón	0.95	(He et al., 2021)
2021	3D U-net, 2D U-net	MRI	Vía aérea superior	0.77	(Ivanovska et al., 2021)
2022	U-net + Resnet+ Modified Inception-Resnet-A	MRI	Vía aérea superior	0.58	(Bommineni et al., 2022)

En la Tabla 1 se observa la comparación de diferentes métodos utilizados para segmentar grasa en diferentes regiones del cuerpo como la vía aérea superior, el corazón, las piernas, el abdomen y en la segmentación de estructuras de la vía aérea superior. La comparación se realizó teniendo en cuenta el coeficiente DICE obtenido al usar cada uno de los métodos y el tipo de imagen utilizada en el estudio, teniendo como referencia el año y la cita bibliográfica de cada estudio.

Figura 4 *Ubicación grasa parafaríngea*

En los últimos años a partir del año 2019 se observa una tendencia en el uso de métodos basados en redes neuronales convolucionales en su mayoría con buenos resultados basados en el coeficiente DICE. Estas redes neuronales convolucionales presentan un beneficio mayor al ser métodos automáticos de segmentación y que pueden aportar en la obtención de mejores resultados en estudios con cantidades mayores de datos en comparación con los métodos semiautomáticos. Las redes neuronales convolucionales han sido utilizadas en su mayoría para la segmentación de grasa en el abdomen en imágenes de tomografía computarizada, razón por la cual se detecta una oportunidad para que sean utilizadas en estudios de segmentación de grasa parafaríngea. En la **¡Error! No se encuentra el origen de la referencia.** se puede observar la vista axial de un corte donde se puede visualizar la segmentación de la grasa parafaríngea.

5 Diseño de la solución

En esta sección se presenta el análisis de la investigación y la solución, en el cual se describe el análisis de los requerimientos, también se presenta el diseño de la arquitectura del sistema y de la red neuronal convolucional utilizada.

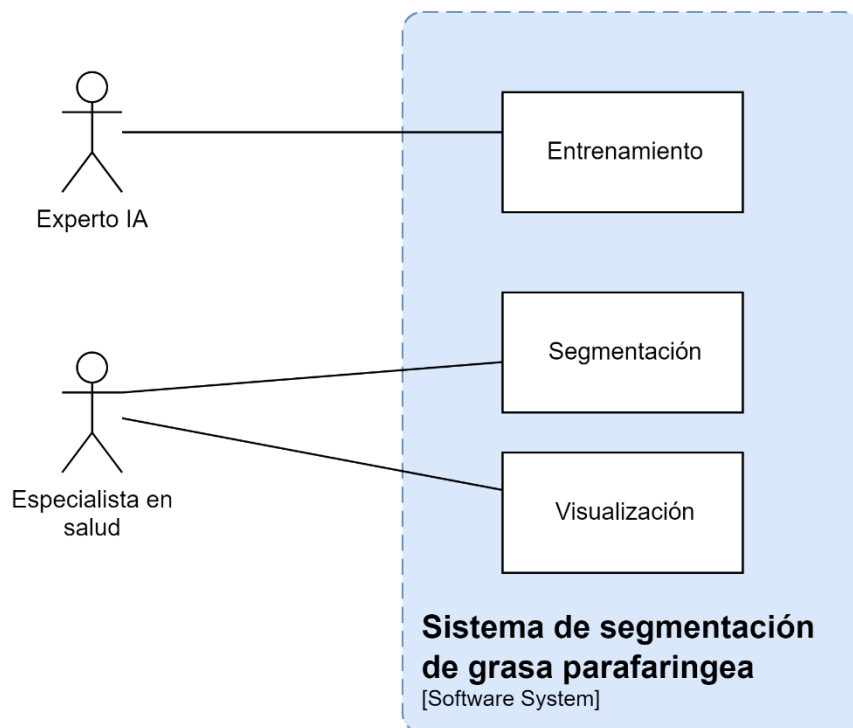
5.1 Análisis de la solución

En esta sección se presenta la descripción de los productos obtenidos en la etapa de definición de objetivos, alternativas, restricciones y riesgos, describiendo los requerimientos del sistema, haciendo el análisis de estos y caracterizando el proceso de detección y segmentación que se llevó a cabo.

5.1.1 Requerimientos del sistema

En esta sección se presenta una breve descripción de los requerimientos del sistema cuya especificación se podrá ver en el documento de Especificación de Requerimientos del Sistema de Segmentación de Grasa Parafaríngea (SRS).

Figura 5 Diagrama del Sistema de Segmentación de Grasa Parafaríngea



El sistema presenta tres funcionalidades principales como se puede observar en la Figura 5 las cuales son el **entrenamiento** de la red neuronal convolucional u-net, la **segmentación** que para el prototipo propuesto hace uso de un modelo que implementa una arquitectura u-

net 3D y la visualización de imágenes médicas que no se consideró para la implementación del prototipo. Los actores que hacen uso del sistema son el experto en IA el cuál puede realizar el entrenamiento del modelo configurando sus hiperparámetros y verifica el rendimiento del modelo obtenido en el entrenamiento. El otro actor es el Especialista en salud que se encarga de ingresar las imágenes médicas al sistema para su segmentación y visualización.

Los requerimientos funcionales del sistema que se implementan en el prototipo se pueden observar en la Tabla 2. En el presente documento se omite el requerimiento relacionado con la visualización de imágenes ya que no es una prioridad para el proyecto.

Tabla 2 *Requerimientos funcionales*

ID	Nombre	Prioridad
RF-1	Segmentar imagen	Alta
RF-2	Entrenar modelo	Alta
RF-3	Preprocesar imagen	Media
RF-4	Configurar entrenamiento	Media
RF-5	Configurar modelo	Media
RF-6	Mostrar avance del entrenamiento	Baja

Los requerimientos no funcionales se definieron a partir del modelo de calidad del producto en (ISO/IEC 25010, 2011). El modelo de calidad establece varias características de calidad que a su vez se dividen en subcaracterísticas. Las características y subcaracterísticas que se tuvieron en cuenta para el análisis de los requerimientos del prototipo se pueden observar en la Tabla 3.

Tabla 3 *Características seleccionadas para el modelo de calidad*

Característica	Subcaracterística	Descripción
Adecuación funcional	Pertinencia Funcional	Capacidad del producto software para proporcionar un conjunto apropiado de funciones para tareas y objetivos de usuario especificados. (ISO/IEC 25010, 2011).
Compatibilidad	Coexistencia	Capacidad del producto para coexistir con otro software independiente, en un entorno común, compartiendo recursos comunes sin detrimento. (ISO/IEC 25010, 2011).
Usabilidad	Operabilidad	Capacidad del producto que permite al usuario operarlo y controlarlo con facilidad. (ISO/IEC 25010, 2011).
	Protección contra errores de usuario	Capacidad del sistema para proteger a los usuarios de cometer errores (ISO/IEC 25010, 2011).
Portabilidad	Capacidad para ser instalado	Facilidad con la que el producto se puede instalar y/o desinstalar de forma exitosa en un determinado entorno (ISO/IEC 25010, 2011)..

Los requerimientos no funcionales que se definieron a partir de las características de calidad se pueden observar en la Tabla 4.

Tabla 4 *Requerimientos no funcionales*

ID	Nombre	Característica	Prioridad
RNF – 1	Facilidad de uso	Usabilidad	Alta
RNF – 2	Control de errores del usuario	Usabilidad	Baja
RNF – 3	Advertir de errores al usuario	Usabilidad	Baja
RNF - 4	Portabilidad	Portabilidad	Media

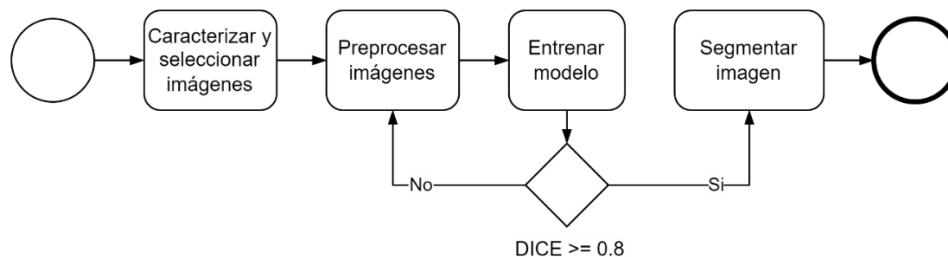
Las restricciones técnicas establecidas son las siguientes:

- **R-1:** El sistema debe ser implementado en Python 3, debido a los diferentes frameworks que se pueden utilizar para implementar funcionalidades relacionadas con inteligencia artificial.
- **R-2:** El sistema debe ser implementado como aplicación de escritorio.
- **R-3:** El sistema debe leer imágenes de resonancia en formato Nifti
-

5.1.2 Caracterización del proceso de segmentación

En esta sección se describe al proceso que se llevó a cabo para realizar la segmentación de grasa parafaríngea como se puede observar en la Figura 6 y los artefactos producto de cada tarea en el proceso.

Figura 6 *Proceso de segmentación de grasa parafaríngea*



En la tarea **Caracterizar y seleccionar imágenes** el objetivo es extraer información de las imágenes relacionada con el tamaño de la imagen, tamaño de los voxels o spacing e intensidades. El producto que se obtuvo en esta tarea es una tabla consolidando mínimos, máximos y media de cada uno de los ítems mencionados anteriormente. A partir de la información de las imágenes se seleccionó de forma aleatoria la muestra de 64 imágenes que hacen parte del estudio.

En la tarea **Preprocesar imágenes** se preparan las imágenes de la muestra y sus segmentaciones manuales previamente verificadas por un radiólogo para ser utilizadas en el entrenamiento y validación del modelo, recortando las imágenes para que tengan un tamaño uniforme y normalizando las intensidades. El producto de esta tarea es el conjunto de imágenes y segmentaciones preprocesadas.

En la tarea **Entrenar modelo** se utilizan las imágenes preprocesadas y los hiperparámetros de la configuración del modelo, se entrena el modelo teniendo como función de optimización la minimización de la pérdida de coeficiente de DICE que implica la maximización del coeficiente. Si en la validación del modelo se obtiene un coeficiente DICE medio mayor o igual a 0.8 se considera finalizado el entrenamiento. En caso de que no se obtenga el coeficiente Dice esperado se procede a realizar un nuevo preprocesamiento de imágenes y modificación de hiperparámetros. El producto resultante de esta tarea son los pesos de la red neuronal que se utilizan en el modelo para la segmentación de nuevas imágenes.

En la tarea **Segmentar imagen** el modelo en funcionamiento recibe una imagen y devuelve la segmentación de la grasa para faríngea en dicha imagen.

5.1.3 Población y Muestra

Para el cálculo del tamaño muestral se utiliza la formula simplificada planteada en (García-García, Reding-Bernal, & López-Alvarenga, 2013), en la cual se acepta un error bilateral alfa del 5% y una potencia estadística del 80%. La ecuación se muestra a continuación:

$$n = \frac{16}{(DE)^2}$$

Donde DE es la diferencia estandarizada, es decir, la diferencia de medias sobre la desviación estándar. Buscando una DE aproximada de 0.5 se tiene que el tamaño muestral para el presente proyecto es de 64. Por esta razón se tomarán 32 imágenes de resonancia magnética de pacientes con apnea obstructiva de sueño y 32 imágenes de resonancia magnética de pacientes sin apnea obstructiva de sueño de una población de 142 imágenes de resonancia magnética obtenidas en el proyecto de investigación titulado *Generación de biomarcadores para el estudio de cambios neurocognitivos asociados a fenómenos cerebrovasculares en individuos con apnea obstructiva del sueño*.

5.2 Diseño del prototipo

En esta sección se describe el diseño arquitectural del prototipo, que se realizó a partir de los requerimientos funcionales, no funcionales y restricciones. También se presenta el diseño de la red neuronal convolucional U-net 3D utilizado en la implementación del prototipo basado en el diseño original presentado en (Çiçek et al., 2016), seguido por los pasos que van desde la selección y caracterización de las imágenes hasta el flujo de preprocesamiento de estas.

5.2.1 Arquitectura del prototipo

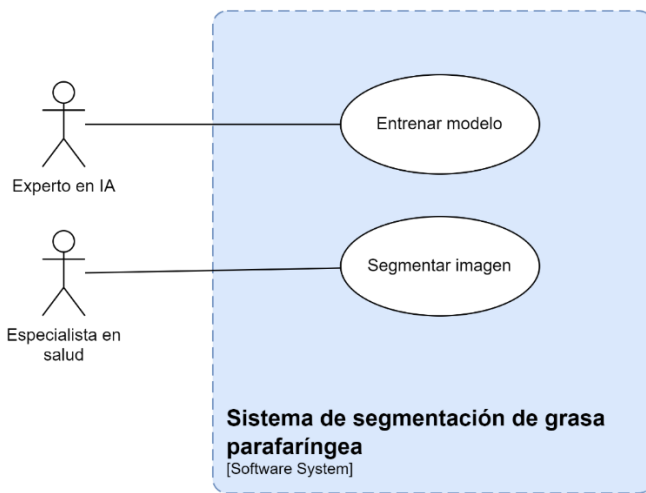
En esta sección se presentan los casos de uso derivados a partir de los requerimientos y los escenarios de calidad, siguiendo una metodología de diseño dirigido por atributos (ADD) propuesta en (Wojcik et al., 2006). También se presentan los patrones arquitecturales usados y algunas vistas arquitecturales.

5.2.1.1 Casos de Uso

Esta sección resalta los casos de uso para los dos actores del sistema de forma separada. Cómo se puede observar la

Figura 7 cada actor está relacionado con un caso de uso y se omitió la funcionalidad de visualización.

Figura 7 Casos de uso del sistema



En el caso de uso **Entrenar modelo** el Experto en IA ingresa las imágenes y su respectiva segmentación, ejecuta el preprocesamiento, ejecuta el entrenamiento y obtiene un archivo con los pesos y datos y entrenamiento del modelo.

En el caso de uso **Segmentar imagen** el Especialista en salud carga la imagen en el sistema, el sistema preprocesa la imagen, la segmenta y guarda la segmentación de la imagen en un archivo.

Para ver la especificación de los casos de uso consultar el SRS.

5.2.1.2 Escenarios de atributos de Calidad

Los escenarios de atributos de calidad y sus requerimientos no funcionales relacionados se describen en la Tabla 5.

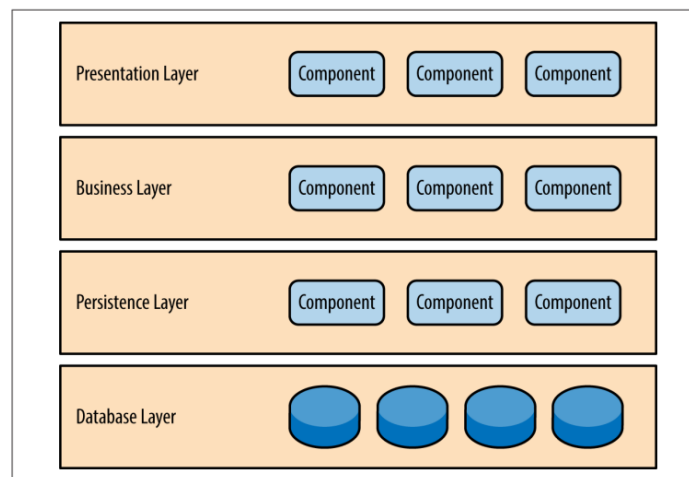
Tabla 5 Escenarios de atributos de calidad

ID	Requerimiento No Funcional	Descripción	Prioridad
AC-1	RNF -1	El usuario al iniciar la aplicación debe poder realizar la segmentación o el entrenamiento máximo con 10 acciones.	Alta
AC-2	RNF- 2	El usuario debe recibir una pregunta de confirmación al seleccionar la opción de realizar el entrenamiento ya que el resultado puede modificar la funcionalidad de segmentación.	Baja
AC-3	RNF-3	El usuario debe recibir una alerta si no selecciona una imagen o un directorio de salida.	Baja
AC-4	RNF-3	El usuario debe recibir una alerta si no se encuentra algún archivo de configuración.	Baja
AC-5	RNF-4	El usuario debe poder usar el sistema en Windows 10 y Windows 11	Media

5.2.1.3 Patrón de arquitectura

Para el diseño del prototipo del sistema se utilizó el patrón de arquitectura en capas para la división del agrupamiento de los componentes. El patrón de arquitectura en capas es uno de los más usados, en este los componentes son organizados en capas horizontales, cada capa con un rol específico, como por ejemplo la capa de presentación o la capa de lógica de negocio (Richards, 2015). En la Figura 8 se puede observar un ejemplo del patrón en el cuál un sistema tiene 4 capas, una de presentación para el manejo de la interfaz de usuario, una capa que maneja la lógica de negocio, otra capa que maneja la persistencia de datos y otra en la que se encuentra una base de datos.

Figura 8 Patrón de arquitectura en capas

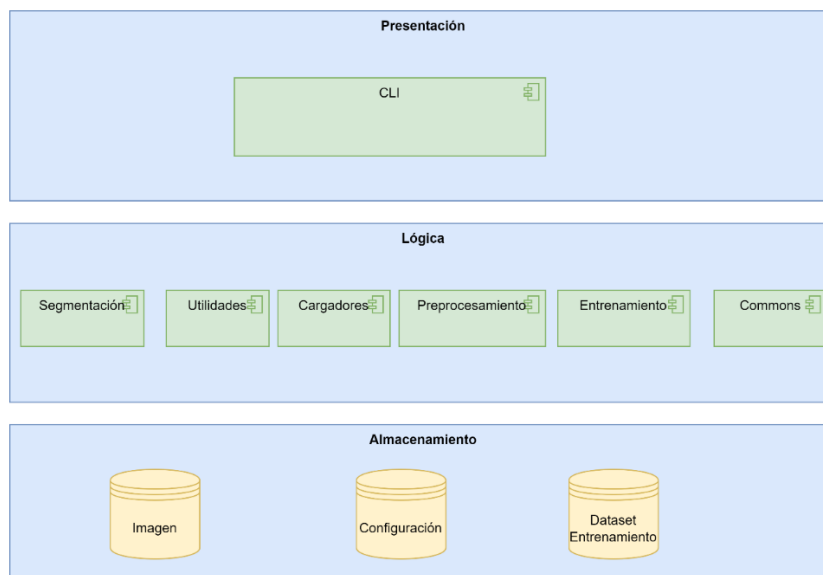


Nota. Adaptado de *Software Architecture Patterns*, (p. 2) por M. Richards, 2015, O'Reilly

5.2.1.4 Vista Lógica

A continuación, se muestra la vista lógica, en la cual se pueden observar los diferentes componentes generales del prototipo.

Figura 9 Vista lógica del prototipo



En la Figura 9 se puede observar la vista lógica del prototipo en la cual se utiliza una arquitectura de 3 capas y 1 nivel. En la **capa de presentación** se tiene un módulo de interfaz de línea de comandos (CLI) la cual sirve como interfaz de usuario para las diferentes funcionalidades.

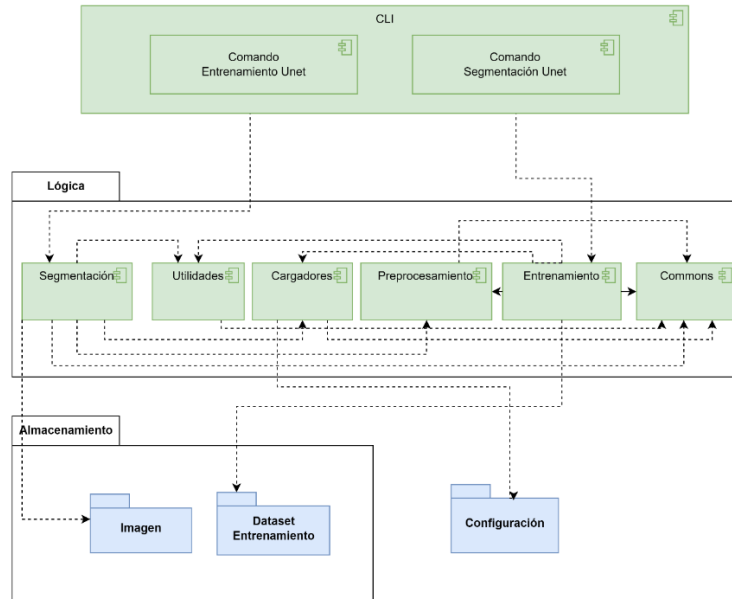
En la **capa lógica** se tienen los componentes de Segmentación, Utilidades, Cargadores, Entrenamiento y Commons. Esta capa está diseñada para a tener la lógica que permite la segmentación por medio de una red neuronal, con el componente de segmentación. El componente de entrenamiento está encargado del entrenamiento de un modelo, el de cargadores está orientado a la obtención de los datos, representados en imágenes, el preprocesamiento de estos está a cargo del componente de preprocesamiento. El componente de utilidades contiene funcionalidades de uso común como la configuración del modelo, el cálculo de la métrica de Dice y el manejo de la configuración del entrenamiento. Por último, se encuentra el componente commons que se funciona como interfaz para acceder a los componentes en librerías externas.

En la **capa de almacenamiento** se utiliza el sistema de archivos para el almacenamiento del conjunto de datos (dataset) de entrenamiento, la configuración utilizada por el modelo y la imagen de entrada al módulo de producción y su correspondiente segmentación de salida.

5.2.1.5 Vista de implementación

En esta sección se describe la vista de implementación, en la cual se da a conocer los componentes de la solución, su organización en paquetes y sus dependencias.

Figura 10 Vista de implementación del prototipo



En la Figura 10 se puede observar la vista de implementación del prototipo en la cual se hace la división en paquetes relacionados con la capa a la que pertenecen cada uno de sus componentes, a excepción del componente de la CLI y la Configuración. Por fuera de los paquetes se encuentra la CLI la cual se encuentra compuesta por los componentes de comandos de entrenamiento Unet y segmentación Unet.

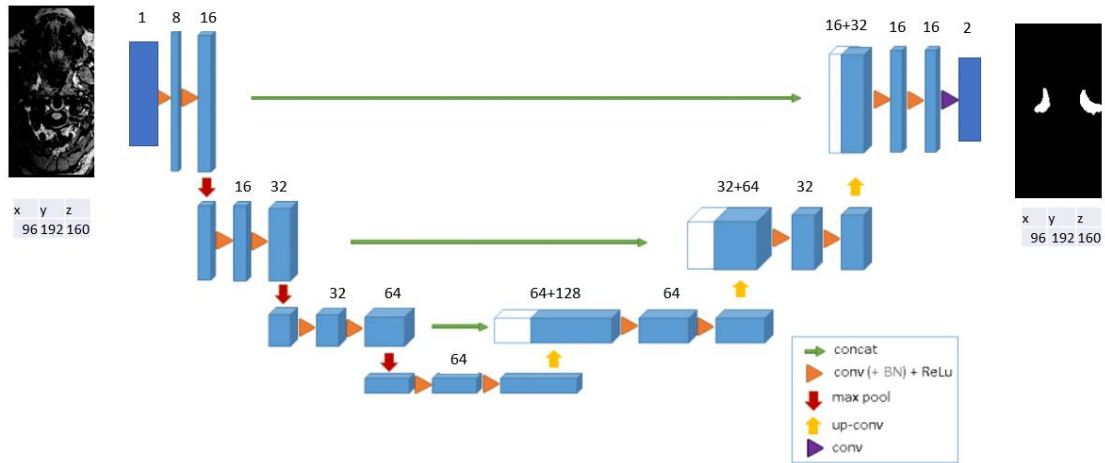
El componente comando de entrenamiento contiene la lógica que permite acceder a la ejecución del entrenamiento del modelo que requiere de un dataset de entrenamiento, una configuración del modelo para realizar el entrenamiento y que guarda los pesos del modelo y configuraciones de preprocesamiento que se utilizarán durante la segmentación productiva.

El componente comando de segmentación contiene la lógica que permite acceder a la ejecución la cual requiere el uso de la configuración del modelo, usada en el entrenamiento, los pesos del modelo y una imagen de entrada que será segmentada.

Arquitectura de la red neuronal

La arquitectura de la red U-net 3D diseñada en este proyecto se encuentra basada en la descrita en (Çiçek et al., 2016), pero por detalles de implementación se realizó una reducción de los canales de salida de cada capa pasando de 32 a 8 canales salida de la primera capa y utilizando la función de pérdida de coeficiente DICE la cual viene dada por $1 - DICE$, buscando minimizar la pérdida para obtener mejores resultados.

Figura 11 Arquitectura U-net 3D implementada



Nota. Figura creada a partir de Figura 3

En la Figura 11 se puede observar el diagrama de la arquitectura de la red neuronal implementada, en esta arquitectura a parte de la diferencia entre los canales de salida también se maneja un número de clases distinto, en este caso se manejan 2 clases, el fondo y los fat pads. La imagen de entrada a la red neuronal es del tamaño de 96x192x160 voxels del mismo modo que la de salida. Para la implementación de la arquitectura diseñada se utilizó la librería unet, que permite la implementación de redes U-net en dos y tres dimensiones, propuesta por (Pérez-García, 2020).

6 Implementación

En esta sección se describe el proceso llevado a cabo para la implementación de la red neuronal como la caracterización, selección y preprocesamiento de las imágenes y se muestran capturas de pantalla del prototipo en funcionamiento durante el entrenamiento y la segmentación.

6.1.1 Caracterización de las imágenes

Se obtuvieron 142 imágenes de resonancia magnética las cuales fueron compartidas por el Hospital Universitario San Ignacio, de las cuales se excluyeron 2 por no tener la calidad suficiente para el estudio debido a artefactos o baja resolución. A partir de las restantes 140 se realizó una caracterización de las imágenes obteniendo los valores mínimos, máximos y medias para los tamaños en los ejes x, y, z, el spacing en cada uno de los ejes y las intensidades mínimas y máximas encontradas en las imágenes de resonancia de magnética.

Tabla 6 Caracterización de imágenes consolidada

	x	y	z	Sx	Sy	Sz	IMin	Imax
Media	341	484	484	0.50	0.50	0.50	0	13980
Mínimo	320	480	480	0.50	0.48	0.48	0	781
Máximo	360	528	528	0.50	0.50	0.50	0	65533

En la Tabla 6 se observan los valores anteriormente descritos de la siguiente manera:

- **X:** Cantidad de voxels en el eje x
- **Y:** Cantidad de voxels en el eje y
- **Z:** Cantidad de voxels en el eje z
- **Sx:** Spacing eje x
- **Sy:** Spacing eje y
- **Sz:** Spacing eje z
- **Imin:** Intensidad mínima
- **Imax:** Intensidad máxima

6.1.2 Selección de las imágenes

Para realizar la selección de las imágenes se tuvo en cuenta el grupo al que pertenecía el sujeto de la imagen que puede ser paciente o control, sin tener en cuenta edad y sexo de los sujetos de estudio, ya que solo se tiene disponible esos datos para algunos de los sujetos del

estudio. Los pacientes son el grupo de personas que han sido diagnosticados con apnea, mientras que los sujetos pertenecientes al grupo de control, no. Cada sujeto del estudio tiene 4 imágenes de resonancia: resonancia T1 de alta resolución, resonancia T1 de baja resolución, resonancia T2 y resonancia T2 FLAIR. Las imágenes seleccionadas para esta investigación, cuya caracterización fue presentada anteriormente, son las T1 de alta resolución ya que en estas el tejido graso presenta altas intensidades y lo hacen mayormente visible y permiten obtener más información sobre la imagen que las de baja resolución.

La muestra de 64 imágenes utilizadas en la investigación se escogió de forma aleatoria teniendo en cuenta que 32 fueran del grupo de control y 32 del grupo de pacientes. Las imágenes de la muestra descrita fueron segmentadas de forma manual utilizando el software ITK Snap por parte de un grupo de radiólogos.

6.1.3 Segmentación manual de imágenes

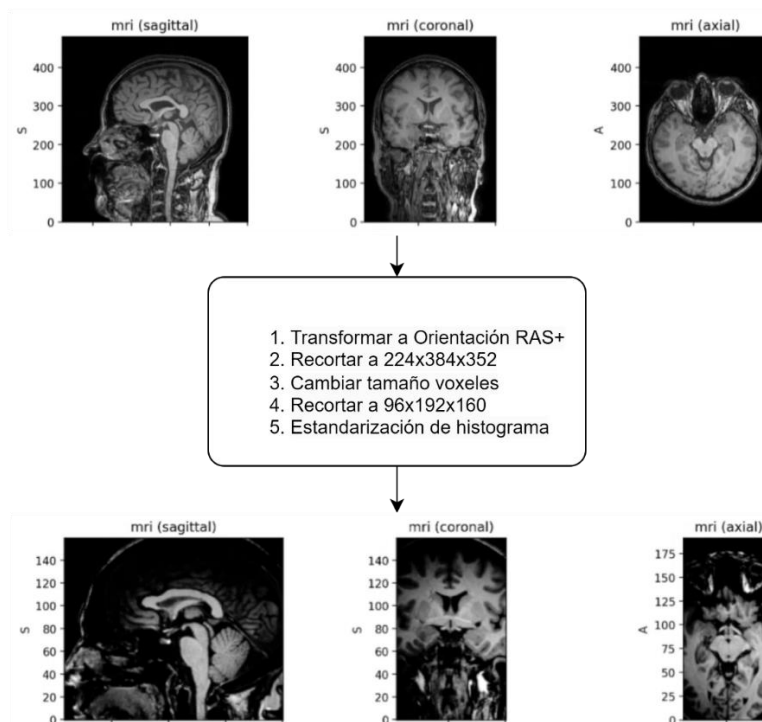
La segmentación manual de imágenes fue realizada por un equipo conformado por una especialista en odontología y un residente de radiología. Para realizar la segmentación manual se probaron las herramientas ITK Snap y 3D Slicer y se seleccionó ITK Snap por la facilidad de uso según lo indicado por los miembros del equipo mencionado anteriormente. Inicialmente fueron segmentadas algunas imágenes de la muestra corte por corte, pero al visualizar el volumen segmentado en 3 dimensiones se observó que la superficie era muy irregular.

A partir de lo anterior se procedió con una segmentación manual asistida por interpolación en la cual la doctora segmentó manualmente algunos cortes que consideró relevantes en las 64 imágenes, segmentando como mínimo los cortes superior e inferior de la grasa parafaríngea y se usó un método de interpolación para segmentar el espacio entre esos cortes. El método utilizado es la interpolación morfológica de contornos (Zukić et al., 2016), que es la implementación provista por ITK Snap del enfoque para la interpolación entre cortes basada en la morfología propuesto en (Albu, Beugeling, & Laurendeau, 2008)

6.1.4 Preprocesamiento de imágenes

Debido a los detalles de implementación que se describirán en la sección de entrenamiento y segmentación se implementó el flujo de preprocesamiento que se puede observar en la Figura 12.

Figura 12 Transformaciones preprocesamiento



La transformación de Orientación RAS+ permite cambiar la orientación de las imágenes a una orientación estándar RAS(Right-Anterior-Superior) del eje de coordenadas desde la orientación LPI(Left-Posterior-Inferior) que tenían las imágenes.

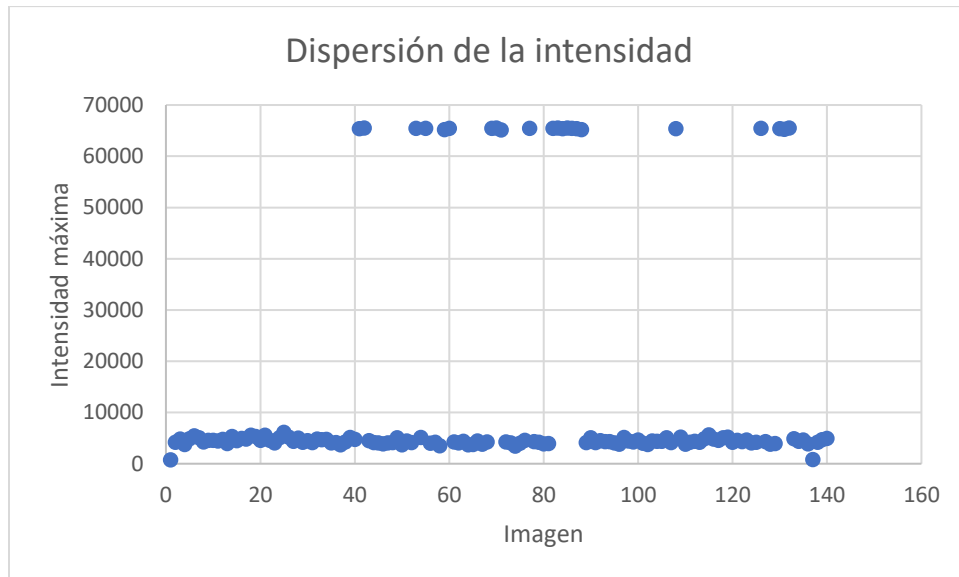
El recorte de la imagen a un tamaño de 224x384x352 voxels se realizó para reducir la imagen a una región de interés donde se encuentran los fat pads.

El cambio del tamaño o espacio de los voxels a 1x1x1 se realizó teniendo en cuenta que todas las imágenes tienen un espacio de voxels de 0.5x0.5x0.5 o muy cercanos a este valor para reducir el tamaño de las imágenes a la mitad. Esta transformación se realizó mediante submuestreo utilizando interpolación lineal, para las imágenes de resonancia magnética y de vecino más cercano para las imágenes etiquetadas, es decir, para la correspondiente segmentación manual.

El último recorte a un tamaño de 96x192x160 se realizó para ajustar el tamaño a múltiplos de 32 para lograr compatibilidad con la red neuronal evitando afectar al máximo la región de interés donde se ubican los fat pads.

Para las segmentaciones de entrenamiento y validación se aplicó la transformación de codificación oneHot, para evitar que el sistema tome el valor de las etiquetas (0: fondo, 1: grasa parafaríngea) como valores numéricos, haciéndolos categóricos.

Figura 13 Gráfica de dispersión de intensidades máximas



La última transformación de aplicada a las imágenes fue la estandarización del histograma, debido a la dispersión de la intensidad máxima identificada en la caracterización de las imágenes y que se puede observar en la Figura 13. A partir de la estandarización se obtienen valores de intensidad con una mayor similitud para cada una de las imágenes y es independiente de los pacientes y de la marca del escáner de resonancia magnética usado (Nyul, Udupa, & Zhang, 2000). Las transformaciones aplicadas descritas en esta sección son parte de la implementación propuesta por (Pérez-García, Sparks, & Ourselin, 2021).

6.1.5 Prototipo

El prototipo se implementó de acuerdo con las restricciones establecidas en el lenguaje de programación Python 3. La interfaz de usuario usada es una interfaz de línea de comando adaptada para que los comandos sean fáciles de usar mediante opciones numéricas. En esta sección se van a mostrar la vista inicial, las vistas intermedias y la vista resultante después de ejecutar cada una de las opciones.

En la Figura 14 se observa la vista inicial del prototipo que el usuario ve el iniciar la aplicación, se presentan 3 opciones, la primera permite cerrar la aplicación, la segunda permite iniciar la segmentación de una imagen y la tercera permite iniciar el entrenamiento.

Por último, se presenta un espacio en el cual el usuario puede escribir el número del comando a utilizar y que se ejecutará cuando el usuario oprima la tecla Enter.

Figura 14 Vista inicial del prototipo

```

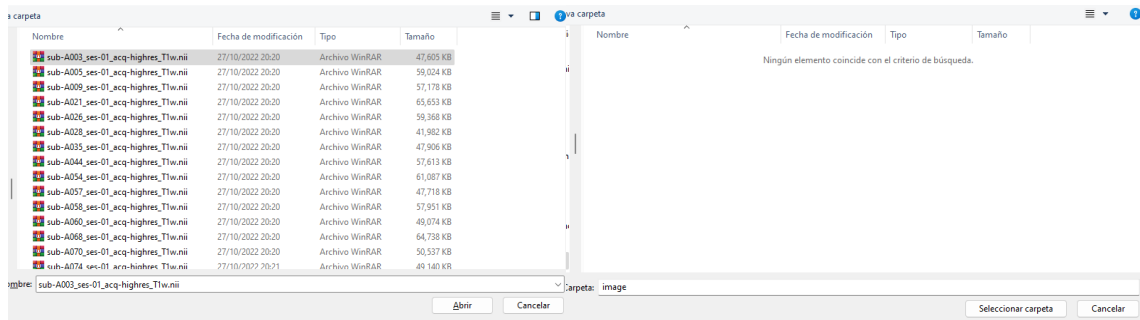
Welcome to model training and image segmentation prototype!

1-exit
2-segment-unet
3-train-model-unet
Please enter a command number:

```

En la Figura 15 se observan las vistas de selección de la imagen a segmentar y la vista de selección de la carpeta donde se guardará la segmentación. La aparición de estas vistas es secuencial, primero se selecciona la imagen a segmentar, la vista se cierra y luego aparece la imagen para seleccionar la ubicación de la segmentación y se cierra.

Figura 15 Vistas intermedias segmentación



Nota. A la izquierda: Vista de selección de imagen, a la derecha: vista de selección de carpeta de salida

En la Figura 16 se muestra la vista final después de la segmentación, en esta vista se puede observar el volumen calculado de la segmentación y un mensaje indicando que la imagen fue segmentada de forma exitosa. El usuario puede encontrar el resultado de la segmentación en la carpeta previamente indicada.

Figura 16 Vista final segmentación

```

1-exit
2-segment-unet
3-train-model-unet
Please enter a command number: 2
Dataset size: 1 subjects
The segmentation volume is: 8544 mm^3
The image: C:/Users/Usuario/Documents/.../images/image/sub-A003_ses-01_acq-highres_T1w.nii.gz
was segmented successfully.

```


se guardan en la ubicación definida en el archivo de configuración, ambos archivos pueden ser usados posteriormente en un entorno tipo laboratorio para poder visualizar los resultados.

El archivo de configuración es un archivo con formato **yaml**, en el cual se definen parámetros del entrenamiento y del modelo, que por defecto son los usados en la prueba de la red U-net 3D con activación ReLU. A continuación, se muestra la configuración por defecto:

```
model:
  out_classes: 2
  dimensions: 3
  num_encoding_blocks: 4
  out_channels_first_layer: 8
  activation: ReLU

training:
  num_epochs: 1
  training_split_ratio: 0.7
  training_data_path: config/whole_image_state_dict.pth
  training_weights_path: config/whole_image_state_dict_80_epochs_64.pth
  landmarks_path: config/landmarks.npy
```

La explicación de cada uno de los parámetros se encuentra en el mismo archivo, el cual se ubica en la carpeta **config** del proyecto y tiene como nombre **config.yaml**.

Las tecnologías y herramientas que se usaron en la implementación son las siguientes:

- Python 3: Lenguaje de programación.
- PyTorch: Es una librería que hace uso de tensores para aprendizaje profundo tanto en GPU como en CPU (Contributors, n.d.).
- Torchio: Es una librería que permite la carga, preprocesamiento, aumento y muestreo basado en parches en imágenes médicas. (Pérez-García et al., 2021)
- Unet: Es una librería que implementa la arquitectura U-net tanto para 2D como para 3D (Pérez-García, 2020).
- Click: Es un paquete en Python que permite crear interfaces de línea de comandos (Pallets, n.d.).
- PyYAML: Es una librería en Python que permite el manejo, lectura y escritura de archivos en formato YAML, el cual es un formato de serialización de datos diseñado para la legibilidad humana y la interacción con lenguajes de programación (“PyYAML · PyPI,” n.d.).
- Schema: Es una librería en Python que permite la validación de estructuras de datos (NumPy Developers, n.d.). Utilizado para validar el archivo de configuración.
- Tqdm: Es una librería en Python que permite implementar barras de progreso. (“Tqdm · PyPI,” n.d.)

- NumPy: Es una librería en Python que provee funcionalidad para operaciones algebraicas multidimensionales, matemáticas y estadística, entre otros (“NumPy Documentation — NumPy v1.24 Manual,” n.d.)

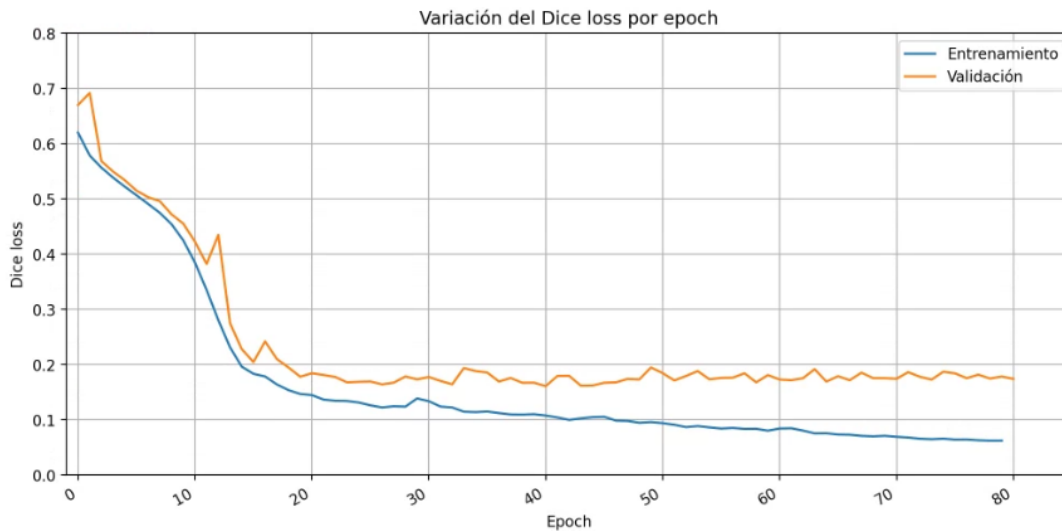
7 Pruebas y resultados

En esta sección se presentan las pruebas realizadas y los resultados del entrenamiento, segmentación y cálculo del volumen de grasa tanto en pacientes con AOS como en el grupo de control.

7.1 Entrenamiento y Segmentación

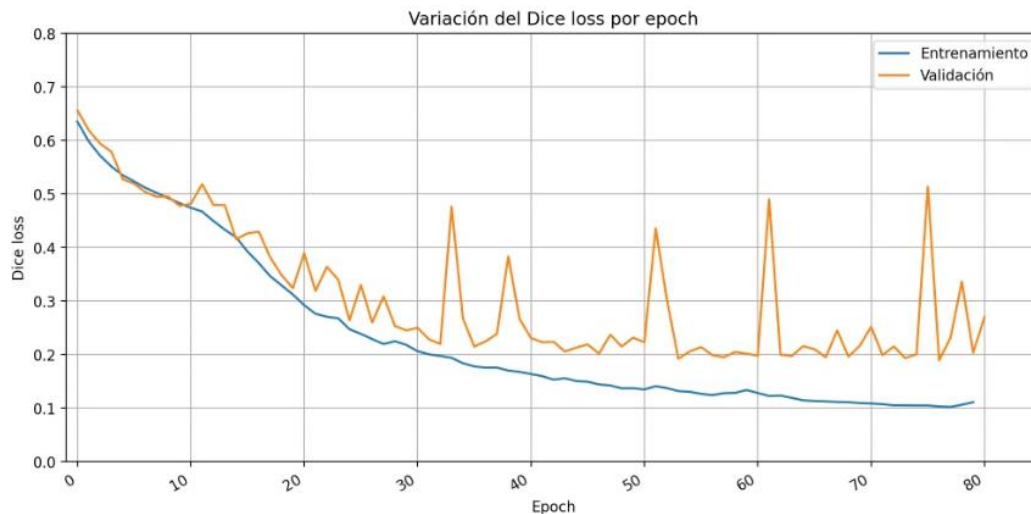
Para la realización del entrenamiento y segmentación se utilizó un equipo con procesador Intel Xeon W-10885M CPU de 8 núcleos de 2.40GHz, con 32 GB de RAM. Debido al límite de memoria se aplicaron las transformaciones, para la reducción del tamaño de las imágenes, durante la etapa de preprocesamiento descritas anteriormente. Para el entrenamiento del modelo se utilizaron 44 imágenes, que corresponden al 70% de la muestra y para la validación se utilizaron 20 imágenes, que corresponden al 30% de la muestra. Para la prueba con la cual se evaluó la segmentación se utilizó el 100% de la muestra. En total se realizaron 2 pruebas una entrenando un modelo con función de activación ReLu y otra con función de activación Sigmoide.

Figura 20 *Perdida de DICE en el entrenamiento y validación con ReLu*



En la Figura 20 se observa la variación de la métrica de pérdida de Dice utilizada para el entrenamiento y validación del modelo con arquitectura U-net 3D con la función de activación ReLu, la cuál es la función usada en la definición original de dicha arquitectura.

Figura 21 Perdida de DICE en el entrenamiento y validación con Sigmoide



En la Figura 21 se puede observar la variación de la métrica de pérdida de Dice utilizada para el entrenamiento y validación del modelo con arquitectura U-net 3D con la función de activación Sigmoide. La función sigmoide se utilizó para verificar si se podía obtener una mejoría en la segmentación con respecto a ReLu debido a las características morfológicas de la región segmentada.

Tabla 7 Comparación de métricas para la validación, entrenamiento y prueba

	Dice Loss Tr (SD)	Dice Loss Val (SD)	DICE (SD)
ReLu	0.0624 (0.0076)	0.1786 (0.0172)	0.8086 (0.1152)
Sigmoide	0.1112 (0.0120)	0.2038 (0.0120)	0.6347 (0.1433)

En la Tabla 7 se observa la consolidación de las métricas obtenidas en el entrenamiento (Dice Loss Tr) y validación (Dice Loss Val), la pérdida de Dice con su desviación estándar (SD) y la media de coeficiente DICE obtenido en las pruebas con su SD.

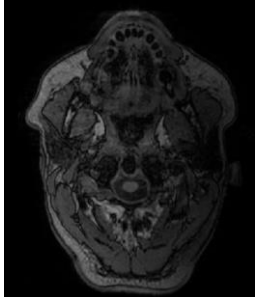
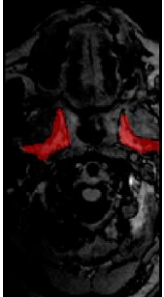
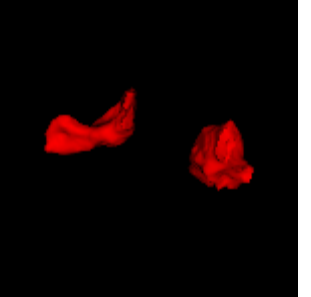
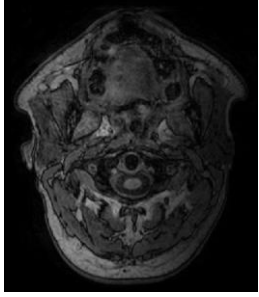
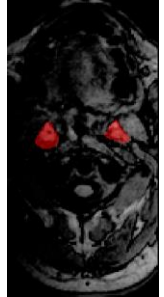
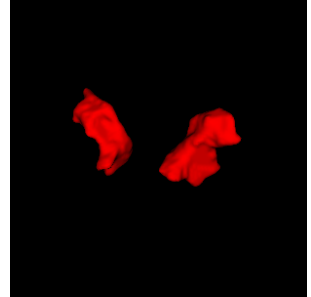
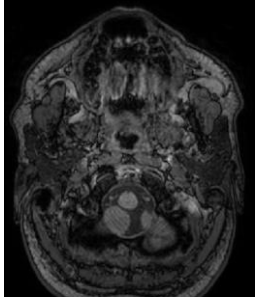
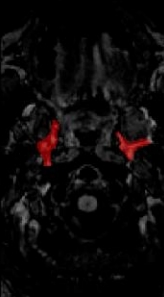

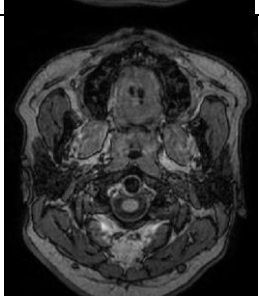
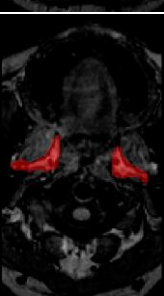

Tabla 8 Comparación volúmenes calculados

Grupo\Segmentación	Manual	Unet-3D
Promedio Pacientes	13812.7003	12577.3750
Promedio Control	10383.6875	10629.0313
Promedio General	12098.1939	11603.2031

En la Tabla 8 se puede observar el volumen promedio de la grasa parafaríngea calculado para cada uno de los grupos a los que pertenecen las imágenes, paciente y control, y el promedio de todas las imágenes. Estos cálculos se realizaron a partir de las segmentaciones obtenidas

en las pruebas usando la función de activación ReLu y a partir de la segmentación manual verificada por el grupo de radiólogos.

Tabla 9 *Imágenes segmentadas y volumen*

Imagen Original	Imagen segmentada con U-net 3D	Volumen segmentado con U-net 3D	Volumen Segmentación (mm^3)
			15375
			11608
			9292
			10321

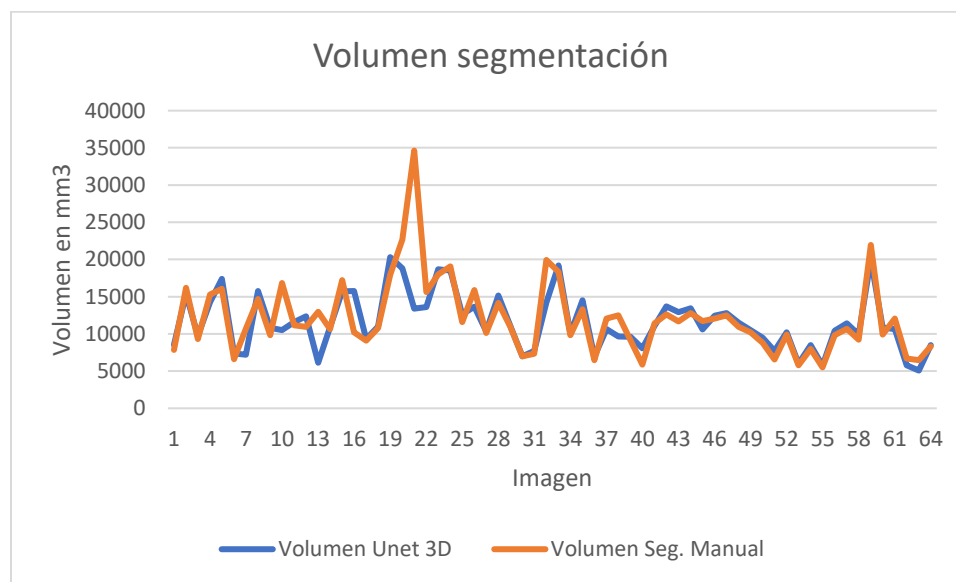
En la Tabla 9 se muestran algunas imágenes usadas en el estudio, la segmentación sobrepuesta sobre la imagen preprocesada, la visualización del volumen segmentado y el volumen calculado. Estas imágenes fueron tomadas a partir de la visualización de estas usando ITK Snap.

8 Discusión

En la presente sección se realizará un análisis de los resultados descritos anteriormente y su comparación con estudios de segmentación de grasa parafaríngea descritos en la revisión de literatura. Como se puede observar en la presentación de los resultados, el comportamiento de la métrica de pérdida de Dice durante el entrenamiento y validación de la red neuronal U-net 3D implementada (Figura 20 y Figura 21) fue similar para la función de activación ReLU y sigmoide, pero con mejores resultados para la primera, lo cual se ve al obtener los coeficientes Dice en la prueba realizada, con 0.8086 para ReLU y 0.6063 para la sigmoide.

A partir de las segmentaciones obtenidas y teniendo en cuenta el coeficiente DICE con mayor valor, se realizó una comparación del volumen de las segmentaciones manuales y la segmentación con U-net 3D con activación ReLU.

Figura 22 *Volumen segmentación de cada imagen*



En la Figura 22 se puede observar volumen calculado para cada una de las segmentaciones manuales y automáticas, observando una similitud. A partir de los datos de los volúmenes se calculó un coeficiente de correlación entre los dos conjuntos de datos de 0.75, lo cual indica una correlación entre el volumen de la segmentación manual y automática, pero sin significancia estadística comprobada.

En cuanto a la comparación con estudios similares se tuvo en cuenta el estudio realizado por (Shahid et al., 2017), que utiliza un método automático basado en análisis de componentes para segmentar la grasa para faríngea y el estudio hecho por (Bommineni et al., 2022), en que utilizan una red neuronal U-net modificada que usa elementos de Resnet para segmentar diferentes estructuras en la vía aérea superior entre los que se encuentra la grasa parafaríngea.

En la Tabla 10 se puede observar el coeficiente Dice obtenido por los métodos utilizados en el presente estudio y en los estudios mencionados.

Tabla 10 Comparación resultados estudios

Estudio	Coefficiente DICE (SD)
U-net 3D ReLU (implementación propuesta)	0.809 (0.115)
(Shahid et al., 2017)	0.779 (0.041)
(Bommineni et al., 2022)	0.580 (0.120)

El mayor coeficiente DICE se obtuvo en el presente estudio, pero con una desviación estándar mayor a la obtenida por (Shahid et al., 2017), lo cual permite interpretar que el método usado en este estudio es menos exacto al del otro estudio pero que se podría mejorar al aumentar el tamaño de la muestra. Comparado con el estudio de (Bommineni et al., 2022) se puede interpretar que el método propuesto usando U-net 3D obtuvo un mejor resultado en cuanto al coeficiente DICE y una desviación estándar similar. Cabe resaltar que los estudios comparados utilizaron distintos conjuntos de datos, número de muestras y métodos de preprocesamiento o no usaron como en (Bommineni et al., 2022). En la Tabla 11 se pueden observar algunas de las características que permiten establecer una diferencia en los estudios comparados.

Tabla 11 Comparación de características de los estudios

Estudio	Número de imágenes	Tamaño de imágenes	Regiones de interés
U-net 3D ReLU (Implementación propuesta)	64	341x484x484	Grasa parafaríngea
(Shahid et al., 2017)	30	176x256x176	Grasa parafaríngea
(Bommineni et al., 2022)	206	512x512x58	<ul style="list-style-type: none"> • Mandíbula • Vía aérea retropalatina • Vía aérea retroglosa • Paladar blando • Lengua • Otra lengua • Epiglotis • Grasa parafaríngea • Paredes laterales retropalatinas • Paredes laterales retroglosales

9 Conclusiones y trabajos futuros

1. El prototipo desarrollado permite la segmentación automática del tejido graso en la región parafaríngea utilizando una red neuronal Unet 3D.
2. La metodología basada en la ciencia del diseño utilizada en este trabajo permitió articular diferentes métodos y hacer uso de la base de conocimiento para construir un prototipo que se podrá utilizar en futuras investigaciones relacionadas con la AOS.
3. El uso de la red neuronal Unet 3D se encuentra restringido o condicionado para el entrenamiento por las capacidades de la máquina en la que se realiza su ejecución, por tal motivo, el preprocesamiento de las imágenes estuvo enfocado en la disminución del tamaño de las imágenes mediante recorte y re-espaciado.
4. La caracterización de las imágenes permitió identificar la diferencia en las intensidades de los voxels, lo cual conllevó a la estandarización del histograma.
5. La red Unet 3D implementada presenta un coeficiente Dice de 0.809, el cuál es superior al de otros estudios de segmentación automática de grasa parafaríngea que presentan un coeficiente Dice de 0.779 y 0.580.

Con respecto a los trabajos futuros, se sugiere la implementación de una interfaz gráfica de usuario, que permita representaciones visuales de las funcionalidades desarrolladas. Adicionalmente se recomienda incorporar un módulo que permita la visualización tanto de las imágenes médicas como de su segmentación.

Para encontrar relaciones entre la presencia de grasa parafaríngea, la existencia y severidad de la AOS, se recomienda ampliar el tamaño de la muestra y generar datos estadísticos que indiquen dicha correlación. Al aumentar el tamaño de la muestra se podría mejorar la exactitud de la red. Se espera que esta muestra ampliada se pueda estratificar por edad, sexo y severidad de la AOS para determinar si hay diferencias en el nivel de grasa parafaríngea en cada uno los grupos, lo que permitiría tener una aplicación clínica de la investigación.

Para permitir la adaptación del prototipo a imágenes de resonancia magnética con diferente tamaño, espaciado e intensidad a las utilizadas, se sugiere el desarrollo de una funcionalidad que permita personalizar el preprocesamiento de las imágenes.

También se propone usar varios métodos de segmentación con el mismo dataset para poder hacer comparaciones con significancia estadística. También se propone cambiar las funciones de activación haciendo uso de una función para el camino de contracción/codificación y otra función distinta para el camino de expansión/decodificación de la red unet 3D.

Finalmente se propone adaptar el flujo de segmentación propuesto para realizar la segmentación de grasa en la lengua que permita encontrar relaciones entre el volumen de esta estructura y la existencia de la AOS, utilizando un dataset del cual se conozcan datos de los sujetos de estudio como edad y sexo.

10 Referencias

- Aggarwal, K., & Panwar Jogi, S. (2020). Automatic Liver and Subcutaneous Fat Segmentation from MRI-PDF Images. *2020 International Conference on Computational Performance Evaluation (ComPE)*, 376–378. <https://doi.org/10.1109/ComPE49325.2020.9200196>
- Albu, A. B., Beugeling, T., & Laurendeau, D. (2008). A Morphology-Based Approach for Interslice Interpolation of Anatomical Slices From Volumetric Images. *IEEE Transactions on Biomedical Engineering*, 55(8), 2022–2038. <https://doi.org/10.1109/TBME.2008.921158>
- Alsufyani, N. A., Flores-Mir, C., & Major, P. W. (2012). Three-dimensional segmentation of the upper airway using cone beam CT: a systematic review. *Dento Maxillo Facial Radiology*, 41(4), 276–284. <https://doi.org/10.1259/dmfr/79433138>
- Amer, R., Nassar, J., Bendahan, D., Greenspan, H., & Ben-Eliezer, N. (2019). Automatic Segmentation of Muscle Tissue and Inter-muscular Fat in Thigh and Calf MRI Images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11765 LNCS, 219–227. https://doi.org/10.1007/978-3-030-32245-8_25
- Boehm, B. (1986). A Spiral Model of Software Development and Enhancement. *SIGSOFT Softw. Eng. Notes*, 11(4), 14–24. <https://doi.org/10.1145/12944.12948>
- Bommineni, V. L., Erus, G., Doshi, J., Singh, A., Keenan, B. T., Schwab, R. J., ... Davatzikos, C. (2022). Automatic Segmentation and Quantification of Upper Airway Anatomic Risk Factors for Obstructive Sleep Apnea on Unprocessed Magnetic Resonance Images. *Academic Radiology*. <https://doi.org/10.1016/j.acra.2022.04.023>
- Bui, N. L., Ong, S. H., & Foong, K. W. C. (2015). Automatic segmentation of the nasal cavity and paranasal sinuses from cone-beam CT images. *International Journal of Computer Assisted Radiology and Surgery*, 10(8), 1269–1277. <https://doi.org/10.1007/s11548-014-1134-5>
- Carter, R., & Watenpaugh, D. E. (2008). Obesity and obstructive sleep apnea: Or is it OSA and obesity? *Pathophysiology*, 15(2), 71–77. <https://doi.org/https://doi.org/10.1016/j.pathophys.2008.04.009>
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 9901 LNCS* (pp. 424–432). https://doi.org/10.1007/978-3-319-46723-8_49
- Contributors, P. (n.d.). PyTorch documentation — PyTorch 1.13 documentation. Retrieved January 2, 2023, from <https://pytorch.org/docs/stable/index.html>
- Dabiri, S., Popuri, K., Ma, C., Chow, V., Feliciano, E. M. C., Caan, B. J., ... Beg, M. F. (2020). Deep learning method for localization and segmentation of abdominal CT.

- Computerized Medical Imaging and Graphics*, 85(February), 101776.
<https://doi.org/10.1016/j.compmedimag.2020.101776>
- Dhaliwal, S. S., Hesabgar, S. M., Haddad, S. M. H., Ladak, H., Samani, A., & Rotenberg, B. W. (2018). Constructing a patient-specific computer model of the upper airway in sleep apnea patients. *Laryngoscope*, 128(1), 277–282.
<https://doi.org/10.1002/lary.26834>
- El, H., & Palomo, J. M. (2010). Measuring the airway in 3 dimensions: A reliability and accuracy study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(4), S50.e1-S50.e9. <https://doi.org/10.1016/j.ajodo.2009.11.010>
- García-García, J. A., Reding-Bernal, A., & López-Alvarenga, J. C. (2013). Cálculo del tamaño de la muestra en investigación en educación médica. *Investigación En Educación Médica*, 2(8), 217–224. [https://doi.org/10.1016/S2007-5057\(13\)72715-7](https://doi.org/10.1016/S2007-5057(13)72715-7)
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gustinel, M. (2022). *A survey on recently proposed activation functions for Deep Learning*. 1–7. Retrieved from <http://arxiv.org/abs/2204.02921>
- He, X., Guo, B. J., Lei, Y., Wang, T., Curran, W. J., Liu, T., ... Yang, X. (2021). Automatic quantification of myocardium and pericardial fat from coronary computed tomography angiography: a multicenter study. *European Radiology*, 31(6), 3826–3836. <https://doi.org/10.1007/s00330-020-07482-5>
- Hevner, A., R, A., March, S., T, S., Park, Park, J., ... Sudha. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, 28, 75.
- Irmakci, I., Hussein, S., Savran, A., Kalyani, R. R., Reiter, D., Chia, C. W., ... Bagci, U. (2019). A Novel Extension to Fuzzy Connectivity for Body Composition Analysis: Applications in Thigh, Brain, and Whole Body Tissue Segmentation. *IEEE Transactions on Biomedical Engineering*, 66(4), 1069–1081.
<https://doi.org/10.1109/TBME.2018.2866764>
- ISO/IEC 25010. (2011). *{ISO}/{IEC} 25010:2011, Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*.
- Ivanovska, T., Daboul, A., Kalentev, O., Hosten, N., Volzke, H., Biffar, R., & Worgotter, F. (2021). On Efficient Segmentation of Parapharyngeal Fat Pads From Population-based MRIs: An Obstructive Sleep Apnea Application. *2021 6th International Conference on Communication, Image and Signal Processing (CCISP)*, 178–182. IEEE. <https://doi.org/10.1109/CCISP52774.2021.9639319>
- Ivanovska, T., Dober, J., Laqua, R., Hegenscheid, K., & Völzke, H. (2013). Pharynx Segmentation from MRI Data for Analysis of Sleep Related Disorders. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 8033 LNCS* (pp. 20–29). https://doi.org/10.1007/978-3-642-41914-0_3
- Kabaliuk, N., Nejati, A., Loch, C., Schwass, D., Cater, J. E., & Jermy, M. C. (2017).

- Strategies for Segmenting the Upper Airway in Cone-Beam Computed Tomography (CBCT) Data. *Open Journal of Medical Imaging*, 07(04), 196–219.
<https://doi.org/10.4236/ojmi.2017.74019>
- Kato, M., Adachi, T., Koshino, Y., & Somers, V. K. (2009). Obstructive sleep apnea and cardiovascular disease. *Circulation Journal : Official Journal of the Japanese Circulation Society*, 73(8), 1363–1370. <https://doi.org/10.1253/circj.cj-09-0364>
- Kazemi, A., Keshtkar, A., Rashidi, S., Aslanabadi, N., Khodadad, B., & Esmaeili, M. (2019). Segmentation of Cardiac Epicardial and Pericardial Fats by Using Gabor Filter Bank Based GLCM. *2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)*, 177–182.
<https://doi.org/10.1109/ICBME49163.2019.9030380>
- Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, & Samir Chatterjee. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77.
- Kim, A. M., Keenan, B. T., Jackson, N., Chan, E. L., Staley, B., Poptani, H., ... Schwab, R. J. (2014). Tongue fat and its relationship to obstructive sleep apnea. *Sleep*, 37(10), 1639-1648D. <https://doi.org/10.5665/sleep.4072>
- Li, Y., Lin, N., Ye, J., Chang, Q., Han, D., & Sperry, A. (2012). Upper Airway Fat Tissue Distribution in Subjects With Obstructive Sleep Apnea and Its Effect on Retropalatal Mechanical Loads. *Respiratory Care*, 57(7), 1098–1105.
<https://doi.org/10.4187/respcare.00929>
- MacLean, M. T., Jehangir, Q., Vujkovic, M., Ko, Y. A., Litt, H., Borthakur, A., ... Witschey, W. R. (2021). Quantification of abdominal fat from computed tomography using deep learning and its association with electronic health records in an academic biobank. *Journal of the American Medical Informatics Association*, 28(6), 1178–1187.
<https://doi.org/10.1093/jamia/ocaa342>
- Maddalo, M., Zorza, I., Zubani, S., Nocivelli, G., Calandra, G., Soldini, P., ... Maroldi, R. (2017). Validation of a free software for unsupervised assessment of abdominal fat in MRI. *Physica Medica*, 37, 24–31.
<https://doi.org/https://doi.org/10.1016/j.ejmp.2017.04.002>
- Masoudi, S., Anwar, S. M., Harmon, S. A., Choyke, P. L., Turkbey, B., & Bagci, U. (2020). Adipose Tissue Segmentation in Unlabeled Abdomen MRI using Cross Modality Domain Adaptation. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 1624–1628.
<https://doi.org/10.1109/EMBC44109.2020.9176009>
- Morsy, N. E., Farrag, N. S., Zaki, N. F. W., Badawy, A. Y., Abdelhafez, S. A., El-Gilany, A.-H., ... BaHammam, A. S. (2019). Obstructive sleep apnea: personal, societal, public health, and legal implications. *Reviews on Environmental Health*, 34(2), 153–169. <https://doi.org/https://doi.org/10.1515/reveh-2018-0068>
- Motamedi, K. K., McClary, A. C., & Amedee, R. G. (2009). Obstructive sleep apnea: a growing problem. *The Ochsner Journal*, 9(3), 149–153.

- MuhammadLaiq Ur Rahman, S., Chitiboi, T., Ivanovska, T., Molchanov, V., Völzke, H., Hahn, H. K., & Linsen, L. (2015). Automatic pharynx segmentation from MRI data for obstructive sleep apnea analysis. *VISAPP 2015 - 10th International Conference on Computer Vision Theory and Applications; VISIGRAPP, Proceedings, 1*, 599–608. <https://doi.org/10.5220/0005315905990608>
- Neelapu, B. C., Kharbanda, O. P., Sardana, V., Gupta, A., Vasamsetti, S., Balachandran, R., ... Sardana, H. K. (2017). A pilot study for segmentation of pharyngeal and sino-nasal airway subregions by automatic contour initialization. *International Journal of Computer Assisted Radiology and Surgery*, *12*(11), 1877–1893. <https://doi.org/10.1007/s11548-017-1650-1>
- NumPy Developers. (n.d.). keleshev/schema: Schema validation just got Pythonic. Retrieved January 10, 2023, from <https://github.com/keleshev/schema>
- NumPy documentation — NumPy v1.24 Manual. (n.d.). Retrieved January 10, 2023, from <https://numpy.org/doc/stable/>
- Nyul, L. G., Udupa, J. K., & Zhang, X. (2000). New variants of a method of MRI scale standardization. *IEEE Transactions on Medical Imaging*, *19*(2), 143–150. <https://doi.org/10.1109/42.836373>
- Pallets. (n.d.). Welcome to Click — Click Documentation (8.1.x). Retrieved January 10, 2023, from <https://click.palletsprojects.com/en/8.1.x/>
- Pérez-García, F. (2020). *fepegar/unet: PyTorch implementation of 2D and 3D U-Net*. <https://doi.org/10.5281/ZENODO.3697931>
- Pérez-García, F., Sparks, R., & Ourselin, S. (2021). TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine*, *208*, 106236. <https://doi.org/https://doi.org/10.1016/j.cmpb.2021.106236>
- PyYAML · PyPI. (n.d.). Retrieved January 10, 2023, from <https://pypi.org/project/PyYAML/>
- Richards, M. (2015). *Software Architecture Patterns*. Retrieved from <http://oreilly.com/safari>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access*, *9*, 16591–16603. <https://doi.org/10.1109/ACCESS.2021.3053408>
- Ruiz, A. J., Rondon Sepúlveda, M. A., Franco, O. H., Cepeda, M., Hidalgo Martinez, P., Amado Garzón, S. B., ... Otero Mendoza, L. (2016). The associations between sleep disorders and anthropometric measures in adults from three Colombian cities at different altitudes. *Maturitas*, *94*, 1–10. <https://doi.org/10.1016/j.maturitas.2016.08.013>
- Shahid, M. L. U. R., Chitiboi, T., Ivanovska, T., Molchanov, V., Völzke, H., & Linsen, L. (2017). Automatic MRI segmentation of para-pharyngeal fat pads using interactive visual feature space analysis for classification. *BMC Medical Imaging*, *17*(1).

<https://doi.org/10.1186/s12880-017-0179-7>

- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, 15, 29. <https://doi.org/10.1186/s12880-015-0068-x>
- tqdm · PyPI. (n.d.). Retrieved January 10, 2023, from <https://pypi.org/project/tqdm/>
- Venkatesan, R., & Li, B. (2017). *Convolutional Neural Networks in Visual Computing* (Vol. 4). Boca Raton ; London : Taylor & Francis, CRC Press, 2017.: CRC Press. <https://doi.org/10.4324/9781315154282>
- Wang, S. H., Keenan, B. T., Wiemken, A., Zang, Y., Staley, B., Sarwer, D. B., ... Schwab, R. J. (2020). Effect of weight loss on upper airway anatomy and the apnea–hypopnea index the importance of tongue fat. *American Journal of Respiratory and Critical Care Medicine*, 201(6), 718–727. <https://doi.org/10.1164/rccm.201903-0692OC>
- Wojcik, R., Bachmann, F., Bass, L., Clements, P., Merson, P., Nord, R., & Wood, B. (2006). *Attribute-Driven Design (ADD), Version 2.0 Software Architecture Technology Initiative*. Retrieved from <http://www.sei.cmu.edu/publications/pubweb.html>
- Zlokolica, V., Krstanović, L., Velicki, L., Popović, B., Janev, M., Obradović, R., ... Babin, D. (2017). Semiautomatic Epicardial Fat Segmentation Based on Fuzzy c-Means Clustering and Geometric Ellipse Fitting. *Journal of Healthcare Engineering*, 2017. <https://doi.org/10.1155/2017/5817970>
- Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M. C., Kaus, M. R., Haker, S. J., ... Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index. *Academic Radiology*, 11(2), 178–189. [https://doi.org/10.1016/s1076-6332\(03\)00671-8](https://doi.org/10.1016/s1076-6332(03)00671-8)
- Zukić, D., Vicory, J., McCormick, M., Wisse, L., Gerig, G., Yushkevich, P., & Aylward, S. (2016). ND morphological contour interpolation. *The Insight Journal*. <https://doi.org/10.54294/achtrg>

11 Anexos

11.1 Anexo 1. Aval de comité de ética

El presente proyecto hace parte del proyecto titulado *Generación de biomarcadores para el estudio de cambios neurocognoscitivos asociado a fenómenos cerebro-vasculares en individuos con apnea obstructiva del sueño*. Este proyecto utiliza imágenes de resonancia magnética que fueron anonimizadas en el proyecto antecesor, por tal motivo no se hace manejo de información personal o sensible de los sujetos del estudio.

11.2 Anexo 2. Documento de Especificación de Requerimientos de Software

En este documento se presentan los requerimientos funcionales generales del sistema de segmentación de grasa parafaríngea y requerimientos no funcionales. Parte de estos requerimientos son utilizados para la elaboración del prototipo presentado en este proyecto.

11.3 Anexo 3. Documento de Arquitectura de Software

En este documento se presenta el diseño arquitectural del prototipo descrito a partir de las vistas de despliegue, lógica e implementación.

11.4 Anexo 4. Repositorio del prototipo

El enlace adjunto a esta sección contiene el código fuente del prototipo desarrollado durante la elaboración del presente proyecto.

https://github.com/OscarDuarte7/segmentation_prototype

11.5 Anexo 5. Manual de usuario

En este documento se describe una guía para el uso del prototipo desarrollado.