

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
SANTIAGO - CHILE



**“Framework de segmentación eficiente de imágenes
de contraste de fase 2D obtenidos por resonancia
magnética”**

JAVIER ALEXANDER VARGAS MONTESINOS

**MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA**

Profesor Guía: Julio Sotelo
Profesor Correferente: Joaquín Mura

Noviembre - 2024

DEDICATORIA

Se lo dedico a mi Padre, cuyo apoyo, motivación y esfuerzo, me permitió llegar a este punto. También a mi Madre que, a pesar de la distancia, siempre se preocupó por mi bien estar. Y a mi Abuela y Abuelo por su apoyo y cariño incondicional.

AGRADECIMIENTOS

Agradezco a todas aquellas personas que me han acompañado a lo largo de mi carrera universitaria: profesores, amigos y familiares, quienes me han motivado e impulsado a continuar este desafío hasta el final.

RESUMEN

Resumen— En este trabajo se presenta la implementación de un *framework* de segmentación para secuencias de imágenes de resonancia magnética de contraste de fase 2D (2D PC-MRI), cuyo objetivo es mejorar la precisión en la segmentación de los principales vasos sanguíneos. El modelo propuesto utiliza una arquitectura 3D Unet, la cual está diseñada para la segmentación de volúmenes. Se realiza una comparación del desempeño del modelo propuesto con métodos de segmentación basados en contornos activos y segmentación manual, mostrando una mejora en la calidad y estabilidad de la segmentación, destacando la relevancia de aplicar técnicas de redes neuronales en el análisis de imágenes médicas para mejorar la calidad y fiabilidad de las mediciones.

Palabras Clave—PC-MRI, segmentación, Unet, Redes Neuronales Convolucionales, flujo cardiovascular.

ABSTRACT

Abstract— In this work we present the implementation of a segmentation framework for 2D phase-contrast magnetic resonance imaging (PC-MRI), aimed at improving the accuracy in segmenting the major blood vessels. The proposed model makes use of a 3D Unet architecture, which is designed for volume segmentation. A performance comparison of the proposed model is carried out against others segmentation methods like active contours and manual segmentation, demonstrating an improvement in the quality and stability of the segmentation. This highlights the relevance of applying neural network techniques in medical image analysis to enhance the quality and reliability of measurements.

Keywords—PC-MRI, segmentation, Unet, Convolutional Neural Networks, cardiovascular flow.

GLOSARIO

MRI: Magnetic Resonance Image

PC-MRI: Phase-Contrast Magnetic Resonance Image

CNN: Convolutional Neural Network

RNN: Recurrent Neural Network

DICOM: Digital Imaging and Communication in Medicine

ROI: Region of Interest

ANN: Artificial neural network

DNN: Deep neural Network

TEP: Tomografía por Emisión de Positrones

CT: Computed Tomography

IA: Inteligencia Artificial

PD: Proton Density

RF: Radiofrecuencia

TR: Tiempo de repetición

TE: Tiempo de eco

NMV: Net Magnetization Vector

INDICE DE CONTENIDOS

RESUMEN.....	4
ABSTRACT	4
INDICE DE TABLAS.....	13
INTRODUCCIÓN	14
CAPÍTULO 1: DEFINICIÓN DEL PROBLEMA	15
1.1 CONTEXTO.....	15
1.2 SEGMENTACIÓN	15
1.3 PROBLEMA	16
1.4 OBJETIVO.....	17
1.4.1 OBJETIVO GENERAL	17
1.4.2 OBJETIVOS ESPECIFICOS	17
CAPÍTULO 2: MARCO CONCEPTUAL	19
2.1 CICLO CARDÍACO	19
2.2 IMÁGENES DE RESONANCIA MAGNÉTICA	19
2.2.1 IMÁGENES DE RESONANCIA MAGNÉTICA POR CONTRASTE DE FASE.....	24
2.2.2 ELECTROCARDIOGRAMA.....	26
2.2.3 CINE MRI	26
2.3 SEGMENTACIÓN DE IMÁGENES MEDICAS	27
2.3.1 SEGMENTACIÓN POR CONTORNO ACTIVO	28
2.4 FORMATO DICOM	29
2.5 REDES NEURONALES	29
2.5.1 NEURONA ARTIFICIAL	30
2.5.2 FUNCIONES DE ACTIVACIÓN.....	31
2.6 REDES NEURONALES CONVOLUCIONALES.....	32
2.6.1 ARQUITECTURA DE UNA CNN.....	33

2.7 APRENDIZAJE DE UNA RED NEURONAL ARTIFICIAL	35
2.7.1 FUNCIÓN DE PERDIDA (LOSS FUNCTION)	35
CAPÍTULO 3: PROPUESTA DE SOLUCION	37
3.1 CONJUNTO DE DATOS	37
3.2 METODO PROPUESTO	38
3.2.1 UNET	40
3.2.2 VOLUMENES DE IMÁGENES.....	41
3.2.3 3D UNET	43
3.2.4 IMPLEMENTACION DE MODELO 3D UNET	43
3.3 ENTRENAMIENTO DEL MODELO	45
3.4 PROCESAMIENTO DE LAS IMÁGENES DE ENTRADA.....	46
3.4.1 EXTRACCIÓN SECUENCIAL	46
3.4.2 EXTRACCIÓN RECURSIVA	48
3.5 REGIÓN DE INTERÉS	49
3.6 MÁSCARA DE SALIDA.....	51
3.7 PROCESAMIENTO DE LA SEGMENTACIÓN	52
3.8 IMAGEN DE FASE.....	53
3.9 FRAMEWORK DE SEGMENTACIÓN.....	54
CAPÍTULO 4: VALIDACIÓN DE LA SOLUCIÓN	56
4.1 METODO EXPERIMENTAL.....	56
4.2 CASOS ESPECÍFICOS.....	57
4.2.1 PRIMER CASO	57
4.2.2 SEGUNDO CASO	61
4.3 PRUEBAS DE USO GENERAL	65
4.4 TIEMPOS DE PREDICCIÓN.....	67
CAPÍTULO 5: CONCLUSIONES	70

5.1 TRABAJOS FUTUROS.....	73
REFERENCIAS BIBLIOGRÁFICAS.....	75

INDICE DE FIGURAS

Figura 1: Imagen de resonancia magnética siendo segmentada por un algoritmo semiautomático. Fuente: Elaboración propia.	16
Figura 2: Segmentación de la aorta descendente durante periodos de alto y bajo contraste. Fuente: Elaboración propia.	17
Figura 3: Gráficos de área y flujo calculados a partir de una segmentación manual y por un método de contorno activo. Fuente: Elaboración propia.	17
Figura 4: Alineamiento de los núcleos de hidrogeno. Fuente: (Westbrook et al., 2011)	19
Figura 5: Vector de magnetización neta (NMV). Fuente: (Westbrook et al., 2011)	20
Figura 6: Precesión en los núcleos de hidrogeno. Fuente: (Westbrook et al., 2011)	20
Figura 7: Transferencia de energía provocado por el pulso de radiofrecuencia. Fuente: (Westbrook et al., 2011).....	21
Figura 8: Efectos de la resonancia en los núcleos. A) Fuente: (Westbrook et al., 2011)	21
Figura 9: Secuencia básica de pulsos de RF. TR es el tiempo de repetición de los pulsos RF, y TE es el tiempo de eco, que corresponde al tiempo entre el pulso RF y la captura de la señal. Fuente: (Westbrook et al., 2011).....	22
Figura 10: Relajación longitudinal T1. Fuente: (Westbrook et al., 2011)	22
Figura 11: Relajación Transversal T2. Fuente: (Westbrook et al., 2011)	22
Figura 12: Gradiente del campo magnético. Fuente: (Westbrook et al., 2011)	23
Figura 13: Diferencias de T1 y T2 entre las grasas y los líquidos. Fuente: (Westbrook et al., 2011)	23
Figura 14: Imagen axial del cerebro humano ponderada en T1 (izquierda), T2 (Centro) y PD (Derecha). Fuente: (Westbrook et al., 2011)	24
Figura 15: Efecto de los gradientes bipolares. El desplazamiento de fase es cero para núcleos estacionarios, pero positiva para núcleos que se mueven en la dirección del gradiente. Fuente: (mriquestions, 2024).....	25
Figura 16: Ejemplo de 2D PC-MRI, Se observa la imagen de fase (izquierda) y la imagen de magnitud (derecha). Fuente: (Wymer et al., 2020)	26
Figura 17: Ejemplo de cine PC-MRI con 9 instantes de tiempo, donde se observan las imágenes de magnitud y fase. Fuente: Elaboración Propia.	27
Figura 18: Segmentación de una imagen MRI con diferentes métodos. En rojo se marca la segmentación manual, mientras que los otros colores corresponden a diferentes métodos de segmentación automatizados. Fuente: (Mahapatra, 2013).....	28
Figura 19: Fotogramas seleccionados desde un video de 2 segundos. El snake se ajusta automáticamente a los movimientos de los labios. Fuente: (Kass et al., 1988)	29
Figura 20: Arquitectura de una red neuronal. Fuente: (Wang S.-C. , 2003)	30

Figura 21: Comparación entre una neurona biología y una neurona artificial. Fuente: (Krenker et al., 2011) 30

Figura 22: Arquitectura de una CNN. Se aprecia la capa convolucional con función de activación ReLu, la capa *pooling*, y dos capas completamente conectadas. Fuente: (O'Shea & Nash, 2015) 33

Figura 23: Representación visual de una capa convolucional. Se observa los datos de entrada los cuales son reducidos en dimensión por una capa de *pooling* para luego ser procesado por el *kernel*. Fuente: (O'Shea & Nash, 2015)..... 34

Figura 24: Mapas de activación aprendidos en cada filtro durante el entrenamiento de una CNN. Fuente: (O'Shea & Nash, 2015)..... 34

Figura 25: Muestra de imágenes para los diferentes tipos de cortes en el conjunto de datos. Fuente: Elaboración Propia. 38

Figura 26: Visualización de una segmentación por contorno activo de la aorta ascendente durante la fase sistólica y diastólica. Para ambas fases se observa una muestra de las imágenes de magnitud. Fuente: Elaboración Propia. 39

Figura 27: Arquitectura del modelo Unet. Fuente: (Ronneberger et al., 2015) 41

Figura 28: Conversión de una secuencia de **L** imágenes de resonancia magnética de resolución **WxH** pixeles, en un volumen de dimensión **LxWxH**. Fuente: Elaboración propia. 42

Figura 29: Segmentación de la aorta ascendente para un volumen MRI con dimensiones **LxWxH**. Se presenta la vista frontal del volumen junto con un corte sagital en el índice 124. Fuente: Elaboración Propia. 43

Figura 30: Arquitectura del modelo 3D Unet propuesto. Fuente: Elaboración propia. 44

Figura 31: Creación de volúmenes a través del método de extracción secuencial. Fuente: Elaboración propia. 47

Figura 32: Creación de volúmenes a través del método de extracción recursiva. Fuente: Elaboración propia. 49

Figura 33: Extracción de las regiones de interés para el entrenamiento de la 3D Unet. Fuente: Elaboración Propia. 50

Figura 34: Entrenamiento de la 3D Unet, el input corresponde a los volúmenes creados y la salida una máscara segmentando la estructura al interior de la región de interés. En el ejemplo, la estructura es la aorta ascendente. Fuente: Elaboración propia..... 51

Figura 35: Extracción y unificación de las máscaras desde los volúmenes segmentados. Fuente: Elaboración propia. 52

Figura 36: Recuperación de la resolución original de la ROI para luego ser posicionada en la secuencia original PC-MRI. Fuente: Elaboración propia. 52

Figura 37: Combinación de las imágenes de Magnitud y Fase en una sola imagen con 2 canales. Fuente: Elaboración propia. 53

Figura 38: Los 5 procedimientos para el framework de segmentación. Fuente: Elaboración propia. 54

Figura 39: Framework de segmentación de secuencias 2D PC-MRI utilizando el modelo 3D Unet. Fuente: Elaboración propia. 55

Figura 40: (Caso 1) Muestra de imágenes de Magnitud, donde se observan las aortas ascendente y descendente a lo largo del ciclo cardiaco (las primeras 3 corresponden a la fase sístole y las otras a la diástole). Fuente: Elaboración propia.	57
Figura 41: (Caso 1) Segmentación de la aorta ascendente utilizando los métodos de segmentación manual (Verde), contorno activo (Azul) y el modelo propuesto 3D Unet Recursivo (Púrpura) y Secuencial (Rojo). Fuente: Elaboración propia.	58
Figura 42: (Caso 1) Gráfico comparativo de las áreas utilizando la segmentación manual (Verde), el método de contorno activo (Azul), y la propuesta 3D Unet con los enfoques recursivos (Púrpura) y secuencial (Rojo). Fuente: Elaboración propia.	59
Figura 43: (Caso 1) Gráfico comparativo de los flujos calculados a partir de la segmentación obtenida por medio de la segmentación manual (Verde), el método de contorno activo (Azul), y la propuesta 3D UNET con los enfoques recursivos (Púrpura) y secuencial (Rojo). Fuente: Elaboración propia.	60
Figura 44: (Caso 2) Muestra de imágenes de Magnitud, donde se observan las aortas ascendente y descendente en las fases de sístole y diástole (las primeras 3 corresponden a la sístole y las otras a la diástole). Fuente: Elaboración propia.	61
Figura 45: (Caso 2) Segmentación de la aorta ascendente utilizando los métodos de segmentación manual (Verde), contorno activo (Azul) y el modelo propuesto 3D Unet Recursivo (Púrpura) y Secuencial (Rojo). Fuente: Elaboración propia.	62
Figura 46: (Caso 2) Gráfico comparativo de las áreas utilizando la segmentación manual (Verde), el método de contorno activo (Azul), y la propuesta 3D Unet con los enfoques recursivos (Púrpura) y secuencial (Rojo). Fuente: Elaboración propia.	63
Figura 47: (Caso 2) Gráfico comparativo de los flujos calculados a partir de la segmentación obtenida por medio de la segmentación manual (Verde), el método de contorno activo (Azul), y la propuesta 3D UNET con los enfoques recursivos (Púrpura) y secuencial (Rojo). Fuente: Elaboración propia.	64
Figura 48: Tiempos de ejecución del modelo 3D Unet, utilizando CPU y GPU. Fuente: Elaboración propia. ..	68
Figura 49: Gráficos de Área y Flujo para el estudio 1. Fuente: Elaboración propia.....	77
Figura 50: Gráficos de Área y Flujo para el estudio 2. Fuente: Elaboración propia.....	77
Figura 51: Gráficos de Área y Flujo para el estudio 3. Fuente: Elaboración propia.....	77
Figura 52: Gráficos de Área y Flujo para el estudio 4. Fuente: Elaboración propia.....	78
Figura 53: Gráficos de Área y Flujo para el estudio 6. Fuente: Elaboración propia.....	78
Figura 54: Gráficos de Área y Flujo para el estudio 7. Fuente: Elaboración propia.....	78
Figura 55: Gráficos de Área y Flujo para el estudio 8. Fuente: Elaboración propia.....	78
Figura 56: Gráficos de Área y Flujo para el estudio 9. Fuente: Elaboración propia.....	79
Figura 57: Gráficos de Área y Flujo para el estudio 10. Fuente: Elaboración propia.....	79
Figura 58: Gráficos de Área y Flujo para el estudio 12. Fuente: Elaboración propia.....	79
Figura 59: Gráficos de Área y Flujo para el estudio 13. Fuente: Elaboración propia.....	79

Figura 60: Gráficos de Área y Flujo para el estudio 14. Fuente: Elaboración propia..... 80

Figura 61: Gráficos de Área y Flujo para el estudio 15. Fuente: Elaboración propia..... 80

Figura 62: Gráficos de Área y Flujo para el estudio 16. Fuente: Elaboración propia..... 80

Figura 63: Gráficos de Área y Flujo para el estudio 21. Fuente: Elaboración propia..... 80

Figura 64: Gráficos de Área y Flujo para el estudio 22. Fuente: Elaboración propia..... 81

INDICE DE TABLAS

Tabla 1: Tipos de corte en las secuencias PC-MRI del conjunto de datos. Fuente: Elaboración Propia.....	37
Tabla 2: Cambio dimensional a lo largo del modelo 3D Unet. Fuente: Elaboración propia.	44
Tabla 3: Características técnicas del equipo que ejecutará las pruebas. Fuente: Elaboración propia.....	56
Tabla 4: (Caso 1) Cálculo del coeficiente de Dice para las segmentaciones obtenidas por los métodos en estudio, siendo comparadas con la segmentación manual. Fuente: Elaboración propia.....	60
Tabla 5: (Caso 2) Cálculo del coeficiente de Dice para las segmentaciones obtenidas por los métodos en estudio, siendo comparadas con la segmentación manual. Fuente: Elaboración propia.....	63
Tabla 6: Calculo del coeficiente de Dice para las segmentaciones obtenidas por los métodos estudiados, para un conjunto de estudios de secuencias PC-MRI. Fuente: Elaboración propia.....	65

INTRODUCCIÓN

La resonancia magnética (MR) es una herramienta avanzada de imagen médica que permite obtener visualizaciones detalladas del interior del cuerpo humano sin recurrir a métodos invasivos ni a la radiación ionizante. En el campo del análisis cardiovascular, las imágenes de resonancia magnética con contraste de fase 2D (2D PC-MRI) han demostrado ser especialmente útiles para medir el flujo sanguíneo en el corazón y los grandes vasos. Sin embargo, para realizar un análisis efectivo de estas imágenes, es esencial un proceso de segmentación preciso que permita delimitar las áreas de interés.

Una segmentación de alta calidad es crucial para comprender mejor la anatomía interna y realizar análisis tanto cualitativos como cuantitativos. No obstante, los métodos automáticos y semiautomáticos de segmentación pueden verse comprometidos por problemas relacionados con la calidad de las imágenes. Factores como la baja resolución espacial y temporal, la presencia de artefactos de flujo pulsátil y la pérdida de señal durante la diástole cardíaca pueden afectar negativamente la precisión del análisis. Estas inconsistencias introducen incertidumbre en las mediciones de área y flujo cardíaco.

En este trabajo se propone la implementación de un *framework* de segmentación para secuencias 2D PC-MRI, basado en un modelo 3D UNet, el cual corresponde a una arquitectura de red neuronal convolucional (CNN) especialmente diseñada para la segmentación de volúmenes. Este enfoque busca mitigar la incertidumbre en la segmentación, particularmente en los momentos de la secuencia en los que la estructura en análisis pierde nitidez o resolución, optimizando así la precisión en el análisis del flujo cardiovascular.

El documento se organiza en cinco capítulos. El capítulo 1 proporciona una introducción al problema de la segmentación y plantea los objetivos del trabajo. En el capítulo 2, se presentan los conceptos clave relacionados con las imágenes 2D PC-MRI, los formatos de archivo utilizados y los conceptos de las redes neuronales, entre otros elementos importantes para entender la propuesta. En el capítulo 3 se detalla la metodología propuesta, incluyendo el conjunto de datos utilizado y la arquitectura UNet clásica, junto a su variante para volúmenes 3D. Además, se explican los distintos procesos involucrados en la implementación del *framework* de segmentación. En el capítulo 4, se pone a prueba el método propuesto, comparando su rendimiento frente al método de segmentación basado en contornos activos y la segmentación manual. Finalmente, en el capítulo 5 se presentan las conclusiones del trabajo y se sugieren guías de investigación para trabajos futuros.

CAPÍTULO 1: DEFINICIÓN DEL PROBLEMA

1.1 CONTEXTO

Las enfermedades cardiovasculares constituyen la principal causa de muerte a nivel global, se estima que cada año aproximadamente 17,9 millones de personas fallecen debido a enfermedades cardiovasculares. Este grupo de enfermedades afecta al corazón y a los vasos sanguíneos e incluyen condiciones como las cardiopatías coronarias, enfermedades cerebrovasculares y cardiopatías reumáticas. Entre las defunciones relacionadas con estas enfermedades, 4 de cada 5 se atribuye a cardiopatías coronarias y accidentes cerebrovasculares, y una tercera parte de estas defunciones se produce en individuos menores de 70 años. (World Health Organization, 2019)

Una de las herramientas utilizadas para el diagnóstico y tratamiento de estas enfermedades es por medio de las imágenes por resonancia magnética (MRI). Esta es una técnica de imagen no invasiva que utiliza campos magnéticos y ondas de radio para generar imágenes detalladas del interior del cuerpo humano. En el ámbito del análisis cardiovascular, las imágenes de resonancia magnética con contraste de fase 2D (2D PC-MRI) son una herramienta muy eficaz para la medición del flujo circulatorio en el corazón y de los grandes vasos. (Mayo Clinic, 2023)

1.2 SEGMENTACIÓN

Para llevar a cabo el análisis de las MRI, resulta necesario realizar un proceso de segmentación que delimite regiones de interés para posteriormente evaluar algún parámetro sobre esa región. Una segmentación precisa proporciona una buena comprensión de la anatomía y de las estructuras al interior del cuerpo humano lo que habilita un análisis cualitativo y cuantitativo de mayor calidad.

La segmentación de las imágenes de resonancia magnética (MRI) se puede realizar de tres formas. La forma manual, que implica que un experto identifique y marque manualmente las regiones de interés en la imagen. La forma semiautomática, que combina la intervención humana junto con un proceso automático, por lo general, la persona marca la posición de la estructura que se desea segmentar y un algoritmo automático comienza a segmentar dicha región. Y la forma totalmente automática, la cual no requiere de intervención humana, por lo que el algoritmo identifica automáticamente la estructura y la segmenta. La Figura 1 muestra un ejemplo de segmentación semiautomática, en la cual el especialista marca manualmente el área de la estructura y el algoritmo se adapta al contorno de la imagen.

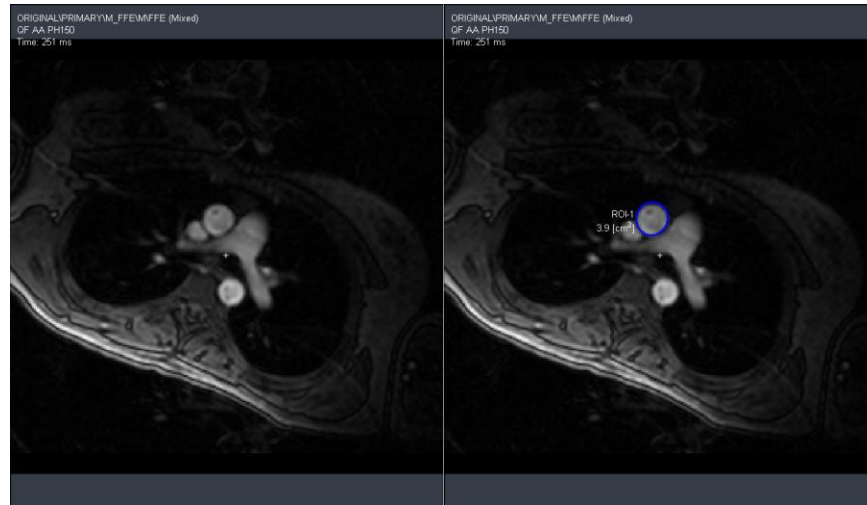


Figura 1: Imagen de resonancia magnética siendo segmentada por un algoritmo semiautomático. Fuente: Elaboración propia.

1.3 PROBLEMA

Un problema que surge de los métodos de segmentación automáticos o semiautomáticos (como los métodos de segmentación basados en contorno activo), es que son sensibles al contraste y ruido que posea la imagen, por lo que la segmentación puede fallar si el algoritmo de segmentación automática no está preparado para predecir la estructura en momentos de incertidumbre.

En las imágenes por resonancia magnética (MRI) son variados los problemas que pueden afectar a la calidad de imagen. Entre los principales se encuentran: la baja resolución espacial y temporal, que limita la capacidad para distinguir detalles finos; la presencia de artefactos de flujo pulsátil, que distorsionan las estructuras anatómicas debido al movimiento; y la pérdida de señal durante fases del ciclo cardíaco. Estas fases cardíacas, que incluyen la sístole (contracción del corazón para expulsar sangre) y la diástole (relajación del corazón para permitir su llenado), son fases que comúnmente se registran en estudios dinámicos, los cuales son obtenidos mediante resonancias magnéticas por contraste de fase (PC-MRI), donde se registran secuencias de imágenes que capturan el flujo sanguíneo durante dichas etapas. En particular, la diástole, al estar asociada con un flujo más bajo, tiende a ser especialmente propensa a la pérdida de información en las imágenes, lo que puede resultar en segmentaciones de mala calidad (Figura 2).

Estas inconsistencias no solo afectan la calidad visual de las imágenes segmentadas, sino que también generan incertidumbre en los datos derivados de estos análisis, como el área y flujo que son calculados tomando el conjunto de píxeles segmentados. Por ejemplo, la cantidad de flujo sanguíneo en un momento específico puede ser subestimada o sobreestimada debido a fallos en la segmentación, lo que impacta negativamente en los estudios de investigación que dependen de estas mediciones (Figura 3).

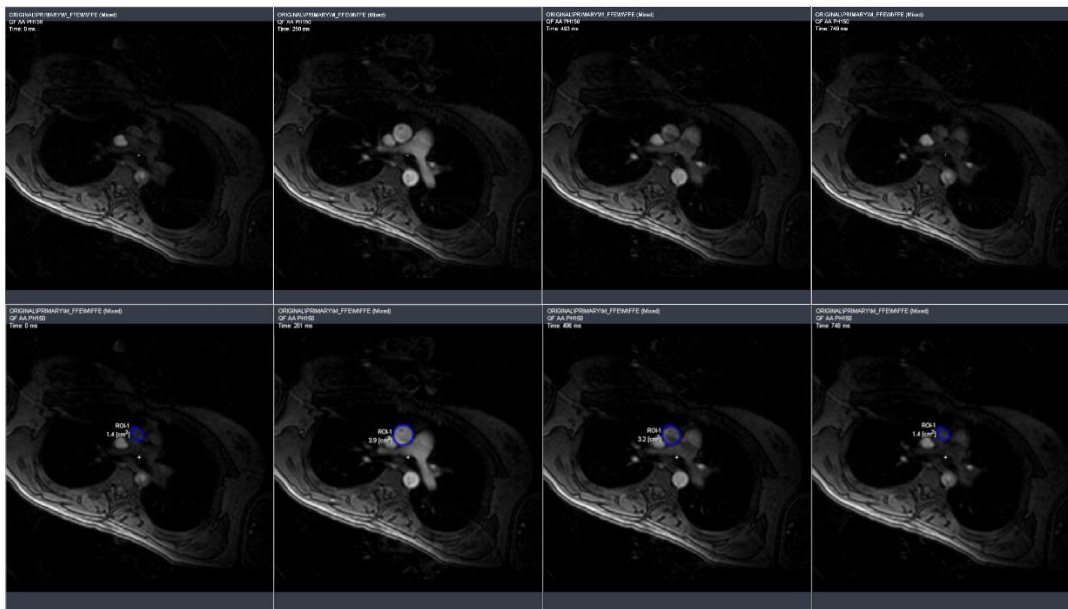


Figura 2: Segmentación de la aorta descendente durante periodos de alto y bajo contraste. Fuente: Elaboración propia.

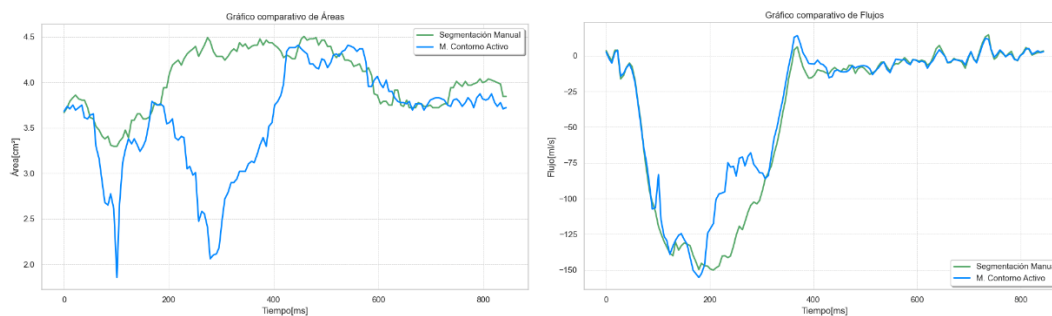


Figura 3: Gráficos de área y flujo calculados a partir de una segmentación manual y por un método de contorno activo. Fuente: Elaboración propia.

1.4 OBJETIVO

1.4.1 OBJETIVO GENERAL

Desarrollar un *framework* de segmentación eficiente de imágenes de contraste de fase 2D obtenidos por resonancia magnética que permite extraer el área y el flujo a partir de la aplicación de técnicas avanzadas de procesamiento de imágenes e inteligencia artificial.

1.4.2 OBJETIVOS ESPECIFICOS

- Diseñar e implementar un modelo de segmentación eficiente de imágenes de contraste de fase 2D obtenidas por resonancia magnética que logre representar el área en momentos de baja intensidad de señal.

- Desarrollar un *framework* para la segmentación de imágenes que integre el modelo de segmentación propuesto y que extraiga los datos del flujo y área de la región segmentada.
- Realizar experimentos utilizando conjuntos de datos clínicos para evaluar la precisión y confiabilidad del método propuesto.

CAPÍTULO 2: MARCO CONCEPTUAL

2.1 CICLO CARDÍACO

El ciclo cardíaco se define como una secuencia de la alternancia entre contracción y relajación de las aurículas y los ventrículos, con el fin de bombear sangre a través del cuerpo. Comienza al inicio de un latido cardíaco y termina al inicio del siguiente.

La **sístole** es la fase del ciclo cardíaco en la que las cámaras cardíacas se contraen y bombean sangre. Durante esta fase, los ventrículos se contraen, enviando sangre a los pulmones y a la periferia por medio de las arterias.

La **diástole** es la fase en la que las cámaras cardíacas están en estado de relajación y se llenan con sangre proveniente de las venas. Los atrios se llenan de sangre, mientras que los ventrículos también se preparan para la próxima contracción. Cuando los atrios están en diástole, los ventrículos están en sístole y viceversa. (kenhub, 2023)

2.2 IMÁGENES DE RESONANCIA MAGNÉTICA

La resonancia magnética (MRI) es una técnica de imagen no invasiva que utiliza campos magnéticos y ondas de radio para obtener imágenes detalladas del interior del cuerpo humano. Esta técnica se basa en principios físicos relacionados con las propiedades magnéticas de los núcleos atómicos, particularmente del hidrógeno, que es abundante en los tejidos biológicos debido a su presencia en el agua y las grasas.

El núcleo del átomo de hidrógeno tiene una propiedad intrínseca llamada **spin** nuclear, que le confiere un momento magnético, similar al de un diminuto imán. Bajo condiciones normales, estos momentos magnéticos están orientados al azar. Sin embargo, cuando el cuerpo es colocado en un campo magnético externo fuerte (B_0), estos momentos se alinean, ya sea paralelos o antiparalelos al campo magnético. Los núcleos se alinean en dos estados de energía: uno de **baja energía**, paralelo al campo, y otro de **alta energía**, antiparalelo al campo (Figura 4).

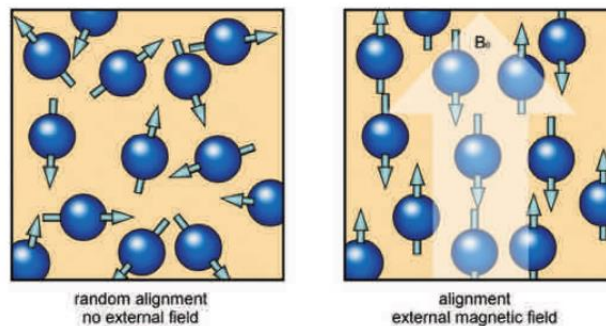


Figura 4: Alineamiento de los núcleos de hidrogeno. Fuente: (Westbrook et al., 2011)

Aunque ambos estados energéticos están presentes, hay una ligera mayoría de núcleos en el estado de baja energía, generando una magnetización neta longitudinal (NMV) en la dirección del campo B_0 (Figura 5).

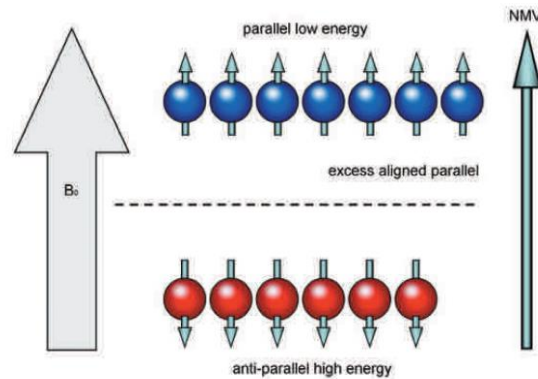


Figura 5: Vector de magnetización neta (NMV). Fuente: (Westbrook et al., 2011)

Los núcleos no se alinean perfectamente con el campo B_0 , en lugar de eso, giran alrededor de su eje en un movimiento conocido como **precesión**, con una frecuencia característica llamada **frecuencia de Larmor** (Figura 6). Esta frecuencia depende de la fuerza del campo magnético y de la **relación giromagnética**.

La frecuencia de Larmor viene dada por la siguiente ecuación:

$$\omega_0 = B_0 \cdot \lambda$$

Donde:

- ω_0 es la frecuencia de Larmor
- B_0 es la fuerza del campo magnético
- λ es la relación giromagnética, aproximadamente 42.58 [MHz/T] para los núcleos de hidrógeno.

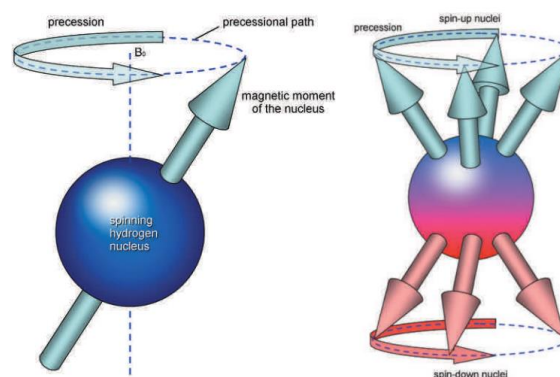


Figura 6: Precesión en los núcleos de hidrógeno. Fuente: (Westbrook et al., 2011)

Para generar una señal detectable, se aplica un pulso de radiofrecuencia (RF) a la frecuencia de Larmor, provocando resonancia en los núcleos. Los núcleos absorben esta energía y cambian de estado incrementando la población de núcleos de alta energía (Figura 7), esto produce una inclinación en el vector de magnetización neta (NMV) desde el plano longitudinal (Paralelo a B_0) hacia el plano transversal (Perpendicular a B_0) (Figura 8, a). Además, el pulso de RF sincroniza las fases de los núcleos, haciendo que precesen de manera alineada y sincrónica en el plano transversal. Este estado sincronizado es análogo a un grupo de relojes que se mueven al unísono, lo que permite generar una señal detectable (Figura 8, b).

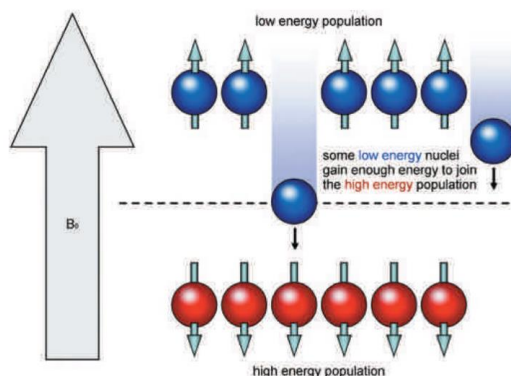


Figura 7: Transferencia de energía provocado por el pulso de radiofrecuencia. Fuente: (Westbrook et al., 2011)

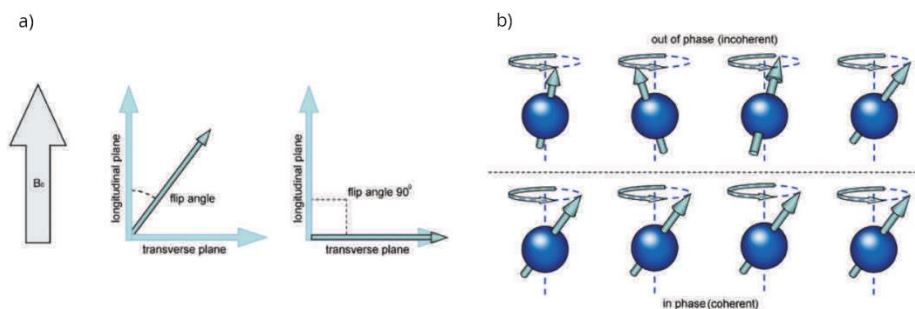


Figura 8: Efectos de la resonancia en los núcleos. A) Fuente: (Westbrook et al., 2011)

Al terminar el pulso de RF, los núcleos regresan gradualmente a su estado original, emitiendo la energía absorbida en forma de señales electromagnéticas (Figura 9). Este proceso se conoce como relajación, de las cuales se describen dos tipos:

- Relajación longitudinal (T1): Ocurre cuando los núcleos liberan la energía adquirida durante el pulso de RF al entorno molecular que los rodea. Este intercambio energético permite que los momentos magnéticos de los núcleos recuperen su magnetización en el plano longitudinal, es decir, la dirección paralela al campo magnético principal (B_0). T1 es el tiempo que tarda el 63% de la magnetización longitudinal en recuperarse en el tejido. (Figura 10).

- Relajación transversal (T2): Es causado por las interacciones entre los campos magnéticos de núcleos vecinos. Estas interacciones generan pequeñas variaciones en las frecuencias de precesión de los núcleos, lo que lleva a una pérdida progresiva de la coherencia en la fase de los protones. Como resultado, se reduce la magnetización en el plano transversal. T2 es el tiempo que tarda en perderse el 63% de la magnetización transversal (quedando un 37%) (Figura 11).

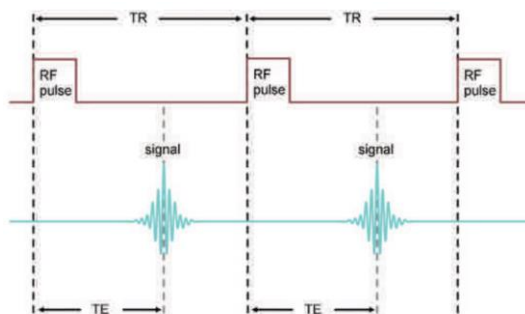


Figura 9: Secuencia básica de pulsos de RF. TR es el tiempo de repetición de los pulsos RF, y TE es el tiempo de eco, que corresponde al tiempo entre el pulso RF y la captura de la señal. Fuente: (Westbrook et al., 2011)

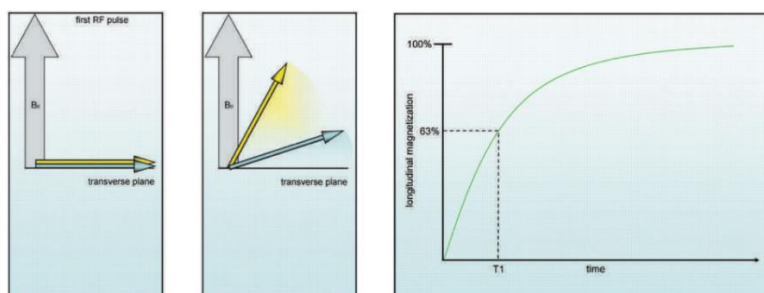


Figura 10: Relajación longitudinal T1. Fuente: (Westbrook et al., 2011)

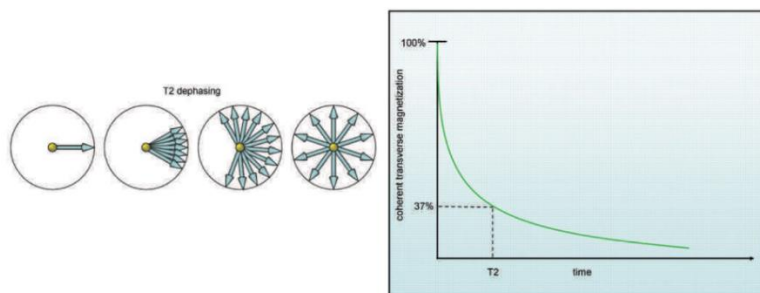


Figura 11: Relajación Transversal T2. Fuente: (Westbrook et al., 2011)

Las señales emitidas por los núcleos durante la relajación son captadas por las bobinas receptoras del escáner. Para saber de dónde proviene cada señal en el cuerpo, se usan gradientes magnéticos. Estos gradientes añaden variaciones al campo magnético en diferentes direcciones, lo que permite etiquetar las señales según su posición al provocar variaciones en las fases en el espacio (Figura 12).

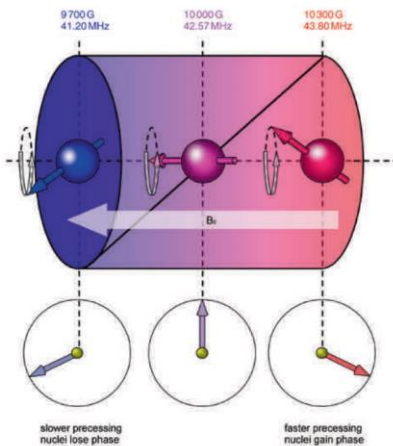


Figura 12: Gradiente del campo magnético. Fuente: (Westbrook et al., 2011)

Los tiempos de relajación T1 y T2 no son iguales para todos los tejidos del cuerpo: la grasa tiene un T1 y T2 breves, mientras que los líquidos presentan un T1 y T2 prolongados. El Tiempo de Repetición (TR) controla el tiempo entre dos pulsos de RF consecutivos, afectando el contraste T1: un TR corto permite que la magnetización longitudinal (T1) no se recupere completamente, resaltando tejidos con T1 corto, como la grasa (Figura 13, a). El Tiempo de Eco (TE) es el tiempo entre el pulso de RF y la captura de la señal, y controla el contraste T2: un TE largo permite que los tejidos con T2 largo, como los líquidos, se muestren con mayor claridad al captar la señal más tarde, cuando la magnetización transversal (T2) ha decaído (Figura 13, b).

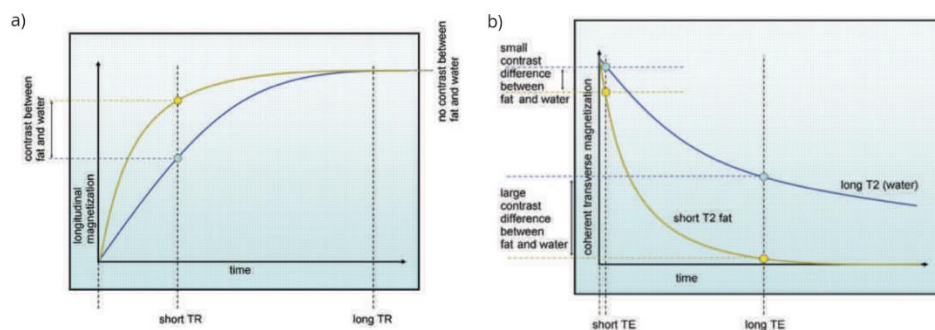


Figura 13: Diferencias de T1 y T2 entre las grasas y los líquidos. Fuente: (Westbrook et al., 2011)

Es posible obtener una imagen de densidad de protones (PD), utilizando un TR largo y un TE corto, el cual minimiza el contraste basado en T1 y T2. Esto resalta la densidad de protones en los tejidos sin hacer énfasis en sus tiempos de relajación, mostrando una imagen donde la diferencia de intensidad refleja principalmente la cantidad de protones (agua y grasa) en cada área.

Las señales detectadas se organizan en un espacio matemático llamado *k-space*. Aquí no se ve una imagen aún, sino que una colección de datos que describen las frecuencias y fases de las señales en todo el cuerpo. Cada punto del *k-space* representa una combinación específica de frecuencia y fase, y juntos contienen la información necesaria

para formar la imagen. Mediante la Transformada de Fourier, los datos del k -space se convierten en datos visuales que reflejan los tejidos del cuerpo (Figura 14). (Westbrook, 2011)

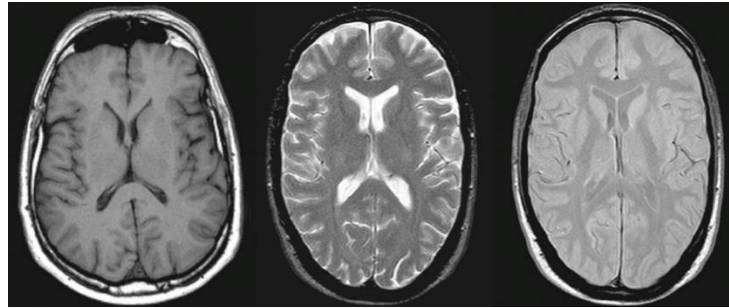


Figura 14: Imagen axial del cerebro humano ponderada en T1 (izquierda), T2 (Centro) y PD (Derecha).

Fuente: (Westbrook et al., 2011)

2.2.1 IMÁGENES DE RESONANCIA MAGNÉTICA POR CONTRASTE DE FASE

La resonancia magnética por contraste de fase (PC-MRI) es una técnica utilizada para medir el movimiento de fluidos dentro del cuerpo, como el flujo sanguíneo o el líquido cefalorraquídeo. A diferencia de las imágenes estáticas convencionales, la PC-MRI es capaz de detectar y cuantificar el movimiento de las moléculas en función de su velocidad y dirección. Al igual que en el método MRI convencional, utiliza el fenómeno de resonancia magnética nuclear a través de campos magnéticos y ondas de radiofrecuencia para obtener imágenes del interior del cuerpo humano.

El componente distintivo de la PC-MRI se encuentra en el uso de gradientes bipolares de campo magnético, los cuales son fundamentales para codificar el movimiento de los protones. Estos gradientes se aplican en la dirección del flujo que se desea medir. El primer gradiente altera la velocidad de precesión de los protones en función de su posición y dirección de movimiento. El segundo gradiente, de igual magnitud, pero polaridad opuesta, idealmente cancela este efecto en los protones en reposo, devolviéndolos a su fase inicial. Sin embargo, los protones en movimiento no logran sincronizarse completamente debido a su desplazamiento físico durante la aplicación de los gradientes, acumulando así un desplazamiento de fase proporcional a su velocidad (Figura 15).

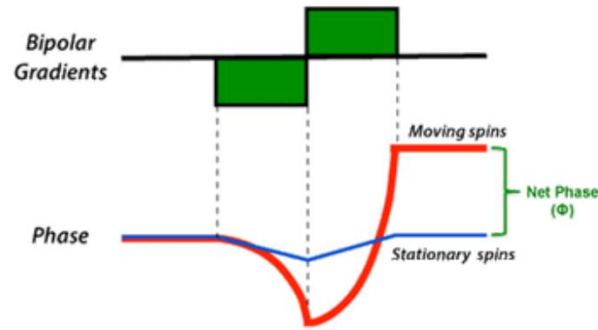


Figura 15: Efecto de los gradientes bipolares. El desplazamiento de fase es cero para núcleos estacionarios, pero positiva para núcleos que se mueven en la dirección del gradiente. Fuente: (*mriquestions*, 2024)

Este desplazamiento de fase es la clave de la PC-MRI. Al adquirir dos imágenes, una con los gradientes activados y otra sin ellos, se puede calcular la diferencia de fase, aislando la contribución del movimiento. La relación matemática que vincula el desplazamiento de fase con la velocidad del flujo es:

$$\Delta\varphi = \gamma \cdot \vec{v} \cdot \Delta M_1$$

Donde:

- $\Delta\varphi$ es el desplazamiento de fase.
- γ es la relación giromagnética.
- \vec{v} es la velocidad del núcleo.
- ΔM_1 es el cambio en el momento magnético.

El resultado de este proceso es la obtención de dos tipos de imágenes: las imágenes de magnitud y fase. La imagen de **magnitud** refleja la intensidad de la señal, similar a las imágenes obtenidas en una resonancia magnética convencional, proporcionando información sobre la estructura anatómica del cuerpo. La imagen de **fase**, por otro lado, muestra el desplazamiento de fase acumulado por los protones debido al movimiento del fluido, lo que permite calcular la velocidad y dirección del flujo (Figura 16). (Wymer et al., 2020)

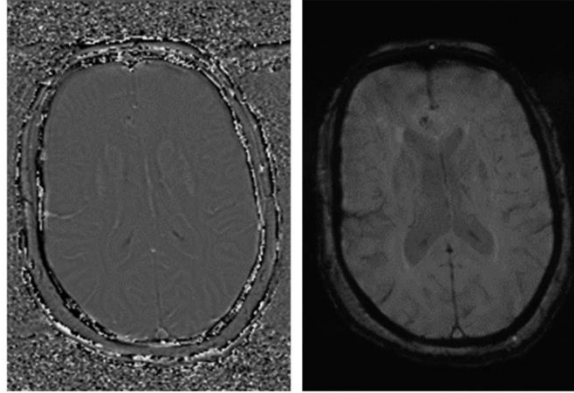


Figura 16: Ejemplo de 2D PC-MRI, Se observa la imagen de fase (izquierda) y la imagen de magnitud (derecha). Fuente: (Wymer et al., 2020)

2.2.2 ELECTROCARDIOGRAMA

Un electrocardiograma (ECG) es un procedimiento simple, indoloro y rápido que registra la actividad eléctrica del corazón. Cada vez que el corazón late, una señal eléctrica circula a través de él, activando las cuatro cámaras cardíacas para que se contraigan al ritmo correcto y así bombear sangre al cuerpo.

El registro de estas señales en el ECG se presenta como líneas onduladas, que los profesionales de la salud interpretan para identificar posibles anomalías o daños cardíacos. El ECG puede revelar: La velocidad a la que late el corazón. Si el ritmo cardíaco es constante o irregular. La fuerza y sincronización de las señales eléctricas que recorren cada parte del corazón. En algunos casos, el tamaño y posición de las cavidades del corazón. (medlineplus, 2023)

2.2.3 CINE MRI

El *cine imaging* o *cine MRI* es una técnica de resonancia magnética utilizada para capturar imágenes en movimiento, particularmente útil para estudiar el corazón y otros órganos que se mueven durante el ciclo cardíaco o respiratorio. Se basa en la adquisición de múltiples imágenes en diferentes fases del ciclo cardíaco, sincronizadas con el electrocardiograma (ECG), lo que permite observar el movimiento del órgano en tiempo real. Para ello, se asignan diferentes líneas del espacio-k a cada fase del ciclo y las imágenes se reconstruyen y muestran en secuencias, como si fueran una película. (radiopaedia, 2024)

Esta técnica también puede extrapolarse a las PC-MRI (Figura 17), que permite medir tanto el flujo como el movimiento de los tejidos (en el informe se tratará el término **cine PC-MRI** como **secuencia PC-MRI**).

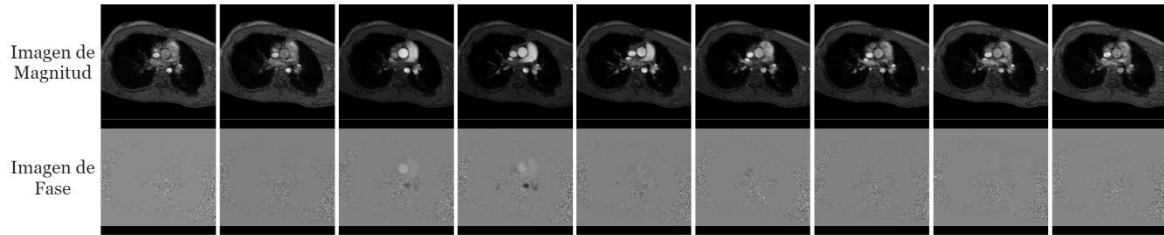


Figura 17: Ejemplo de cine PC-MRI con 9 instantes de tiempo, donde se observan las imágenes de magnitud y fase. Fuente: Elaboración Propia.

2.3 SEGMENTACIÓN DE IMÁGENES MEDICAS

La segmentación de imágenes médicas es un proceso en el cual se delimitan regiones de interés en la imagen con el propósito de extraer información relevante, como la representación y descripción de estructuras internas, la medición de características y la clasificación de objetos. La segmentación cumple un rol fundamental para el análisis y procesamiento de imágenes médicas, ya que los resultados de la segmentación afectan a todos los procesos posteriores al análisis de la imagen.

En (Patil & Deore, 2013) se da a entender que la segmentación manual de imágenes médicas por parte del médico especialista es un proceso tedioso e impreciso al depender exclusivamente del criterio del experto, especialmente con el aumento de modalidades de imágenes médicas y la cantidad inmanejable de imágenes médicas que deben examinarse. Por lo que se vuelve necesario revisar las metodologías actuales de segmentación de imágenes utilizando algoritmos automatizados que sean precisos y requieran la menor interacción del usuario posible, especialmente en el caso de imágenes médicas.

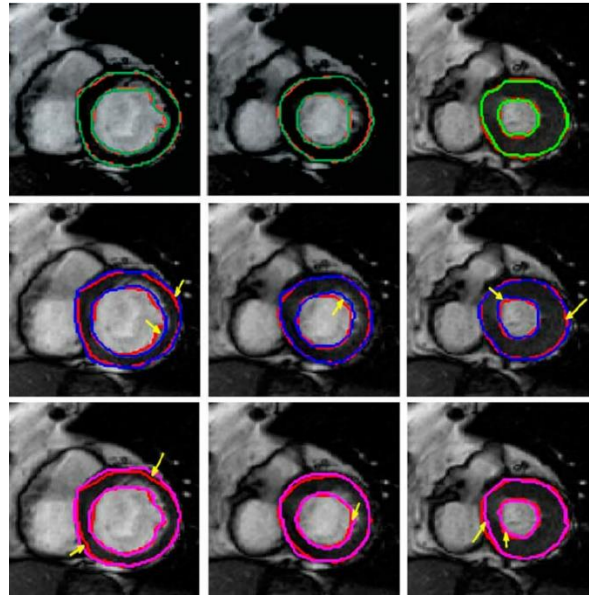


Figura 18: Segmentación de una imagen MRI con diferentes métodos. En rojo se marca la segmentación manual, mientras que los otros colores corresponden a diferentes métodos de segmentación automatizados. Fuente: (Mahapatra, 2013)

2.3.1 SEGMENTACIÓN POR CONTORNO ACTIVO

La segmentación por contorno activo, también conocida como *snakes*, es una técnica utilizada para identificar y delinear características específicas en imágenes, como bordes y líneas. Esta metodología se basa en un modelo matemático que representa el contorno deseado como una curva que puede deformarse. El contorno activo es guiado por vectores de gradiente (fuerzas externas), que lo atraen hacia características importantes de la imagen, lo que le permite "atrapar" contornos cercanos y adaptarse dinámicamente a las variaciones en los datos visuales.

El funcionamiento de los contornos activos implica la minimización de una función de energía que se compone de diferentes términos. Estos términos pueden representar tanto las **fuerzas internas**, que actúan como resortes dentro del contorno para mantenerlo suave y continuo, evitando que tenga deformaciones bruscas o se rompa, y las **fuerzas externas**, que lo empujan o tiran hacia las características de interés de la imagen, como bordes y líneas. La combinación de estas fuerzas permite que el contorno se ajuste a las formas de interés en la imagen, facilitando la segmentación de los objetos. (Kass, 1988)

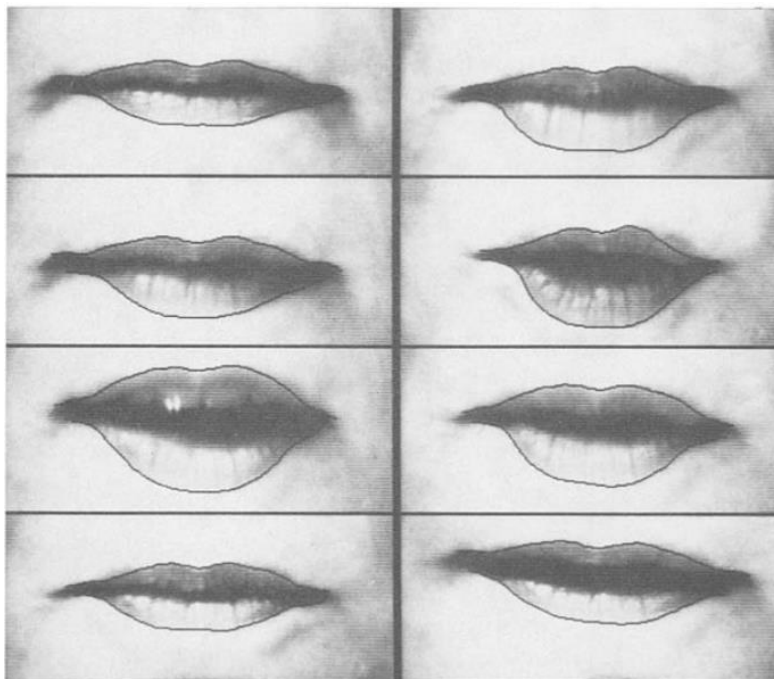


Figura 19: Fotogramas seleccionados desde un video de 2 segundos. El *snake* se ajusta automáticamente a los movimientos de los labios. Fuente: (Kass et al., 1988)

2.4 FORMATO DICOM

El estándar DICOM es un tipo de formato de imágenes que es utilizado para almacenar imágenes de resonancia magnética, esta responde a las siglas en inglés de *Digital Imaging and Communications in Medicine*. Se trata de un protocolo de comunicación estándar de carácter internacional que unifica los formatos de almacenamiento de las imágenes médicas y la transmisión de los mismos de un hardware a otro.

El formato DICOM no solo incluye información visual, también incluye la información relacionada con el paciente del que se ha obtenido. Por ejemplo, el ID del paciente, su ubicación geográfica, su grupo demográfico, qué médico pidió esa prueba o en qué fecha se realizó. Gracias a que esta información permanece con la imagen, no hay posibilidad de confusión o error, y atribuir esa prueba a otro paciente (Nubimed, 2021).

2.5 REDES NEURONALES

Las redes neuronales son sistemas computacionales inspirados en el cerebro humano, diseñados para aprender de ejemplos y resolver problemas complejos. Estas redes están compuestas por capas de neuronas interconectadas, que procesan la información para realizar tareas como clasificación, predicción o aproximación de funciones.

Las redes neuronales se componen de tres partes principales: la capa de entrada, las capas ocultas y la capa de salida. La capa de entrada recibe los datos iniciales que la red debe procesar. Las capas ocultas realizan los cálculos intermedios, donde cada neurona aplica

una función de activación para procesar la información. Finalmente, la capa de salida proporciona el resultado final, que puede ser un valor numérico o una categoría (Figura 20).

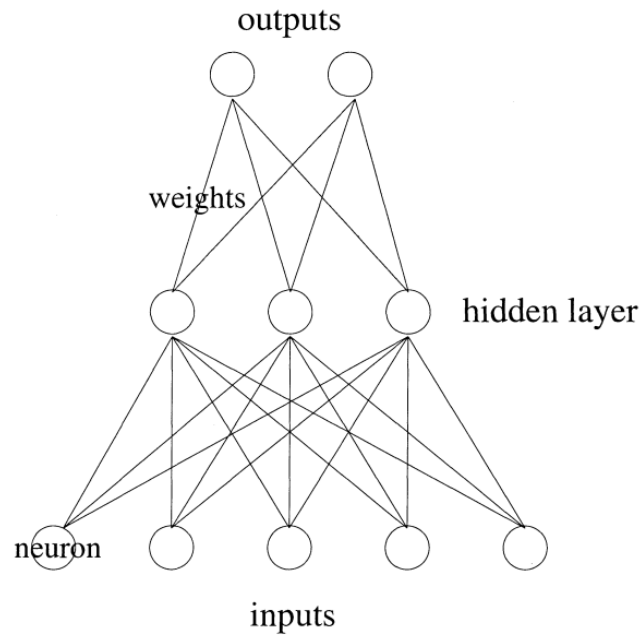


Figura 20: Arquitectura de una red neuronal. Fuente: (Wang S.-C. , 2003)

Las neuronas de las redes están conectadas entre sí mediante pesos, que son ajustados durante el proceso de entrenamiento. Este ajuste se realiza para minimizar un error, generalmente basado en la diferencia entre las salidas predichas y las salidas deseadas. (Wang S.-C. , 2003)

2.5.1 NEURONA ARTIFICIAL

El Nodo o Neurona Artificial es la unidad básica de una red neuronal. Funciona como una versión simplificada de una neurona biológica. Cada nodo recibe una entrada, realiza una operación matemática en esa entrada y produce una salida.

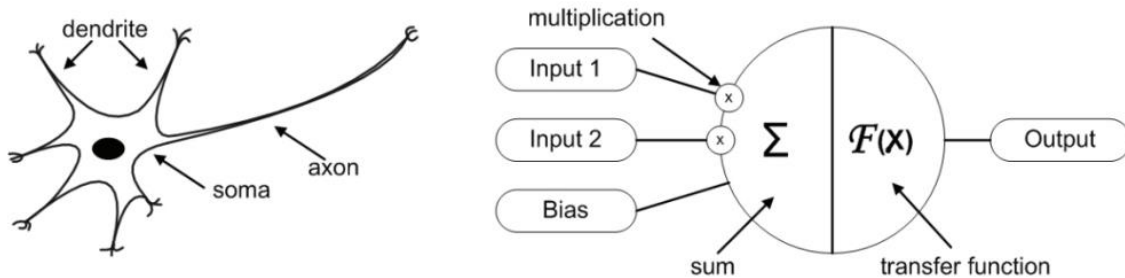


Figura 21: Comparación entre una neurona biológica y una neurona artificial. Fuente: (Krenker et al., 2011)

En el caso de la neurona biológica, como es explicado en (Krenker et al., 2011), la información ingresa a la neurona a través de la dendrita, el soma procesa la información y la transmite a través del axón. En el caso de la neurona artificial, la información ingresa a una neurona artificial por medio de entradas ponderadas (cada valor de la entrada tiene un peso w asociado). En el interior de la neurona artificial, se suman las entradas ponderadas junto al umbral, y el resultado se ingresa a una función de activación. Al final, la neurona artificial transmite la información procesada a través de sus salidas. La descripción matemática de esta neuronal artificial se puede ver a continuación:

$$y(k) = F\left(\sum_{i=0}^m w_i(k) * x_i(k) + b\right)$$

Donde:

- $x_i(k)$ es un valor de entrada en un tiempo k .
- $W_i(k)$ es un valor de ponderación en un tiempo discreto k .
- b es un umbral de activación.
- $F(x)$ es la función de activación.
- $y(k)$ es el valor de salida para un tiempo k .

2.5.2 FUNCIONES DE ACTIVACIÓN

Una función de activación en una red neuronal artificial es un componente que se utiliza para transformar un tipo de señal por otro. Su función principal es introducir no linealidades en la red, lo que permite a la red aprender patrones más complejos y realizar tareas más avanzadas. En otras palabras, las funciones de activación ayudan a la red a aprender y modelar relaciones no lineales entre las entradas y las salidas. (Sharma et al., 2017)

Algunas de las funciones de activación más conocidas son las siguientes:

- **Función de activación Lineal:** La función de activación lineal es directamente proporcional a su entrada. Aunque evita la derivada constante cero¹, no mejora la capacidad de la red para identificar patrones complejos. Es utilizada para problemas sencillos.

$$f(x) = ax$$

¹ se refiere a la situación en la que la derivada de la función de activación de una neurona es cero, al ocurrir esto, la neurona no puede actualizar sus pesos durante el entrenamiento utilizando el algoritmo de retropropagación, lo que puede llevar a una red neuronal que no aprende correctamente.

- **Función de activación sigmoide:** La función sigmoide es ampliamente utilizada debido a su naturaleza no lineal y su capacidad para transformar valores en un rango acotado de 0 a 1. Su suavidad y continuidad facilitan el entrenamiento, pero su falta de simetría alrededor de cero puede resultar en que todas las neuronas tengan el mismo signo de salida.

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

- **Función de activación tangente hiperbólica:** Similar a la función sigmoide, la función tangente hiperbólica (Tanh) es simétrica alrededor del origen y produce salidas en el rango de -1 a 1. Su ventaja radica en gradientes más pronunciados en comparación con la función sigmoide, siendo preferida por tener gradientes no restringidos en una dirección específica.

$$f(x) = 2\text{sigmoid}(2x) - 1$$

- **Función de activación ReLU:** La función de unidad lineal rectificadora (ReLU) es no lineal y ampliamente utilizada. Su eficiencia proviene de la activación selectiva de neuronas, pero puede sufrir el problema de "ReLU muerta", donde algunas neuronas se vuelven inactivas durante el entrenamiento debido a gradientes cero para ciertos valores de entrada.

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

- **Función de activación softmax:** La función softmax es una combinación de múltiples funciones sigmoideas. La sigmoide al entregar valores entre 0 y 1, esta puede ser tratada como una probabilidad. La función de activación softmax, es utilizada en problemas de clasificación multiclase.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \text{ para } j = 1, \dots, K$$

2.6 REDES NEURONALES CONVOLUCIONALES

Las Redes Neuronales Convolucionales (CNN) son análogas a las Redes Neuronales Artificiales tradicionales en el sentido de que están compuestas por neuronas que se optimizan automáticamente mediante el proceso de aprendizaje. Cada neurona aún recibirá una entrada y realizará una operación.

La única diferencia notable entre las CNN y las RNA tradicionales es que las CNN se utilizan principalmente en el campo del reconocimiento de patrones dentro de imágenes. Esto permite codificar características específicas de la imagen en la arquitectura, haciendo que la red sea más adecuada para tareas centradas en imágenes, al mismo tiempo que se reducen aún más los parámetros necesarios para configurar el modelo.

2.6.1 ARQUITECTURA DE UNA CNN

En (O'Shea & Nash, 2015) se observa que, en una CNN, las neuronas en las capas están organizadas en tres dimensiones: la dimensionalidad espacial de la entrada (altura y anchura) y la profundidad. Con profundidad no se refiere al número de capas al interior de la red, como en una ANN. Si no a una tercera dimensión de un volumen de activación. A diferencia de las ANN estándar, las neuronas dentro de cualquier capa solo se conectarán a una pequeña región de la capa que la precede.

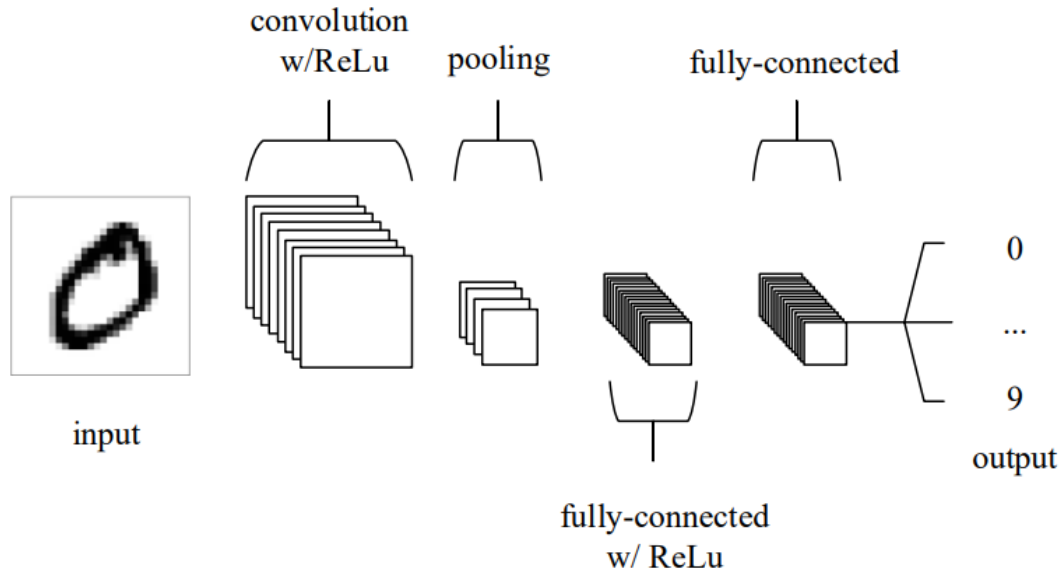


Figura 22: Arquitectura de una CNN. Se aprecia la capa convolucional con función de activación ReLu, la capa *pooling*, y dos capas completamente conectadas. Fuente: (O'Shea & Nash, 2015)

En la Figura 22 se distinguen las siguientes áreas de una arquitectura CNN:

- **Capa de entrada (*input*):** La capa de entrada recibe la imagen original, que se representa como un tensor tridimensional (altura, ancho, profundidad). La profundidad se refiere a los canales de color de la imagen (por ejemplo, 3 canales para una imagen RGB). Como ejemplo: Una entrada típica puede ser una imagen de 64x64x3, donde el volumen corresponde a las dimensiones espaciales de altura y anchura de 64 píxeles cada una, y una tercera dimensión correspondiente a los colores RGB. Si la imagen es en blanco y negro el volumen sería 64x64x1.
- **Capa convolucional:** La capa convolucional es la encargada de procesar los datos de entrada, como imágenes, mediante el uso de filtros llamados *kernels*. Estos *kernels* son pequeños en dimensiones espaciales, y son los que se ajustan en el proceso de aprendizaje. Cuando los datos ingresan a una capa convolucional, el filtro comienza a realizar operaciones de convolución al deslizar el filtro a lo largo de la toda dimensión espacial de la entrada (Figura 23). Para cada posición, se calcula el producto escalar

entre los valores de la entrada y los pesos del *kernel*. Este proceso genera mapas de activación en 2D que representan las características aprendidas por cada filtro, como se ve en la figura 24.

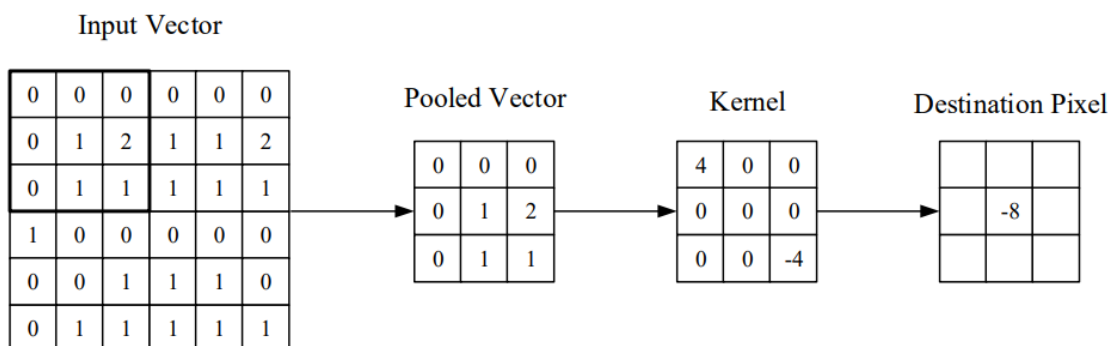


Figura 23: Representación visual de una capa convolucional. Se observa los datos de entrada los cuales son reducidos en dimensión por una capa de *pooling* para luego ser procesado por el *kernel*. Fuente: (O'Shea & Nash, 2015)

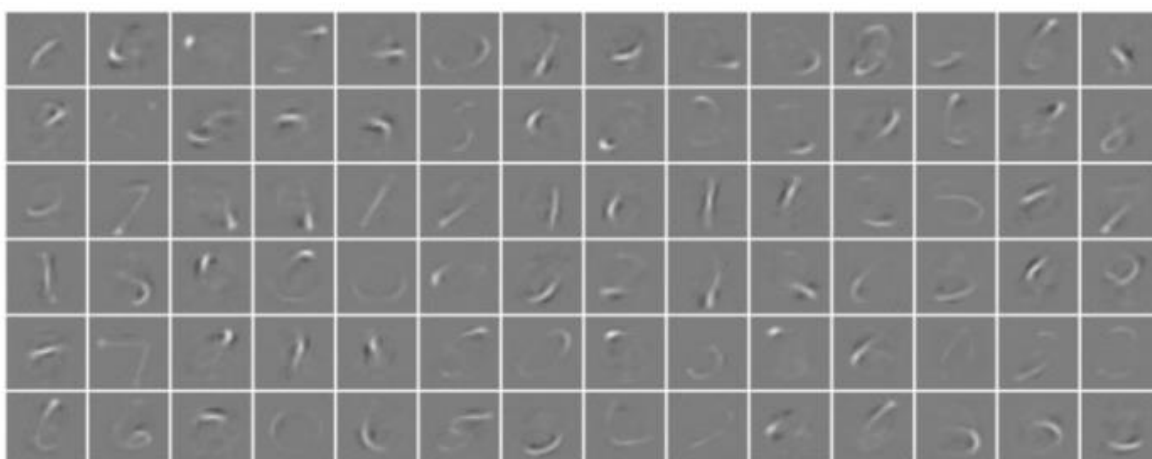


Figura 24: Mapas de activación aprendidos en cada filtro durante el entrenamiento de una CNN. Fuente: (O'Shea & Nash, 2015)

- **Capa de *pooling*:** La capa de *pooling* es una capa que se utiliza en las CNN para reducir el tamaño del mapa de características. El objetivo de la capa de *pooling* es reducir la cantidad de parámetros y cálculos necesarios en la red, lo que a su vez reduce el riesgo de sobreajuste².
- **Capa completamente conectada:** La capa completamente conectada, es una capa que se utiliza en las CNN para realizar la clasificación final de la imagen. Esta capa se encarga de tomar las características extraídas por las capas convolucionales y de

² El sobreajuste es un problema que ocurre cuando un modelo de aprendizaje automático se ajusta demasiado a los datos de entrenamiento y no generaliza bien a los datos nuevos.

pooling para transformarlas en una salida que corresponde a las clases de la tarea de clasificación.

2.7 APRENDIZAJE DE UNA RED NEURONAL ARTIFICIAL

El aprendizaje o entrenamiento de una red neuronal artificial se refiere al proceso mediante el cual la red neuronal artificial ajusta sus pesos y umbrales internos para mejorar su capacidad para realizar una tarea específica, como predecir, clasificar, generar imágenes u otra actividad. Para que la red aprenda, esta requiere de un conjunto de datos de entrenamiento, estos datos pueden venir etiquetados o no etiquetados. Se dice que un conjunto de datos está etiquetado cuando se explicita la salida esperada junto a los datos de entrada. Caso contrario, se dice no etiquetados cuando la salida esperada no está incluida. Según (Berry et al., 2019), el entrenamiento de una red puede ser clasificada en:

- **Aprendizaje supervisado:** Es cuando la red se entrena con datos etiquetados. Durante el proceso de entrenamiento, el algoritmo utiliza esta información para aprender la relación entre las entradas y las salidas, con el objetivo de generalizar esta relación y realizar predicciones precisas sobre datos no etiquetados. Este enfoque es utilizado para tareas de clasificación, donde la red debe asignar una clase a una entrada de datos específica, o para problemas de regresión, donde se busca predecir un valor numérico en base a unos valores de entrada.
- **Aprendizaje no supervisado:** Es cuando la red se entrena con datos no etiquetados. En este tipo de enfoque, la red en lugar de buscar relaciones específicas entre entradas y salidas, se centra en descubrir patrones o estructuras inherentes en los datos. Este enfoque es utilizado para tareas como la identificación de grupos similares en conjuntos de datos (*clustering*) o la extracción de reglas que describen relaciones entre variables (*association mining*). Además, los algoritmos de aprendizaje no supervisado también pueden ser utilizados para la generación de etiquetas en datos no etiquetados, las cuales posteriormente pueden ser utilizados en un entrenamiento supervisado.

2.7.1 FUNCIÓN DE PERDIDA (LOSS FUNCTION)

Una función de pérdida es una función matemática que evalúa la discrepancia entre las predicciones de un modelo y los valores reales en el conjunto de entrenamiento. En el aprendizaje supervisado, la función de pérdida guía el proceso de entrenamiento al evaluar la precisión de las predicciones y ajustar los parámetros del modelo para minimizar esta pérdida. En redes neuronales profundas, la función de pérdida se utiliza en el proceso de retropropagación (*Backpropagation*) para calcular el gradiente y ajustar los pesos de la red, mejorando la capacidad del modelo para generalizar a datos no vistos. La elección de la función de pérdida es fundamental, ya que influye directamente en el entrenamiento de la red y en su capacidad de aprender patrones. (Wang, et al., 2020)

En (Shruti, 2020) se destacan las siguientes funciones de perdida, las cuales son utilizadas en el área de la segmentación de imágenes:

- **Binary Cross-Entropy Loss:** Es una medida de la diferencia entre dos distribuciones de probabilidad, en este caso, la distribución de probabilidad de la predicción del modelo y la distribución de probabilidad de la etiqueta verdadera. Es ampliamente utilizada para la clasificación.

$$L_{BCE}(y, \hat{y}) = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y}))$$

- “ y ” es el valor verdadero.
- “ \hat{y} ” es el valor estimado.

- **Dice Loss:** Es una función de perdida que se basa en el coeficiente de similitud de *Dice*, el cual mide la similitud entre dos imágenes.

$$DL(y, \hat{p}) = 1 - \frac{2y\hat{p} + 1}{y + \hat{p} + 1}$$

- Se añade 1 al numerador y denominador para asegurar que la función no se indefina en escenarios borde donde $y = \hat{p} = 0$.
- “ \hat{p} ” es la predicción del modelo.

- **Tversky Loss:** Esta función se basa en el índice de *Tversky*, que es una generalización del coeficiente de similitud de *Dice*. En esta función se añade peso a los falsos positivos (FP) y falsos negativos (FN) con la ayuda del coeficiente β .

$$TI(p, \hat{p}) = \frac{p\hat{p}}{p\hat{p} + \beta(1 - p)\hat{p} + (1 - \beta)p(1 - \hat{p})}$$

CAPÍTULO 3: PROPUESTA DE SOLUCION

En este capítulo se aborda la implementación de un *framework* para la segmentación de secuencias de imágenes de contraste de fase 2D obtenidas por resonancia magnética (2D PC-MRI). Se describe el conjunto de datos utilizados en el trabajo y se propone una metodología para abordar los problemas de segmentación de las secuencias de 2D PC-MRI. Junto con ello, se estudia la arquitectura Unet y su adaptación para la segmentación de volúmenes, donde se propone una implementación de la arquitectura 3D Unet.

3.1 CONJUNTO DE DATOS

En este trabajo se utilizará un conjunto de datos que consta de 140 secuencias 2D PC-MRI en formato DICOM, obtenidas de aproximadamente 15 pacientes con coartación aortica (Sotelo et al., 2015). Estos archivos contienen imágenes de magnitud y fase a lo largo del ciclo cardiaco, y cubren las secciones axiales, sagitales y coronales de la arteria aorta. Las imágenes fueron adquiridas principalmente para estudiar la distribución de flujo de estos pacientes.

Las secuencias 2D PC-MRI contienen entre 13 a 152 imágenes con un promedio 59, las cuales registran el flujo cardiaco durante las fases cardiacas de sístole y diástole. Además, cada imagen tiene una resolución de 256x256 píxeles en escala de grises.

Tabla 1: Tipos de corte en las secuencias PC-MRI del conjunto de datos. Fuente: Elaboración Propia.

Tipo corte PC-MRI	Cantidad
Abdominal	50
Sagital	25
Aorta Ascendente/Descendente	50
Coronal	3
Supra	12
Total	140

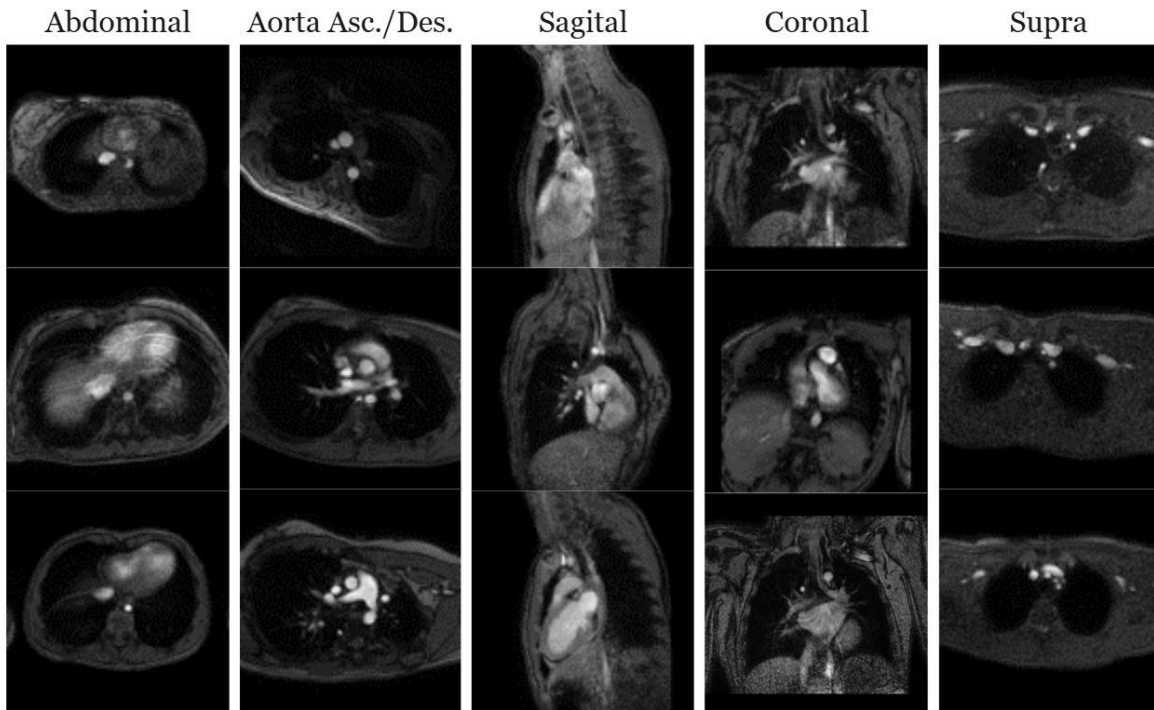


Figura 25: Muestra de imágenes para los diferentes tipos de cortes en el conjunto de datos. Fuente: Elaboración Propia.

3.2 METODO PROPUESTO

Uno de los principales problemas de los métodos de segmentación basados en contornos activos es que su precisión depende en gran medida de la calidad de la imagen. Si los contornos de la estructura a segmentar son poco definidos, la segmentación será imprecisa, requiriendo de la intervención manual para reparar la segmentación. En las MRI, esta limitación se vuelve más problemática, ya que los bordes de los tejidos suelen aparecer borrosos o interrumpidos. Además, la presencia de artefactos como el movimiento del paciente, las interferencias del equipo o las variaciones en la intensidad de señal entre diferentes regiones dificulta aún más la identificación precisa de los contornos, provocando errores en la segmentación.

Este problema se vuelve más complejo cuando se trata de segmentar secuencias de imágenes PC-MRI, ya que la variabilidad entre fotogramas y las deformaciones dinámicas de los tejidos pueden dificultar aún más la consistencia y la precisión de los contornos segmentados.

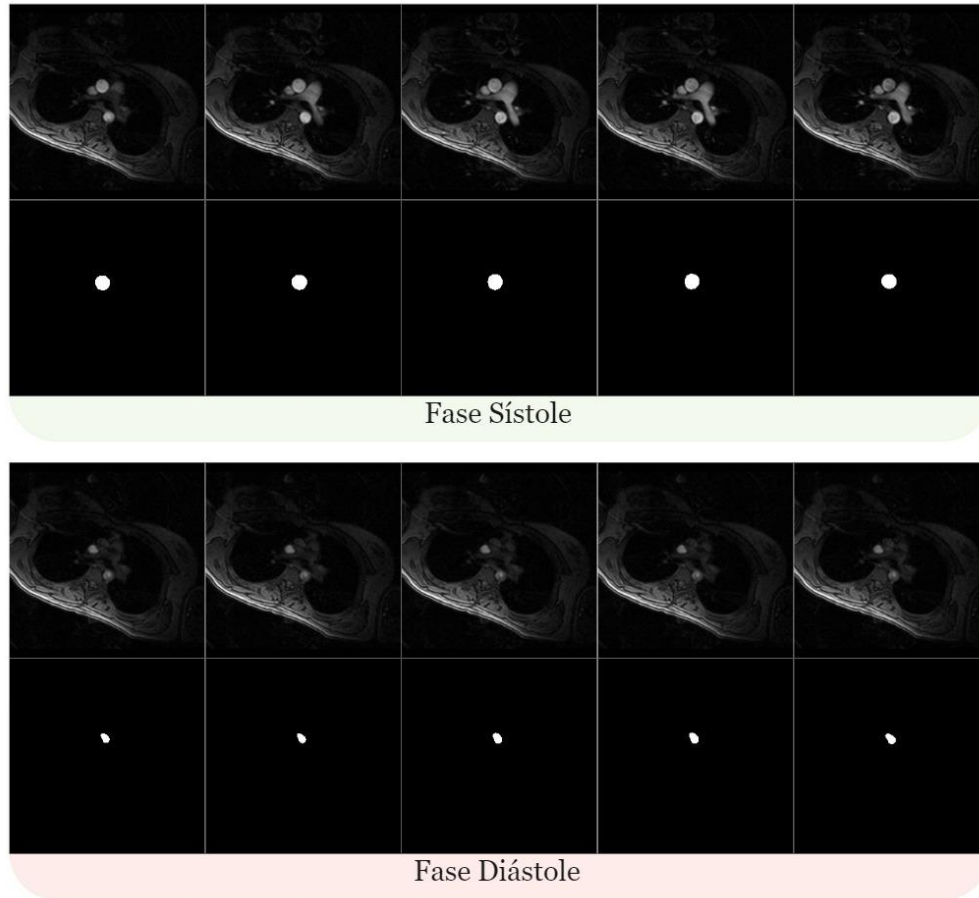


Figura 26: Visualización de una segmentación por contorno activo de la aorta ascendente durante la fase sistólica y diastólica. Para ambas fases se observa una muestra de las imágenes de magnitud. Fuente: Elaboración Propia.

En la Figura 26, se observa como la segmentación durante la fase de sístole, el contorno de la aorta ascendente adquiere una buena definición, lo que permite al método de contorno activo una segmentación de buena calidad. Pero durante la fase de diástole, el contorno se comienza a difuminar con el fondo, provocando que el método de segmentación pierda el contorno real de la estructura aortica, dando como resultado una segmentación deficiente que no logra segmentar toda el área requerida durante dicha fase.

Bajo observación humana, es posible segmentar manualmente el área de la aorta durante la fase de diástole, tomando como información los contrastes que aún siguen remanentes en las imágenes, y también observando la forma de la aorta en los momentos donde presentó mayor contraste. Siguiendo esta idea, una buena forma para segmentar la región de interés (ROI) cuando la imagen pierda claridad en los contornos, es aprovechar la información de la ROI en otros momentos del ciclo cardiaco, donde pueda tener un mejor contraste y permita predecir el área de mejor forma cuando la claridad se pierda. Cosa

que el método de contorno activo clásico no puede lograr ya que cada imagen es procesada individualmente.

Para lograr un método que consiga dichos objetivos, sería a través del uso de técnicas de *Deep Learning*, específicamente utilizando redes neuronales convolucionales (CNN). Estas redes son particularmente útiles para tareas de procesamiento de imágenes debido a su capacidad para aprender patrones espaciales y contextuales complejos. Para el caso de la segmentación de imágenes MRI, una CNN puede entrenarse para identificar patrones y relaciones espaciales en las imágenes de alta claridad y luego aplicarlos a momentos donde la ROI presente contornos más difusos.

3.2.1 UNET

La arquitectura Unet es una red neuronal convolucional ampliamente utilizada en tareas de segmentación de imágenes, donde el objetivo es clasificar cada píxel en una categoría específica. La Unet se caracteriza por su estructura simétrica en forma de U, compuesta por dos partes fundamentales: el *encoder* y el *decoder*.

- **Encoder:** El *encoder*, también conocido como la fase de contracción, se asemeja a una CNN tradicional. Consiste en una serie de capas convolucionales seguidas de capas de *pooling*, cuyo propósito es capturar el contexto de la imagen reduciendo su dimensionalidad. A medida que se avanza en esta parte de la red, se disminuye la resolución espacial, pero se incrementa la profundidad, es decir, el número de filtros aplicados.
- **Decoder:** El *decoder*, o fase de expansión, realiza el proceso inverso. Emplea capas de convolución transpuesta, también conocidas como convoluciones de *upsampling*, para aumentar la resolución espacial y reconstruir la imagen a su tamaño original. Además, en cada etapa del *decoder*, se concatenan características (*feature maps*) de la parte correspondiente del *encoder* mediante conexiones directas (*skip connections*), lo que permite recuperar detalles finos que podrían haberse perdido durante la contracción.

Este enfoque permite que la UNet capture tanto el contexto global como los detalles locales de la imagen, lo que la convierte en una herramienta excelente para la segmentación de imágenes en diversas aplicaciones. (Yin, Sun et al., 2022)

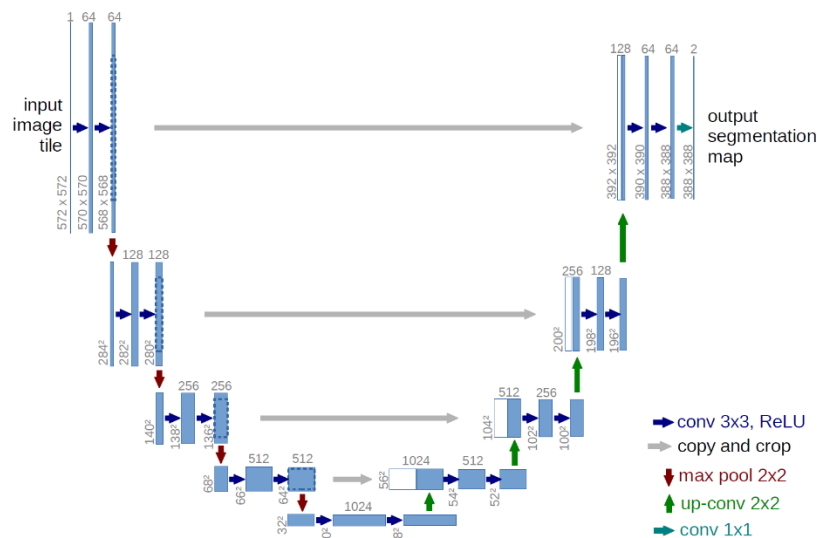


Figura 27: Arquitectura del modelo Unet. Fuente: (Ronneberger et al., 2015)

3.2.2 VOLUMENES DE IMÁGENES

La arquitectura UNet es una buena base para un modelo de segmentación que logre el objetivo planteado de aprovechar la información espacial de la imagen MRI para la segmentación. Pero la arquitectura UNet convencional solo procesa imágenes individuales y no secuenciales, por lo que la información temporal de la ROI no se estaría aprovechando.

Una arquitectura adecuada para este tipo de tarea es una Red Neuronal Recurrente (RNN), que está diseñada para procesar secuencias de datos. A diferencia de las redes neuronales tradicionales, las RNN tienen conexiones entre las neuronas de las capas ocultas, lo que les permite mantener y utilizar información de pasos anteriores. Esto las capacita para capturar dependencias temporales o secuenciales en los datos. Dentro de las RNNs se encuentran los modelos *Long Short-Term Memory* (LSTM), los cuales están diseñados para abordar el problema del desvanecimiento de la gradiente³ que tienen las RNNs convencionales. Las LSTM emplean una estructura interna basada en compuertas que regulan la información que se retiene, actualiza o descarta a medida que se procesan las secuencias, lo que les permite recordar dependencias a largo plazo. (DiPietro & D. Hager, 2020)

Para aprovechar tanto la información espacial como la temporal para segmentar secuencias de imágenes, una posibilidad sería combinar la arquitectura UNet y LSTM en un modelo híbrido. De esta manera el modelo UNet actuaría como extractor de características espaciales, mientras que las capas LSTM procesarían la información temporal extraída de las secuencias. Sin embargo, la implementación de un modelo

³ El desvanecimiento del gradiente ocurre cuando los gradientes se vuelven muy pequeños durante la retropropagación, dificultando el aprendizaje de dependencias a largo plazo en secuencias. Esto afecta especialmente a las RNNs, ya que pierden información relevante de pasos anteriores.

híbrido de esta naturaleza añadiría mucha complejidad a un *framework* diseñado para segmentar estructuras simples en imágenes 2D PC-MRI.

Una alternativa más práctica sería utilizar solo la arquitectura UNet, pero integrando la información espacial y temporal en una entrada única. Esto se puede conseguir si se logra transformar la información temporal de la secuencia en información espacial, permitiendo que la UNet procese toda la información en un solo paso sin necesidad de incorporar entradas secuenciales.

Para lograr transformar la información temporal de las secuencias en información espacial, la propuesta consistiría en reorganizar la secuencia completa de imágenes 2D PC-MRI en un solo volumen tridimensional, donde la tercera dimensión estaría conformada por las imágenes individuales de la secuencia. Este proceso sería equivalente al de apilar un conjunto de cartas dispersas para formar un mazo.

En la Figura 28, se tiene una secuencia de imágenes de tamaño L , donde cada imagen tiene una resolución de $W \times H$ píxeles en escala de grises, al combinar todas las imágenes en una tercera dimensión, el resultado sería un volumen de dimensión $L \times W \times H$. Aquí, L (*length*) corresponde al número de imágenes de la secuencia original, mientras que W (*width*) y H (*height*) son el ancho y alto de cada imagen respectivamente.

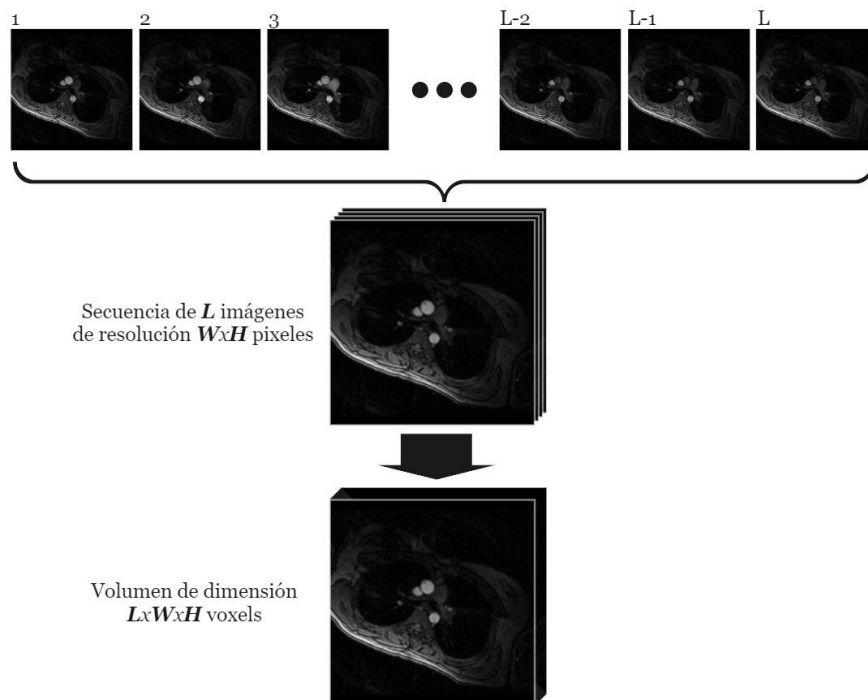


Figura 28: Conversión de una secuencia de L imágenes de resonancia magnética de resolución $W \times H$ píxeles, en un volumen de dimensión $L \times W \times H$. Fuente: Elaboración propia.

Desde esta perspectiva, segmentar una ROI en este volumen tridimensional equivale a segmentar cada imagen individual de la secuencia original (Figura 29). Además, este enfoque permite que la segmentación de cada imagen esté implícitamente influenciada

por la información temporal, la cual está integrada en la tercera dimensión del volumen. Esto permite mitigar la influencia de una imagen con mal contraste al tener disponible la información de la ROI durante las otras fases cardiacas dentro de este volumen.

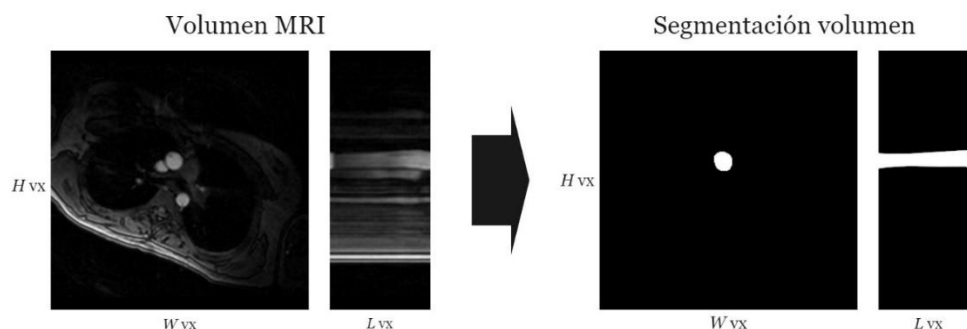


Figura 29: Segmentación de la aorta ascendente para un volumen MRI con dimensiones $L \times W \times H$. Se presenta la vista frontal del volumen junto con un corte sagital en el índice 124. Fuente: Elaboración Propia.

Entonces, en lugar de segmentar imágenes individuales en dos dimensiones, lo mejor es segmentar un volumen de tres dimensiones. Permitiendo de esta manera aprovechar la información espacial y temporal de la secuencia de imágenes 2D PC-MRI, logrando mitigar los problemas de segmentación derivados de un mal contraste de imagen durante determinadas fases cardiacas en la secuencia. Sin embargo, el modelo UNet clásico sigue requiriendo como entrada imágenes en dos dimensiones, por lo que aún se necesita una modificación de esta arquitectura para que pueda procesar volúmenes.

3.2.3 3D UNET

Una 3D Unet es una variante de la red neuronal Unet, diseñada específicamente para procesar datos tridimensionales, Al igual que la Unet clásica, la 3D Unet tiene una estructura de red en forma de "U", donde la parte descendente de la red reduce la dimensionalidad espacial utilizando convoluciones y agrupamientos (*pooling*) para extraer características relevantes, mientras que la parte ascendente restaura la dimensionalidad original utilizando capas de convolución transpuesta (o *upsampling*) para reconstruir la información espacial. La clave es que en la 3D Unet, las operaciones de convolución, agrupamiento y *upsampling* se realizan en 3D. (Çiçek, et al., 2016)

3.2.4 IMPLEMENTACION DE MODELO 3D UNET

En este trabajo, la implementación del modelo 3D UNet será una arquitectura estándar con 4 niveles de profundidad para el *encoder* y *decoder*. Cada nivel del *encoder* constará de dos capas convolucionales 3D con *kernels* de tamaño $3 \times 3 \times 3$ y funciones de activación ReLU. Posteriormente, se aplicará una capa de *MaxPooling* 3D con un tamaño de ventana de $2 \times 2 \times 2$ para reducir la dimensionalidad y aumentar la profundidad de las características extraídas. Entre el *encoder* y el *decoder* se implementará una capa "*bridge*" que actúa como un cuello de botella. En el *decoder*, se utilizarán capas de *Conv3DTranspose* para restaurar la resolución espacial, seguidas por la concatenación de las características del

encoder a través de *skip connections*. En la capa final, se utilizará una convolución 3D con *kernel 1x1x1* para reducir el número de canales de salida, dependiendo de si la tarea es segmentación binaria o multclasificación (Figura 30, Tabla 2).

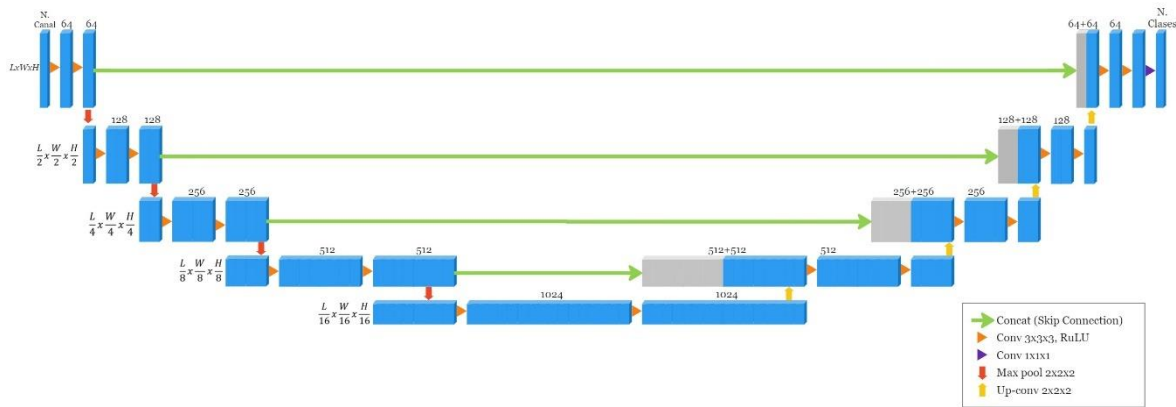


Figura 30: Arquitectura del modelo 3D Unet propuesto. Fuente: Elaboración propia.

Tabla 2: Cambio dimensional a lo largo del modelo 3D Unet. Fuente: Elaboración propia.

Bloque	Dimensiones	Skip Connection
Entrada	(L, W, H, n_canales)	
Encoder Block 1	(L/2, W/2, H/2, 64)	(D, H, W, 64)
Encoder Block 2	(L/4, W/4, H/4, 128)	(D/2, H/2, W/2, 128)
Encoder Block 3	(L/8, W/8, H/8, 256)	(D/4, H/4, W/4, 256)
Encoder Block 4	(L/16, W/16, H/16, 512)	(D/8, H/8, W/8, 512)
Bridge	(L/16, W/16, H/16, 1024)	
Decoder Block 1	(L/8, W/8, H/8, 512)	
Decoder Block 2	(L/4, W/4, H/4, 256)	
Decoder Block 3	(L/2, W/2, H/2, 128)	
Decoder Block 4	(L, W, H, 64)	
Salida Final	(L, W, H, n_clases)	

3.3 ENTRENAMIENTO DEL MODELO

Para entrenar el modelo 3D UNet es necesario que todos los volúmenes tengan una dimensión fija. Esto significa que no es posible entrenar el modelo con volúmenes de dimensión variable dentro de un mismo conjunto (64x256x256, 32x512x512, 128x64x64, etc.). Debido a esto, es necesario definir un valor fijo para las dimensiones **LxWxH** de los volúmenes, donde **L** representa a la cantidad de imágenes de la secuencia, y **WxH** a la resolución de las imágenes. Por ejemplo, si se establecen las dimensiones en 64x256x256, esto implica que todas las secuencias deben contener exactamente 64 imágenes, y cada imagen individual debe tener una resolución de 256x256 píxeles (esto trae consigo un problema el cual se aborda en la siguiente sección).

La dimensión que se establezca depende del contexto en el que se planea aplicar el modelo de segmentación. Incrementar los valores de **W** y **H** proporciona un mayor detalle espacial al modelo, mientras que aumentar el valor de **L** incrementa la cantidad de información temporal. Sin embargo, incrementar cualquiera de estas dimensiones significa un mayor costo en términos de memoria y requiere un conjunto de datos más amplio para garantizar un entrenamiento efectivo. Por lo que es importante **priorizar** que información es más **valiosa** para la segmentación. Este trabajo está enfocado en segmentar vasos sanguíneos, por lo tanto, la información espacial no es tan valiosa como la información temporal, debido a que la forma estructural de un vaso es simple y no requiere mayor detalle, en cambio, se requiere mayor información temporal para mantener la figura a lo largo de la segmentación.

Si bien es posible entrenar diferentes modelos 3D UNet con diferentes dimensiones para los volúmenes, se deben imponer algunas restricciones para los valores que **L**, **W** y **H** pueden tomar, ya que no todas las combinaciones son compatibles con el modelo 3D UNet. En particular, estas dimensiones deben cumplir dos requisitos fundamentales: ser divisibles por **potencias de 2** y tener un **valor mínimo** que evite el colapso de las características durante el proceso de reducción de dimensiones.

El primer requisito, que las dimensiones sean **potencias de 2**, proviene del uso de operaciones como *MaxPooling3D*, que reduce las dimensiones espaciales de los volúmenes a la mitad en cada capa de *pooling*. Esta reducción ocurre de manera secuencial, por lo que, si alguna de las dimensiones no es divisible por 2, se producirán errores al aplicar el *pooling*, y el tamaño de los tensores en las capas profundas de la red no sería válido. Además, el proceso de *upsampling* que ocurre en la fase de decodificación también requiere que las dimensiones sean simétricas para poder restaurar adecuadamente el tamaño original del volumen. Esto es importante en las *skip connections*, donde los tensores de la fase de codificación se concatenan con los de la fase de decodificación.

El segundo requisito es que **L**, **W** y **H** tengan un **valor mínimo** adecuado. En una 3D UNet típica, se aplican múltiples niveles de *pooling* (generalmente 4 o más), lo que reduce las dimensiones a la mitad en cada nivel. Por ello, las dimensiones de entrada deben ser lo

suficientemente grandes como para que, después de cada reducción por *pooling*, las dimensiones no colapsen a menos de 1. Por lo tanto, la resolución mínima requerida en cada dimensión está dada por 2^n , donde n es el número de niveles de *pooling*. En este caso, el valor mínimo para L , W y H sería 16, ya que $2^4=16$. Las dimensiones que sean menores a 16, en cualquiera de las dimensiones espaciales, provocarán que el entrenamiento falle.

Dado lo anterior, para garantizar el correcto entrenamiento del modelo 3D UNet, las dimensiones L , W y H deben ser **potencias de 2** y, para la implementación propuesta en este trabajo, el valor mínimo debe ser de **16 píxeles** en cada una de las dimensiones espaciales.

3.4 PROCESAMIENTO DE LAS IMÁGENES DE ENTRADA

Un problema que surge al establecer valores fijos para las dimensiones de los volúmenes de entrada al modelo 3D UNet, es que las secuencias de imágenes 2D PC-MRI pueden tener una cantidad variable de imágenes y resoluciones, lo que hace imposible crear directamente los volúmenes con la dimensión requerida.

Si la cantidad de imágenes en la secuencia es mayor a L , entonces, solo se podrá extraer una cantidad L de imágenes para generar el volumen, lo que implica que las imágenes fuera del volumen no serán segmentadas por el modelo. Por otro lado, si la secuencia contiene menos de L imágenes, no será posible construir un volumen válido. Lo mismo sucede si la dimensión $W \times H$ no coincide con la resolución de las imágenes. Por lo tanto, para garantizar la segmentación completa de las secuencias 2D PC-MRI, es necesario realizar un preprocesamiento que adapte las secuencias de imágenes a las dimensiones L , W y H del modelo 3D UNet.

Para las dimensiones W y H , que corresponden a la resolución de cada imagen individual en la secuencia, pueden ajustarse fácilmente al modelo 3D UNet aplicando un **reescalado lineal**. Con este método simplemente se toma cada imagen en la secuencia y se reescala su resolución para adaptarlas a los valores de W y H .

Por otro lado, el ajuste al valor L es más complejo de manejar, ya que es necesario determinar un procedimiento adecuado para abordar tanto el exceso como la falta de imágenes en la secuencia, a modo de garantizar que cada imagen entre en un volumen que pueda ser segmentado por el modelo 3D UNet, manteniendo al mismo tiempo las características espaciales y temporales que son necesarias para la propuesta.

3.4.1 EXTRACCIÓN SECUENCIAL

Una posible solución sería realizar una extracción secuencial de imágenes dentro de la secuencia 2D PC-MRI, donde con cada extracción se crea un volumen individual. Los pasos a seguir para este método se definen a continuación:

1. Extraer la cantidad L de imágenes de la secuencia. Si luego de la extracción aún quedan imágenes restantes, extraer las L siguientes.
2. Si el número de imágenes restantes es **insuficiente** para completar la cantidad L , se debe rellenar con las primeras imágenes de la misma secuencia (asumiendo que la secuencia 2D PC-MRI es cíclica; de lo contrario, es recomendado completar repitiendo la última imagen).

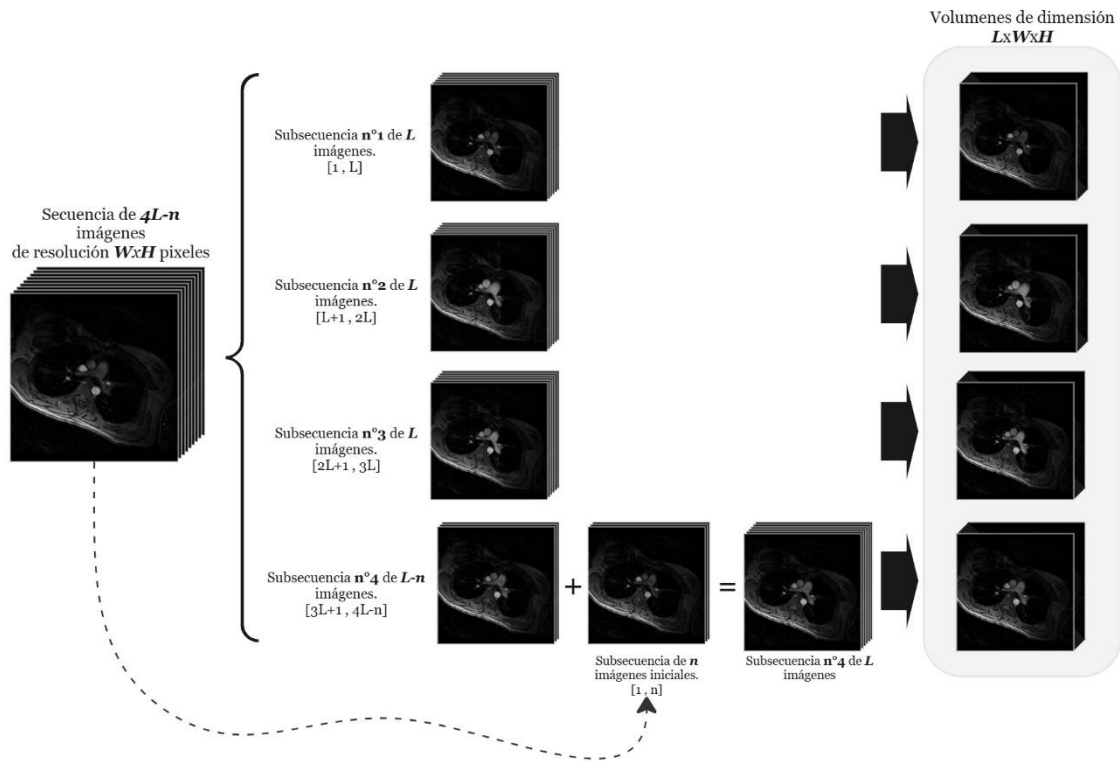


Figura 31: Creación de volúmenes a través del método de extracción secuencial. Fuente: Elaboración propia.

Para aclarar mejor la idea, en la Figura 31, se muestra una secuencia de imágenes 2D PC-MRI la cual contiene una cantidad de $4L-n$ imágenes a una resolución $W \times H$ cada una (todas las variables positivas mayor a cero, donde $L > n$). Siguiendo el método de extracción secuencial para crear volúmenes, se aplica el **Paso 1** para extraer las primeras L imágenes con las cuales se crea el primer volumen de dimensión $L \times W \times H$. Al quedar imágenes disponibles, se repite el proceso y se extraen los L siguientes para crear el segundo volumen, y luego se aplica nuevamente para crear el tercer volumen. Al quedar $L-n$ imágenes disponibles no es posible crear un volumen válido, por lo que se pasa al **Paso 2**, tomando la cantidad faltante de n imágenes, correspondientes al inicio de la secuencia principal, añadiéndolas a la subsecuencia n°4 lo que permite completar el cuarto volumen, logrando de esta manera tener de todas las imágenes de la secuencia en un volumen que puede ser segmentado por el modelo 3D UNet.

Con este procedimiento, es posible extraer múltiples volúmenes con las dimensiones requeridas a partir de una sola secuencia de imágenes 2D PC-MRI. Cada volumen

abarcaría un subconjunto de imágenes que deberá ser segmentado individualmente por el modelo.

3.4.2 EXTRACCIÓN RECURSIVA

La extracción secuencial es un buen método para abordar el problema dimensional en secuencias 2D PC-MRI que contienen un número de imágenes **cercano** al valor de **L** establecido para el modelo 3D UNet. Sin embargo, si una secuencia 2D PC-MRI tiene una cantidad de imágenes significativamente mayores al valor **L**, se puede dar la situación que los volúmenes extraídos solo incluyan imágenes de mal contraste. Esto presenta un problema para la propuesta, ya que el objetivo es que los volúmenes incluyan información temporal de las imágenes en momentos con mayor nitidez para amortiguar los momentos de baja calidad de imagen.

Para afrontar este problema, otra alternativa para generar los volúmenes sería extrayendo muestras de imágenes a lo largo de toda la secuencia, dividiendo la secuencia original en subsecuencias conformadas por las imágenes en posiciones pares e impares. La idea de este método es obtener subsecuencias que logren mantener imágenes de todas las fases registradas en la secuencia original, manteniendo la información temporal lo más diversa posible. Por lo tanto, los pasos a seguir son los siguientes:

1. Se extraen las imágenes de las posiciones pares e impares de la secuencia original y se crean dos subsecuencias. Se comprueba que todas las subsecuencias creadas tengan un largo menor o igual al valor **L**. Si se cumple el criterio se pasa al paso 2, caso contrario se repite el paso para cada subsecuencia que haya sido creada.
2. Al llegar al criterio de parada, se comprueba que cada subsecuencia creada tenga la cantidad **L** de imágenes, para las subsecuencias que tengan una cantidad de imágenes inferior a **L**, se completan con las imágenes iniciales de la **misma subsecuencia** hasta llegar a **L** (similar a la extracción secuencial).

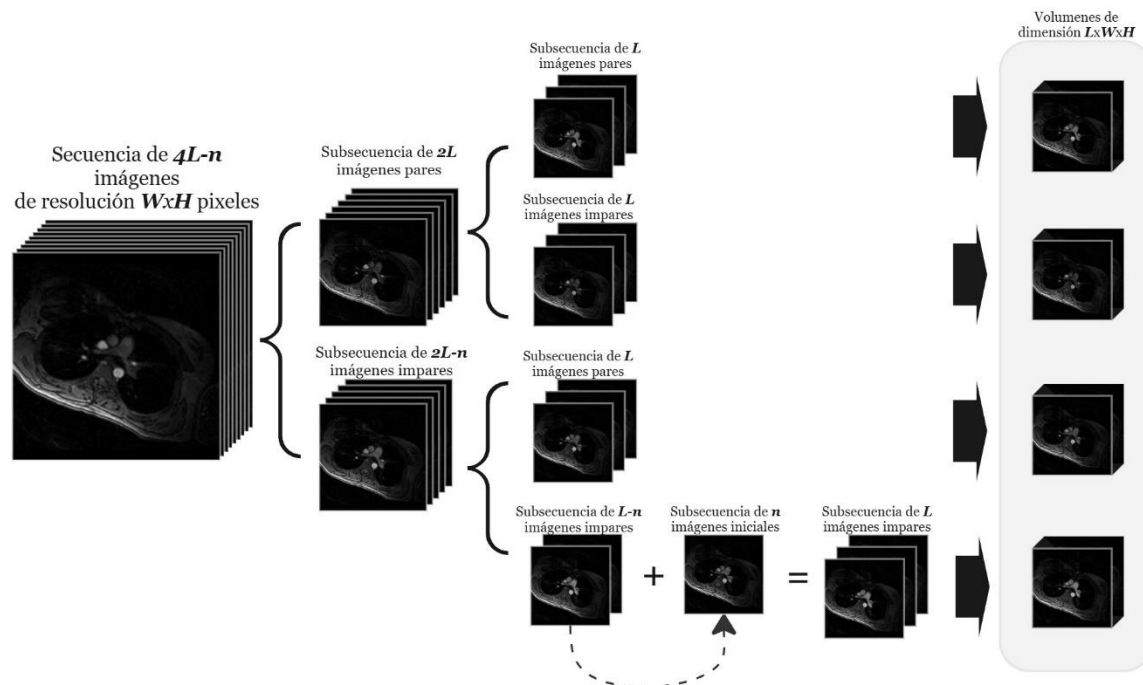


Figura 32: Creación de volúmenes a través del método de extracción recursiva. Fuente: Elaboración propia.

Como ejemplo, en la figura 32 se tiene la misma secuencia 2D PC-MRI vista en la extracción secuencial, con $4L-n$ imágenes de resolución $W \times H$. Durante el **Paso 1**, la secuencia original se divide en dos subsecuencias: la primera contiene $2L$ imágenes ubicadas en las posiciones pares, mientras que la segunda incluye $2L-n$ imágenes en las posiciones impares. Dado que ambas subsecuencias tienen más de L imágenes, se aplica nuevamente el Paso 1, subdividiendo ambas subsecuencias en cuatro nuevas subsecuencias, quedándose con las imágenes en las posiciones pares e impares de sus respectivas subsecuencias padres. Esta vez, todas las subsecuencias creadas tienen un largo igual o menor a L , por lo que se procede al **Paso 2**. En este caso de ejemplo, solo una subsecuencia tiene una cantidad de imágenes inferior a L , por lo que se procede a completar con imágenes de la misma subsecuencia. Finalmente, se logran formar 4 volúmenes con las dimensiones compatibles con el modelo 3D UNet.

De este modo, se obtienen volúmenes que contienen información temporal de todas las fases cardiacas, mitigando así el problema asociado con la extracción secuencial.

3.5 REGIÓN DE INTERÉS

Hasta ahora, para la creación de los volúmenes utilizados en el entrenamiento de la 3D UNet, se ha empleado como referencia la imagen **completa** de la resonancia magnética. Sin embargo, este enfoque no es el más adecuado, ya que el objetivo del *framework* es permitir al usuario definir qué región específica de la PC-MRI debe ser segmentada. Al

entrenar la 3D UNet con la imagen completa, se le fuerza a la red a identificar y segmentar todas las estructuras presentes en la imagen, lo que no solo aumenta la complejidad del problema, sino que también limita la precisión de la segmentación en la ROI que realmente importa al especialista.

En lugar de utilizar la imagen completa, lo más adecuado es recortar un área rectangular que encierre únicamente la región de interés, que, para el contexto de este trabajo, sería el área donde se encuentra algún vaso sanguíneo. Para ello, es necesario extraer la región que incluya todo el vaso a lo largo de toda la secuencia de imágenes.

El extraer estas regiones generan nuevas secuencias de imágenes 2D PC-MRI con resoluciones diferentes a la original, pero con igual cantidad de imágenes. Estas secuencias tienen una información espacial y temporal más enfocada en la región de interés, lo que permite al modelo 3D UNet una mayor capacidad para aprender las características importantes de las estructuras anatómicas que se quieren segmentar.

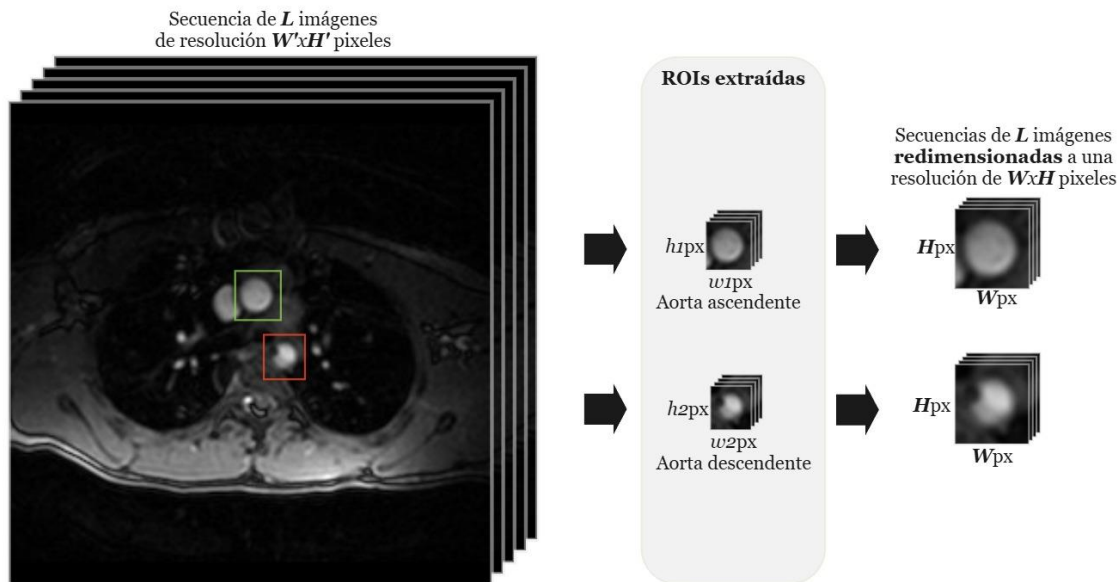


Figura 33: Extracción de las regiones de interés para el entrenamiento de la 3D UNet. Fuente: Elaboración Propia.

Para una mayor comprensión, en la Figura 33, se presenta una secuencia de imágenes 2D PC-MRI la cual contiene L imágenes con una resolución $W' \times H'$ ($W \leq W'$, $H \leq H'$; todas las variables mayores a cero). De estas imágenes se extraen dos áreas rectangulares las cuales encierran dos regiones de interés (aorta ascendente y descendente). Estas nuevas secuencias mantienen la misma cantidad L de imágenes, pero la resolución de cada imagen cambia. Estas nuevas secuencias 2D PC-MRI, al igual que con la secuencia 2D PC-MRI original, deben ser preprocesadas para que se puedan crear volúmenes compatibles con el modelo 3D UNet. Con relación al valor L , esta secuencia ya posee dicha cantidad, así que no se debe realizar ningún tipo de extracción. Con respecto a los valores W y H , se

aplica un reescalado lineal en ambas secuencias para ajustar cada imagen a la resolución **WxH**.

Todo el procesamiento visto en secciones anteriores para el ajuste dimensional de **LxWxH**, es aplicable a estas nuevas secuencias 2D PC-MRI.

3.6 MÁSCARA DE SALIDA

La salida del modelo 3D UNet consiste en un volumen tridimensional que representa la segmentación de la estructura dentro del volumen de entrada. La máscara volumétrica resultante tiene las mismas dimensiones que el volumen de entrada (Figura 34).

La máscara volumétrica se caracteriza por un fondo negro y una segmentación en blanco, donde el valor 0 representa el color negro y el valor 1 corresponde al blanco. De esta manera, el entrenamiento del modelo 3D UNet se centra en predecir si cada píxel (o voxel, en el caso de volúmenes 3D) pertenece a la estructura al interior de la ROI. En este contexto, los valores 1 y 0 indicarán, respectivamente, la pertenencia y la no pertenencia a la estructura.

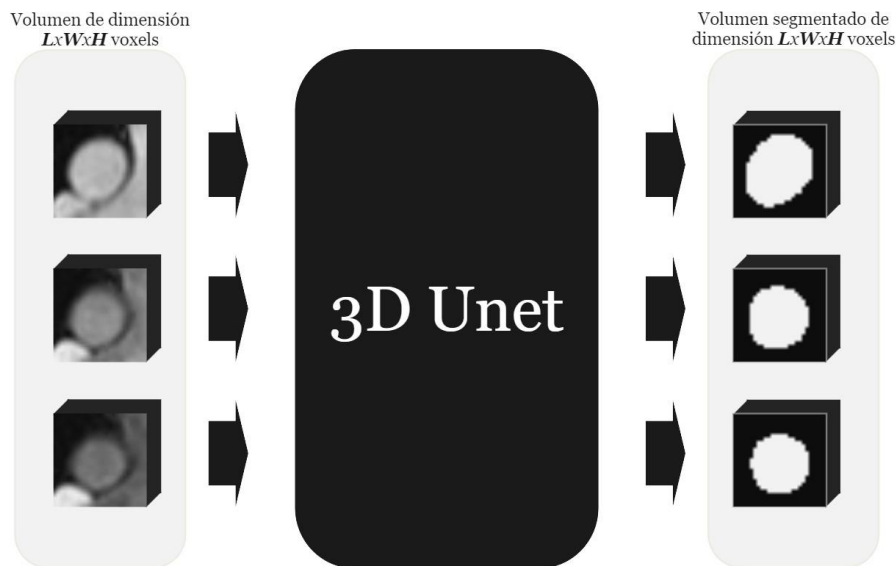


Figura 34: Entrenamiento de la 3D U-net, el input corresponde a los volúmenes creados y la salida una máscara segmentando la estructura al interior de la región de interés. En el ejemplo, la estructura es la aorta ascendente. Fuente: Elaboración propia.

Para guiar el entrenamiento del modelo, se empleará la función de pérdida *Dice Loss*, que evalúa la similitud entre la segmentación predicha y la segmentación real (ver sección 2.4.1).

3.7 PROCESAMIENTO DE LA SEGMENTACIÓN

Una vez segmentados los volúmenes utilizando el modelo 3D UNet, es necesario procesar el resultado y posicionar la segmentación en la secuencia de imágenes 2D PC-MRI original. Para ello, se deben descomponer las máscaras volumétricas de dimensión $L \times W \times H$ en máscaras en 2D, estas máscaras se deben unificar según el orden con el que fueron creados los volúmenes, esto dependerá del método utilizado (**secuencial** o **recursivo**) (Figura 35).

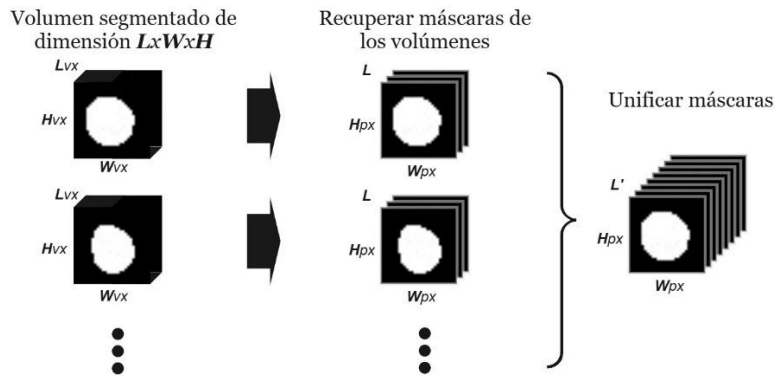


Figura 35: Extracción y unificación de las máscaras desde los volúmenes segmentados. Fuente: Elaboración propia.

En este paso, las máscaras aún mantienen una resolución de $W \times H$, que corresponde a la dimensión establecida para el modelo 3D UNet. Sin embargo, para que la segmentación se posicione correctamente en la secuencia de imágenes original, es necesario restaurar cada máscara a su resolución original de extracción. Finalmente, cada máscara debe ubicarse en su posición correspondiente dentro de la imagen original de la secuencia 2D PC-MRI, completando así el proceso de segmentación (Figura 36).

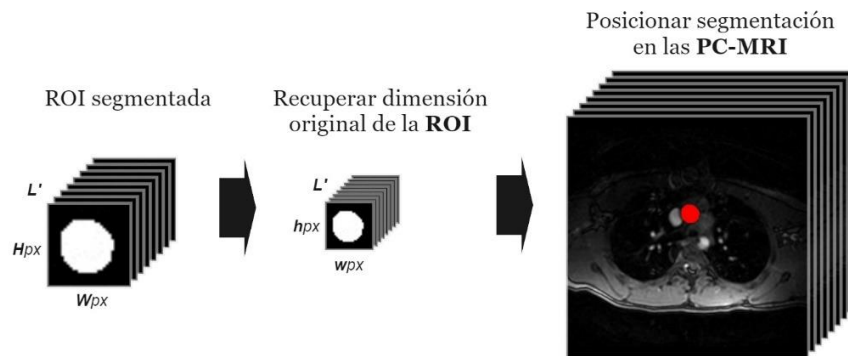


Figura 36: Recuperación de la resolución original de la ROI para luego ser posicionada en la secuencia original PC-MRI. Fuente: Elaboración propia.

3.8 IMAGEN DE FASE

A lo largo del desarrollo del método de segmentación, se ha estado utilizando únicamente la **imagen de magnitud** de una 2D PC-MRI, ya que resulta más fácil de interpretar para el ojo humano, sin considerar la **imagen de fase**. Sin embargo, la imagen de fase proporciona información valiosa sobre el flujo sanguíneo en los vasos durante la secuencia, lo que significa que su inclusión en el método podría ofrecer al modelo información relevante sobre la estructura arterial en situaciones donde la imagen de magnitud no sea suficiente.

Si se considera la representación dimensional de una imagen RGB (Red, Green, Blue) de resolución 512×512 , entonces su dimensión a nivel computacional será de $512 \times 512 \times 3$, donde la última dimensión tiene un valor 3 correspondiente a los canales RGB, los cuales almacenan la ponderación de los tres colores en cada pixel. Si se aplica la misma lógica, se puede combinar la imagen de **Magnitud** y la de **Fase** en una sola imagen de **2 canales**. Al ser ambas imágenes en escala de grises, esto permite fusionarlas quedando con una dimensión de $W \times H \times 2$, donde cada canal almacena el valor original del pixel de cada tipo de imagen. Aunque esta representación no tendrá un significado visual claro, permitirá entrenar el modelo 3D UNet utilizando la información de ambas imágenes (Figura 37).

Con esta nueva configuración, los volúmenes destinados al entrenamiento del modelo 3D UNet tendrán una dimensión $L \times W \times H \times 2$. Esto no representa un inconveniente para el modelo, ya que la dimensión de los canales ya está contemplada en la entrada del 3D UNet.

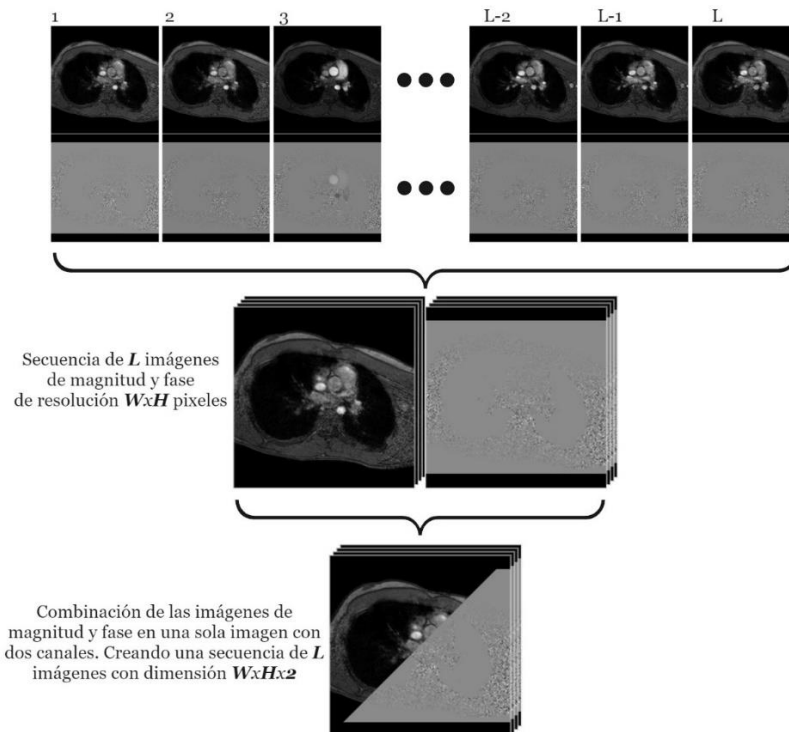


Figura 37: Combinación de las imágenes de Magnitud y Fase en una sola imagen con 2 canales. Fuente: Elaboración propia.

3.9 FRAMEWORK DE SEGMENTACIÓN

Entendido todo lo explicado en las secciones anteriores, el *framework* de segmentación basado en la segmentación de volúmenes con el modelo 3D Unet se puede estructurar en cinco procedimientos:



Figura 38: Los 5 procedimientos para el *framework* de segmentación. Fuente: Elaboración propia.

- **Procesamiento de Imágenes:** En esta fase, se procesan las imágenes para convertirlas al formato adecuado. Para el contexto de este trabajo, se trata de procesar archivos DICOM, extraer de ellos las imágenes de Magnitud y Fase y normalizar los valores de los píxeles a un rango entre 0 a 1. Posteriormente, las imágenes de Magnitud y Fase se combinan en imágenes con 2 canales, donde cada canal representa los valores de los píxeles de Magnitud y Fase.
- **Extracción de la ROI:** Este paso implica la identificación y extracción de una región de interés específica dentro de la secuencia de imágenes 2D PC-MRI para segmentarla. Esta región puede abarcar una pequeña estructura dentro de la imagen o incluir la imagen completa. En el contexto de este trabajo, la ROI extraída corresponde al área que encierra un vaso sanguíneo.
- **Creación de los Volúmenes:** En esta fase, se generan volúmenes a partir de la secuencia de imágenes con las dimensiones compatibles con el modelo 3D UNet, ajustando las dimensiones a los valores WxH mediante un reescalado lineal y, para el valor L, se emplea uno de los dos métodos propuestos: recursivo o secuencial.
- **Segmentación con 3D Unet:** En este paso, se reciben los volúmenes y se segmentan utilizando el modelo 3D UNet. El modelo debe estar entrenado con los volúmenes que contengan la información espacial y temporal de la estructura deseada.
- **Procesamiento de la segmentación:** Este es el último paso y consiste en procesar los volúmenes segmentados, posicionando la segmentación en la secuencia de imágenes original.

En resumen, el *framework* de segmentación de secuencias de PC-MRI que se ha presentado se resume en la figura 39. Esta figura muestra todos los pasos necesarios para segmentar una secuencia de imágenes 2D PC-MRI, organizando cada procedimiento según la fase correspondiente en el *framework*.

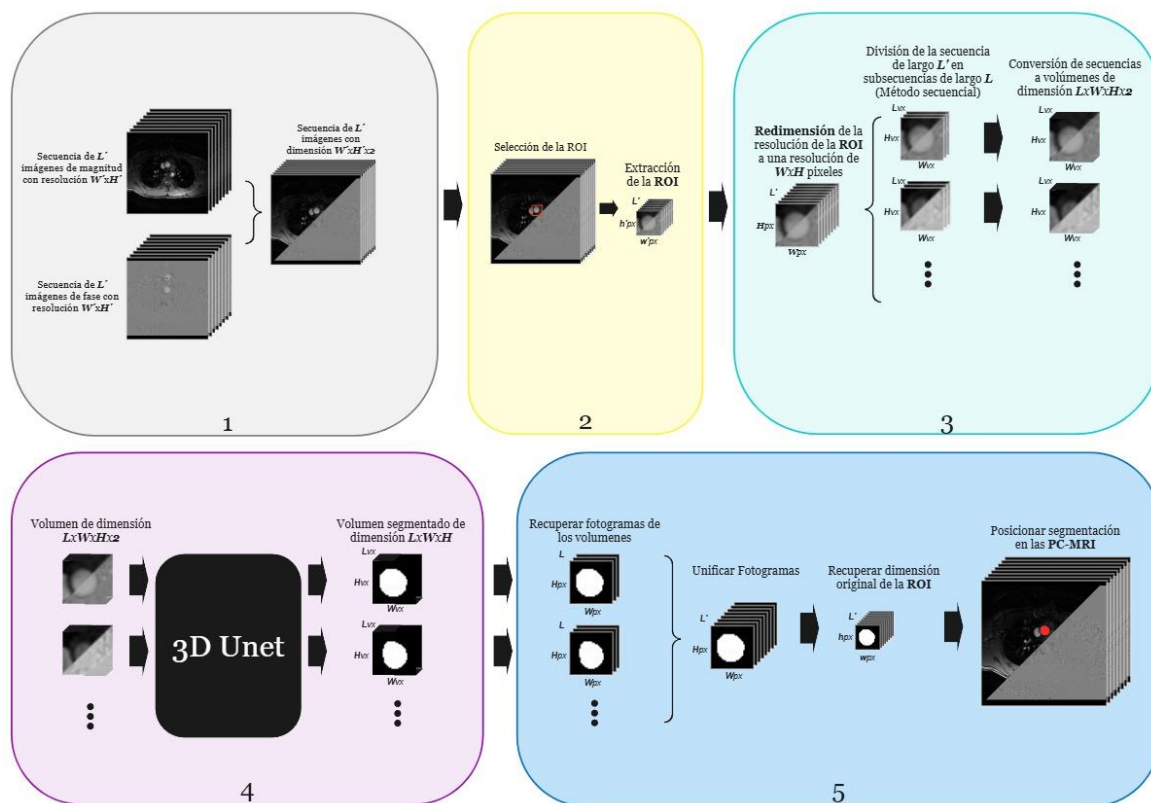


Figura 39: Framework de segmentación de secuencias 2D PC-MRI utilizando el modelo 3D U-Net. Fuente: Elaboración propia.

En la figura se muestra la implementación presentada en este trabajo para segmentar vasos sanguíneos. Sin embargo, este *framework* puede adaptarse para segmentar otras estructuras anatómicas, cuya información temporal sea un factor relevante para su correcta segmentación.

CAPÍTULO 4: VALIDACIÓN DE LA SOLUCIÓN

En este capítulo, se evaluará el modelo 3D Unet propuesto. Para ello, se comparará su desempeño con un algoritmo de segmentación basado en contorno activo⁴ y con una segmentación manual de las imágenes, los cuales servirán como referencia para el análisis comparativo. Las pruebas evaluarán cualitativa y cuantitativamente la segmentación en diferentes secuencias PC-MRI.

Las pruebas se ejecutarán en un equipo con las siguientes características técnicas:

Tabla 3: Características técnicas del equipo que ejecutará las pruebas. Fuente: Elaboración propia.

Hardware	Descripción
CPU	AMD Ryzen 7 7800X3D 8-Core
GPU	NVIDIA GeForce RTX 4070
RAM	32GB
VRAM	12GB (Limitado a 3GB)

El modelo 3D UNet estará entrenado para volúmenes con dimensión **128x32x32**. De esta manera se prioriza la dimensión **L** para obtener mayor información temporal de las secuencias 2D PC-MRI.

4.1 METODO EXPERIMENTAL

El método experimental consistirá en tres secciones. En la primera, se analizará el desempeño del modelo 3D Unet en dos casos específicos, donde se sabe de antemano que el método de segmentación basado en contornos activos falla al segmentar al segmentar correctamente el vaso sanguíneo a lo largo del ciclo cardiaco. En la segunda sección, el modelo se evaluará en un conjunto más amplio de casos, con el objetivo de evaluar su precisión para un uso generalizado. Finalmente, en la tercera sección, se analizarán los tiempos de predicción del modelo 3D Unet.

Para todos los casos, se calculará el coeficiente de *Dice* para cuantificar la cercanía entre la segmentación realizada por el modelo propuesto, el método basado en contorno activo y la segmentación manual. Junto con ello, se evaluará de manera cualitativa el área y flujo

⁴ El método de contorno activo utilizado se encuentra en el siguiente enlace <https://github.com/Cardiac-MR-Group-Lund/segment-open>

calculados a lo largo de la secuencia de imágenes. Además, el modelo 3D Unet se evaluará utilizando los enfoques de creación de volúmenes **secuencial** y **recursivo** (ver sección 3.4).

4.2 CASOS ESPECÍFICOS

4.2.1 PRIMER CASO

Para el primer caso, se tiene un corte transversal donde se observa el flujo en las aortas ascendente y descendente durante las fases de sístole y diástole.

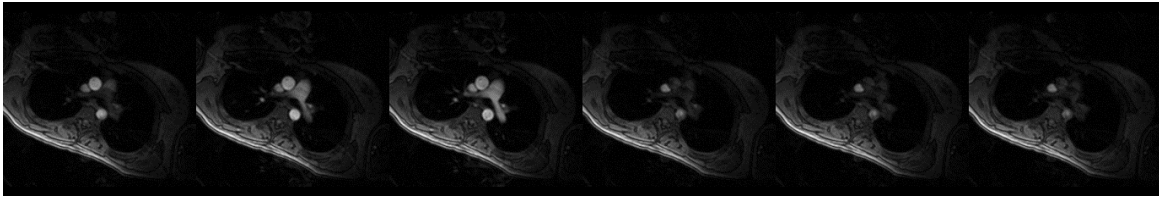


Figura 40: (Caso 1) Muestra de imágenes de Magnitud, donde se observan las aortas ascendente y descendente a lo largo del ciclo cardíaco (las primeras 3 corresponden a la fase sístole y las otras a la diástole). Fuente: Elaboración propia.

En la Figura 40 se observa como el contorno de las arterias está bien definido durante la fase de sístole, pero al entrar en fase diástole, los contornos se difuminan con el fondo. El problema surge al segmentar la aorta ascendente con un método basado en contorno activo, que durante la fase diastólica pierde el contorno de la arteria, provocando una segmentación irregular en dicho periodo.

La Figura 41 presenta una muestra de la segmentación utilizando el método manual, contorno activo y la propuesta 3D Unet recursivo y secuencial:

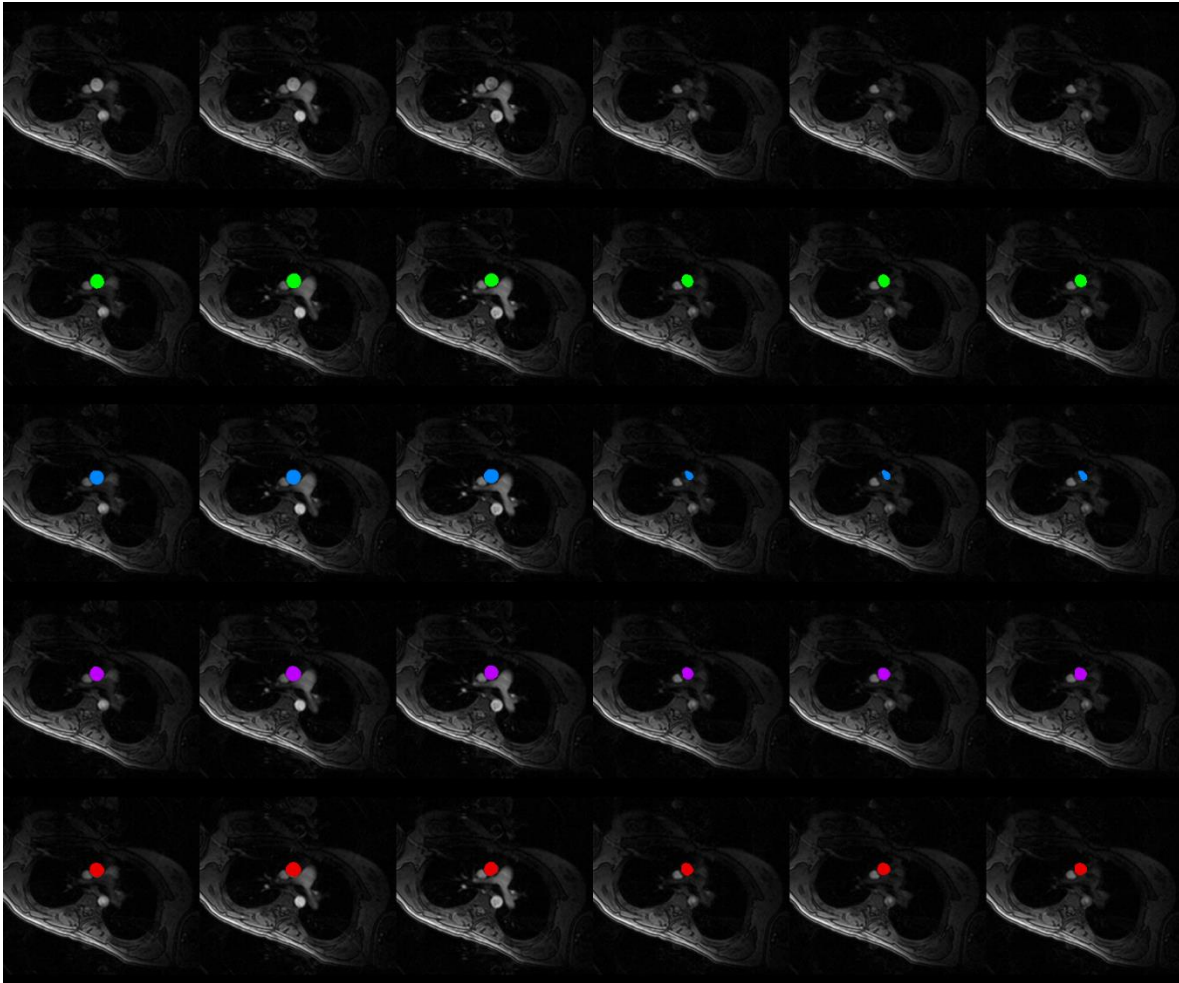


Figura 41: (Caso 1) Segmentación de la aorta ascendente utilizando los métodos de segmentación manual (Verde), contorno activo (Azul) y el modelo propuesto 3D U-Net Recursivo (Púrpura) y Secuencial (Rojo).

Fuente: Elaboración propia.

Se observa que, durante la fase sistólica, los métodos de segmentación por contorno activo y el propuesto presentan resultados similares. Sin embargo, en la fase diástole, el método de contorno activo no logra preservar la forma del vaso, al ser comparado con la segmentación manual. En contraste, el método propuesto, basado en 3D U-Net recursivo y secuencial, mantiene una forma más cercana a la segmentación manual durante este periodo.

El siguiente gráfico muestra el área segmentada a lo largo de la secuencia PC-MRI, utilizando el método manual, contorno activo y la propuesta 3D U-Net recursiva y secuencial.

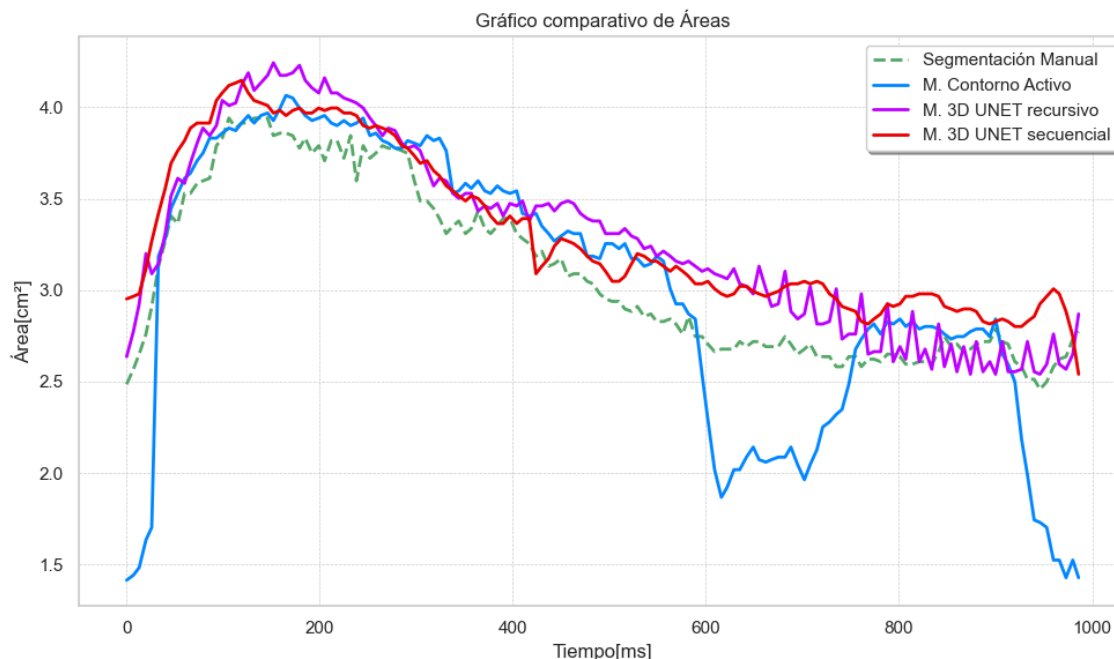


Figura 42: (Caso 1) Gráfico comparativo de las áreas utilizando la segmentación manual (**Verde**), el método de contorno activo (**Azul**), y la propuesta 3D Unet con los enfoques recursivos (**Púrpura**) y secuencial (**Rojo**).
 Fuente: Elaboración propia.

Del gráfico (Figura 42) se puede ver con mayor detalle el comportamiento del área a lo largo de la secuencia. Durante el periodo de 50[ms] a 500[ms], todos los métodos muestran resultados similares (se puede asociar dicho periodo a la fase sístole). Pero, a medida que avanza el tiempo (comienza la fase diástole), el método de contorno activo se vuelve inestable y colapsa en dos instancias, durante el periodo de 600 [ms] a 800 [ms] y luego durante el final de la secuencia a los 900 [ms], que se propaga al inicio durante los primeros 20 [ms] aproximadamente.

Por otro lado, el método propuesto 3D Unet con los enfoques recursivo y secuencial, muestran resultados más estable y consistente a lo largo de toda la secuencia. Al observar más en detalle ambos gráficos del método 3D Unet, se puede notar que la curva del enfoque **recursivo** presenta una forma de sierra, que es más evidente en el periodo de 600 [ms] hasta el final. Esto se debe a cómo este método procesa los volúmenes. Recordando, el enfoque recursivo divide la secuencia original en fotogramas pares e impares, que segmenta de forma individual. Al combinar los fotogramas segmentados, se genera una fluctuación, ya que las predicciones para ambas subdivisiones muestran áreas ligeramente diferentes, especialmente en fases donde los contornos de la estructura son más difusos.

Tabla 4: (Caso 1) Cálculo del coeficiente de *Dice* para las segmentaciones obtenidas por los métodos en estudio, siendo comparadas con la segmentación manual. Fuente: Elaboración propia.

Método	Coficiente de <i>Dice</i>
Contorno activo	0,82 ± 0,13
3D Unet Recursivo	0,94 ± 0,02
3D UNet Secuencial	0,92 ± 0,02

La Tabla 4 resume los resultados observados en los gráficos, mostrando el coeficiente de *Dice* para las segmentaciones obtenidas por los métodos comparados. El 3D Unet Recursivo es el más preciso y consistente con un valor de **0,94**, seguido por el 3D Unet Secuencial con un valor de **0,92**, ambos con una baja variabilidad de **0,02**. En contraste, el método de Contorno Activo presenta un desempeño inferior, con un coeficiente de *Dice* de **0,82** y una variabilidad considerable de **0,13**.

A continuación, se presenta el gráfico de flujos calculados a partir del área segmentada:

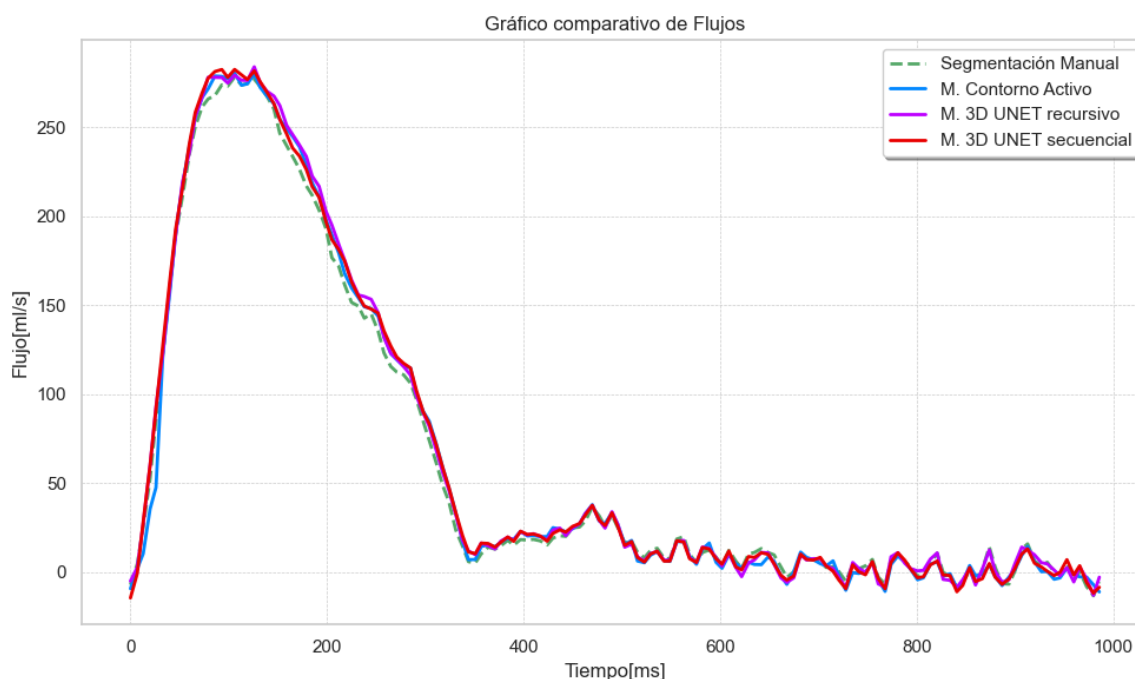


Figura 43: (Caso 1) Gráfico comparativo de los flujos calculados a partir de la segmentación obtenida por medio de la segmentación manual (Verde), el método de contorno activo (Azul), y la propuesta 3D UNET con los enfoques recursivos (Púrpura) y secuencial (Rojo). Fuente: Elaboración propia.

En la figura 43, se observa que todos los métodos muestran una curva de flujo idéntica. A pesar de que la segmentación del método de contorno activo colapsa en ciertos períodos de la secuencia, esto no impacta significativamente el cálculo del flujo, ya que ocurre durante un intervalo en el que el flujo es mínimo.

4.2.2 SEGUNDO CASO

En el segundo caso, se observa un corte transversal que muestra el flujo en las aortas ascendente y descendente durante las fases de sístole y diástole. A diferencia del primer caso, las imágenes mantienen mejor la nitidez en la fase de diástole, pero los contornos de la aorta descendente no están bien contrastados. Esto genera que, al aplicar el método de segmentación por contorno activo, se pierda temporalmente la delineación de la arteria, causando irregularidades en la segmentación.

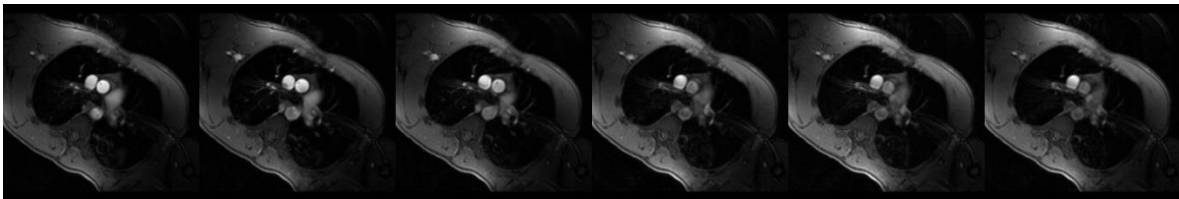


Figura 44: (Caso 2) Muestra de imágenes de Magnitud, donde se observan las aortas ascendente y descendente en las fases de sístole y diástole (las primeras 3 corresponden a la sístole y las otras a la diástole). Fuente: Elaboración propia.

La Figura 45 muestra una comparación entre la segmentación manual, el método de contorno activo y la propuesta basada en un 3D Unet recursivo y secuencial:

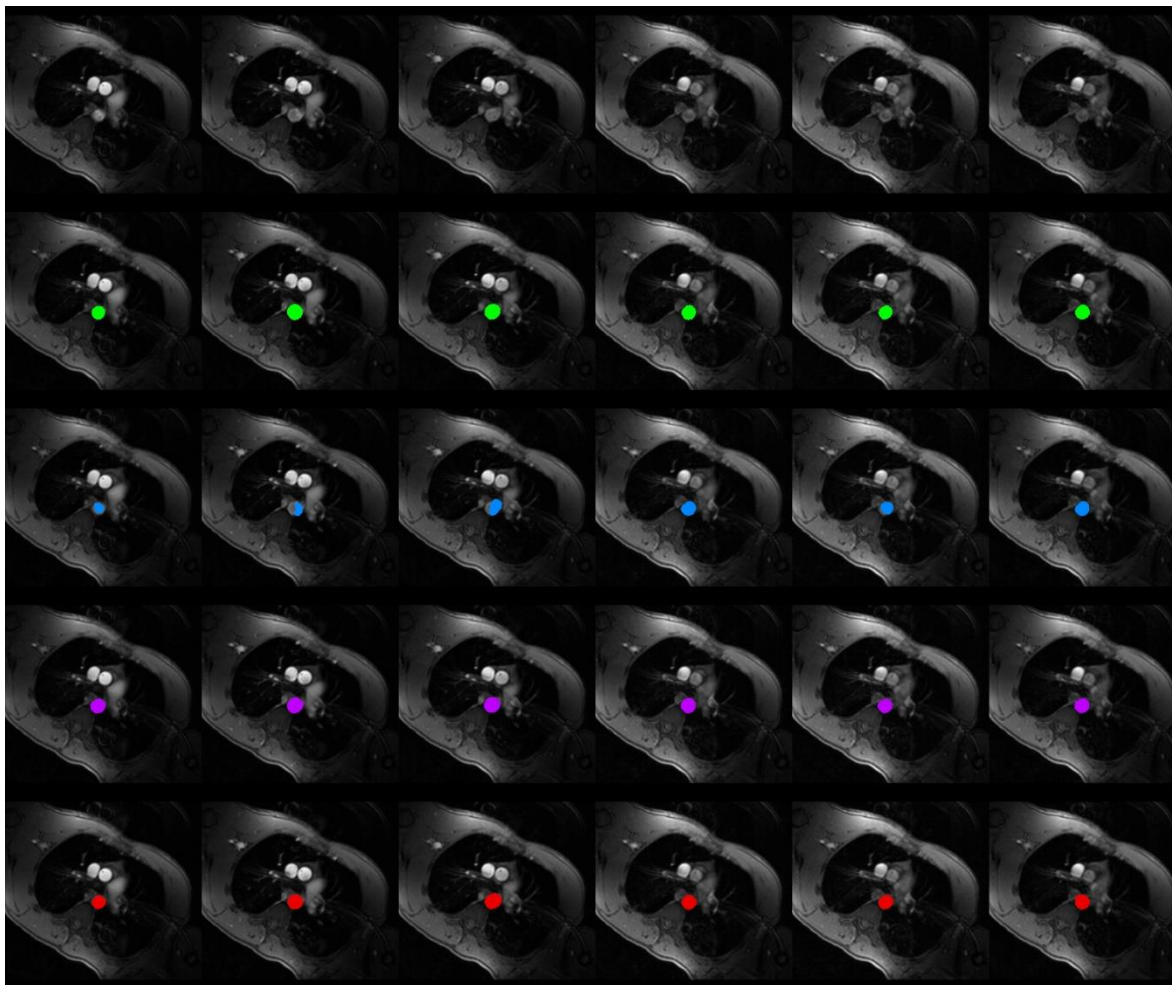


Figura 45: (Caso 2) Segmentación de la aorta ascendente utilizando los métodos de segmentación manual (Verde), contorno activo (Azul) y el modelo propuesto 3D Unet Recursivo (Púrpura) y Secuencial (Rojo).
Fuente: Elaboración propia.

En la imagen se observa cómo la segmentación por contorno activo pierde el borde de la aorta descendente durante la fase de sístole, mientras que en la fase de diástole conserva una segmentación similar a la realizada manualmente, siendo la situación contraria a lo visto en el caso 1. En cambio, el método basado en el modelo 3D UNet utilizando los enfoques recursivo y secuencial, logra una segmentación cercana a la manual en ambas fases.

La Figura 46 muestra el área segmentada a lo largo de la secuencia PC-MRI, comparando el método manual, el contorno activo y la propuesta del modelo 3D Unet en sus enfoques recursivo y secuencial.

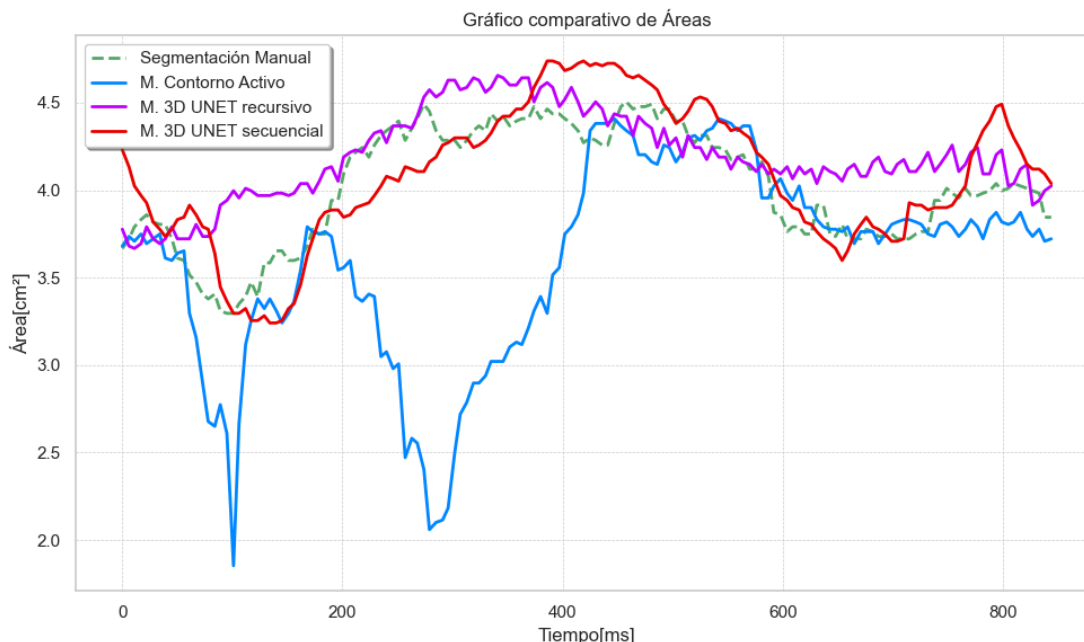


Figura 46: (Caso 2) Gráfico comparativo de las áreas utilizando la segmentación manual (**Verde**), el método de contorno activo (**Azul**), y la propuesta 3D Unet con los enfoques recursivos (**Púrpura**) y secuencial (**Rojo**).

Fuente: Elaboración propia.

En el gráfico se observa con mayor claridad la pérdida del contorno de la aorta descendente en el método de contorno activo, específicamente entre los 50 [ms] y 400 [ms]. Sin embargo, durante el resto de la secuencia, los resultados son similares a los de la segmentación manual. Por otro lado, el método 3D Unet, en ambos enfoques, muestra una cercanía con la curva de segmentación manual a lo largo de ambas fases cardíacas.

Similar al caso 1, el método 3D Unet con el enfoque recursivo, presenta la misma curva con forma de sierra, pero esta vez, tiene mayor presencia en ambas fases cardíacas. Esto se debe a que el contorno de la aorta descendente no está bien definido en gran parte de la secuencia, lo que provoca pequeñas variaciones en la segmentación de las subdivisiones.

Tabla 5: (Caso 2) Cálculo del coeficiente de *Dice* para las segmentaciones obtenidas por los métodos en estudio, siendo comparadas con la segmentación manual. Fuente: Elaboración propia

Método	Coefficiente de <i>Dice</i>
Contorno activo	0,86 ± 0,11
3D Unet Recursivo	0,95 ± 0,01
3D Unet Secuencial	0,96 ± 0,01

La Tabla 5 resume los resultados observados en el gráfico de áreas, mostrando el coeficiente de *Dice* para las segmentaciones obtenidas por los métodos comparados. En este análisis, el 3D Unet Secuencial se posiciona como el más preciso y consistente, con un valor de 0,96 y una baja variabilidad de 0,01. Le sigue de cerca el 3D Unet Recursivo, con un valor de 0,95 con la misma variabilidad. En contraste, el método de Contorno Activo muestra el peor desempeño, con un coeficiente de Dice de 0,86 y una alta variabilidad de 0,11.

La figura 47 presenta el gráfico de flujos calculados a partir del área segmentada:

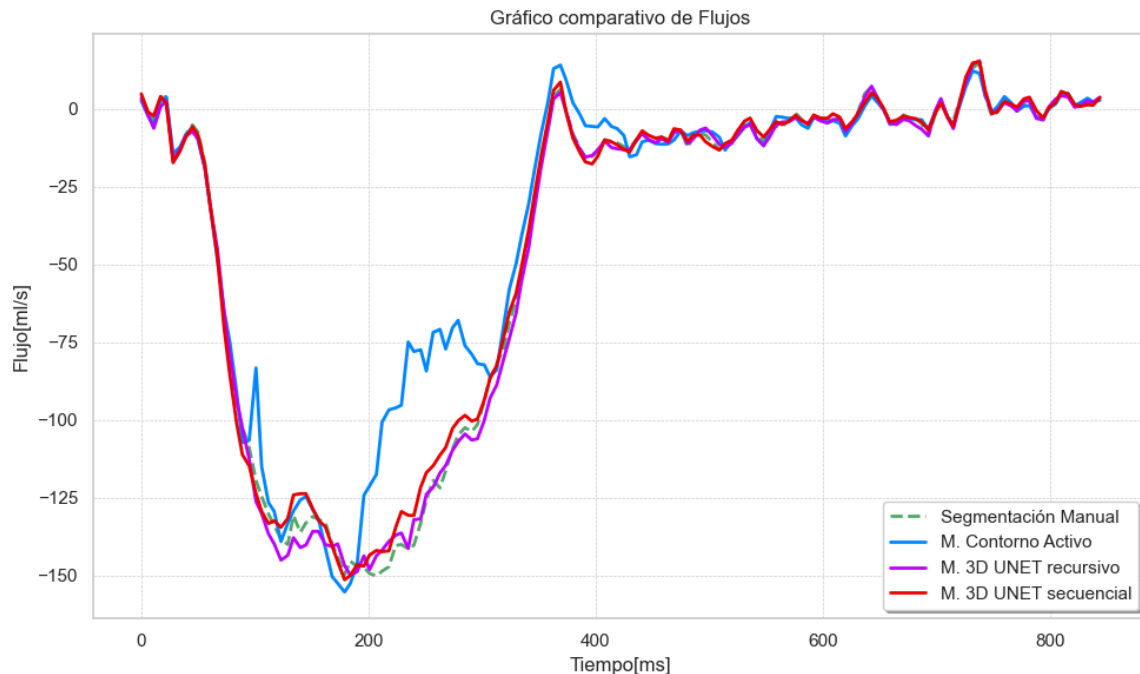


Figura 47: (Caso 2) Gráfico comparativo de los flujos calculados a partir de la segmentación obtenida por medio de la segmentación manual (**Verde**), el método de contorno activo (**Azul**), y la propuesta 3D UNET con los enfoques recursivos (**Púrpura**) y secuencial (**Rojo**). Fuente: Elaboración propia.

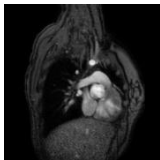
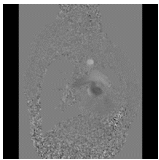
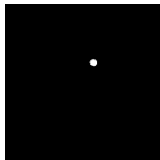
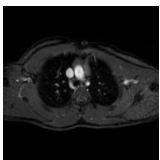
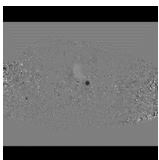
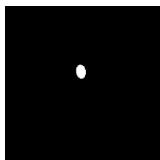
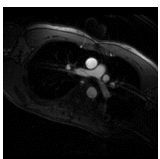
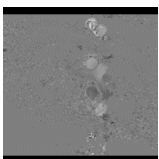
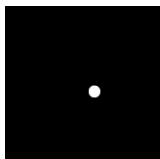
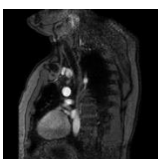
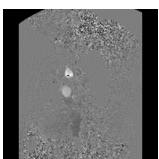
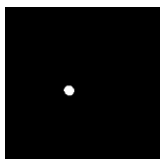
A diferencia del caso 1, la pérdida del contorno de la aorta descendente impacta significativamente las mediciones del flujo. Se observa una gran diferencia entre las curvas de flujo obtenidas por la segmentación manual y las generadas por el método de contorno activo dentro del rango de 100 a 300 [ms], periodo en el que el método de contorno activo pierde el contorno de la aorta descendente. Por otro lado, las curvas de flujo del modelo 3D Unet, en ambos enfoques, muestran curvas similares a la segmentación manual.

4.3 PRUEBAS DE USO GENERAL

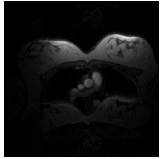
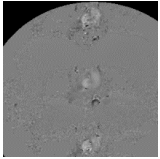
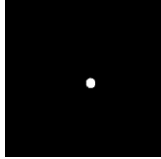
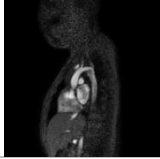
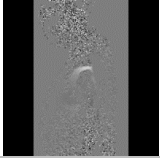
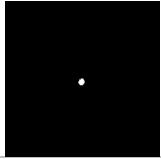
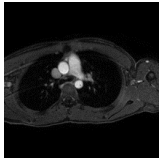
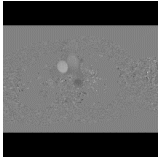
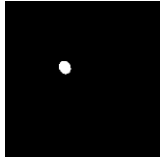
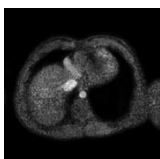
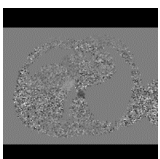
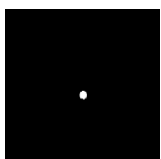
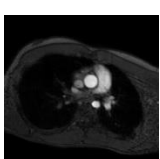
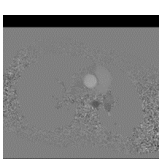
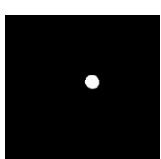

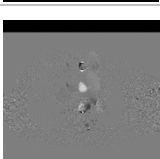
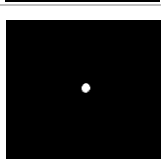
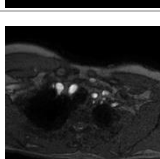
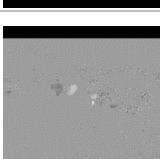
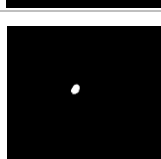

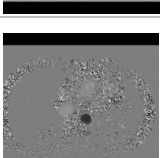
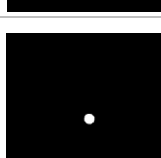
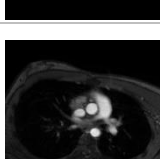
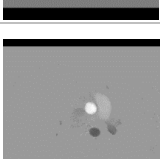
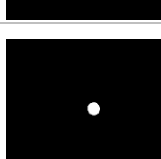
Para evaluar el modelo en casos de uso general, se analizará su rendimiento en un conjunto de casos variados. A diferencia de los casos específicos, estas secuencias 2D PC-MRI no presentan necesariamente problemas de pérdida de nitidez o contornos difusos. Al igual que en la sección anterior, se utilizará el coeficiente de *Dice* como medida comparativa, que expresa el porcentaje de similitud entre dos imágenes. En este contexto, se evaluará qué tan precisa es la segmentación de los métodos estudiados en comparación con la segmentación manual.

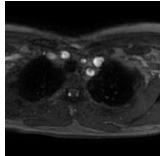
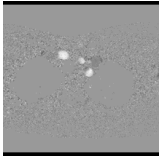
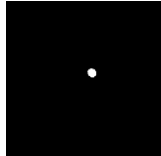
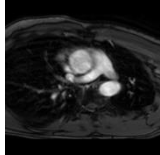
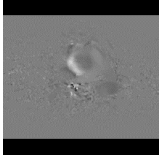
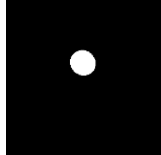
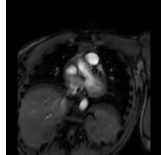
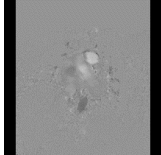
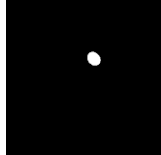
Los resultados se presentan en la tabla 6, la cual incluye: el número de estudio, un fotograma de la secuencia 2D PC-MRI con las imágenes de Magnitud y Fase, la máscara de segmentación que marca el área que segmentarán los métodos en comparación, el número de fotogramas en la secuencia, la resolución de cada imagen y, finalmente, los coeficientes de *Dice* calculados para la segmentación por contorno activo y el modelo 3D Unet, con el enfoque recursivo y secuencial.

Tabla 6: Calculo del coeficiente de *Dice* para las segmentaciones obtenidas por los métodos estudiados, para un conjunto de estudios de secuencias PC-MRI. Fuente: Elaboración propia.

Estudio	PC-MRI		Segmentación	N° Foto-gramas	Resolución imagen	C.D. Contorno Activo	C.D. Unet Recursivo	C.D. Unet Secuencial
1				35	256x256	0,90 ± 0,03	0,95 ± 0,02	0,94 ± 0,02
2				18	256x256	0,86 ± 0,03	0,90 ± 0,03	0,91 ± 0,03
3				140	256x256	0,88 ± 0,05	0,95 ± 0,02	0,95 ± 0,02
4				21	256x256	0,93 ± 0,01	0,96 ± 0,02	0,96 ± 0,02

FRAMEWORK DE SEGMENTACIÓN EFICIENTE DE IMÁGENES DE CONTRASTE DE FASE 2D OBTENIDOS POR
 RESONANCIA MAGNÉTICA

6				142	256x256	$0,90 \pm 0,04$	$0,94 \pm 0,02$	$0,95 \pm 0,02$
7				40	256x256	$0,90 \pm 0,03$	$0,95 \pm 0,02$	$0,94 \pm 0,02$
8				40	256x256	$0,93 \pm 0,01$	$0,94 \pm 0,02$	$0,95 \pm 0,01$
9				34	256x256	$0,91 \pm 0,03$	$0,95 \pm 0,02$	$0,95 \pm 0,02$
10				80	256x256	$0,95 \pm 0,02$	$0,97 \pm 0,02$	$0,95 \pm 0,02$
12				100	256x256	$0,92 \pm 0,02$	$0,96 \pm 0,02$	$0,96 \pm 0,01$
13				40	256x256	$0,89 \pm 0,03$	$0,94 \pm 0,03$	$0,95 \pm 0,02$
14				40	256x256	$0,92 \pm 0,02$	$0,95 \pm 0,01$	$0,96 \pm 0,01$
15				100	256x256	$0,94 \pm 0,02$	$0,95 \pm 0,02$	$0,96 \pm 0,01$

16				40	256x256	$0,91 \pm 0,02$	$0,95 \pm 0,01$	$0,96 \pm 0,01$
21				30	256x256	$0,90 \pm 0,05$	$0,94 \pm 0,01$	$0,94 \pm 0,01$
22				30	256x256	$0,88 \pm 0,03$	$0,92 \pm 0,04$	$0,91 \pm 0,04$

La tabla muestra que el método propuesto, basado en el modelo 3D Unet con enfoques recursivo y secuencial, obtiene un coeficiente de *Dice* superior en todos los estudios comparados con el método de contorno activo. Esto evidencia que el 3D Unet ofrece una segmentación bastante precisa y consistente en casos generales. Para un análisis más detallado del desempeño de las segmentaciones, consulte los gráficos de área y flujo en el **anexo**.

4.4 TIEMPOS DE PREDICCIÓN

Para evaluar los tiempos de ejecución del modelo 3D Unet, se medirá el tiempo necesario para convertir las secuencias de imágenes en volúmenes y realizar la predicción de la segmentación en cada volumen. Se analizarán los tiempos de procesamiento tanto para el enfoque recursivo como para el secuencial del modelo 3D Unet, utilizando tanto la GPU como la CPU.

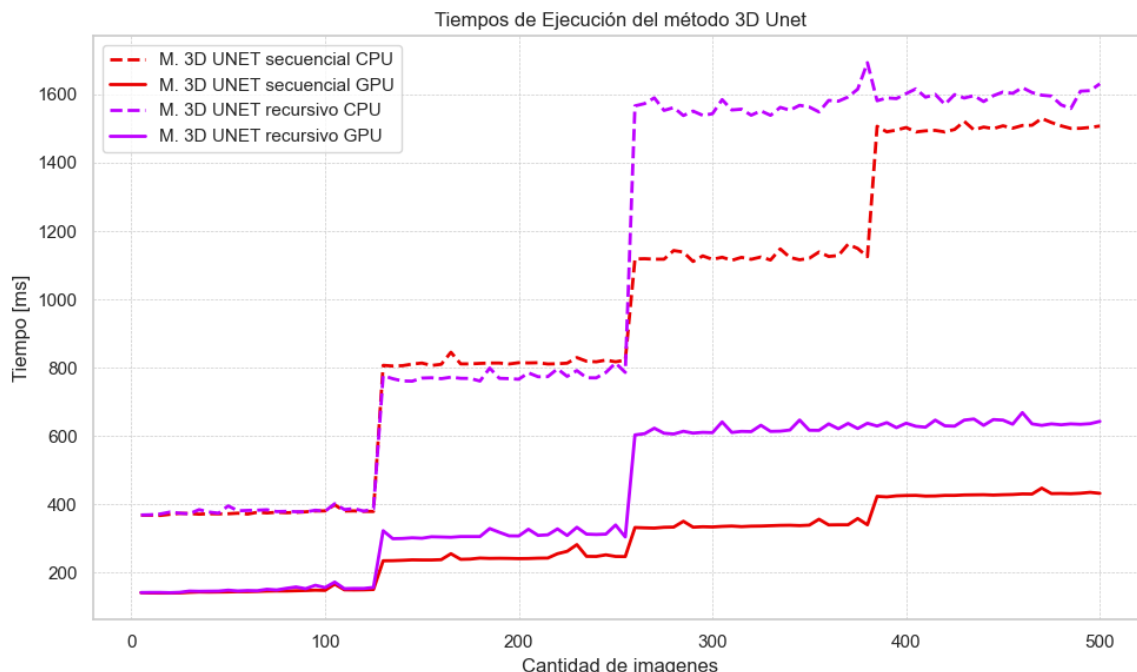


Figura 48: Tiempos de ejecución del modelo 3D Unet, utilizando CPU y GPU. Fuente: Elaboración propia.

El gráfico de tiempos (Figura 48) muestra un incremento escalonado en los tiempos para los enfoques recursivo y secuencial. En particular, al superar los **128** fotogramas, los tiempos aumentan de manera abrupta. Esta variación se debe a la naturaleza de la implementación del modelo, ya que, como se explicó en el capítulo 3, el modelo 3D Unet solo puede procesar entradas con dimensiones fijas y, por lo tanto, se deben crear múltiples volúmenes que logren abarcar todas las imágenes dentro de la secuencia. Para la evaluación, se fijó una dimensión de entrada de **128x32x32**, lo que obliga a subdividir la secuencia completa de imágenes en bloques de **128** imágenes, los cuales se procesan individualmente.

En el enfoque secuencial, cada vez que la cantidad de imágenes supera un múltiplo de 128 (128, 256, 384, 512, etc.), se debe crear un volumen adicional para cubrir las imágenes restantes y permitir su segmentación por el modelo. Esto genera un incremento en el tiempo de procesamiento, que corresponde al tiempo necesario para que el modelo 3D Unet segmente un solo volumen, aproximadamente **100 [ms]** con **GPU** y **400 [ms]** con **CPU**.

En el enfoque recursivo, la secuencia de imágenes se divide recursivamente en imágenes pares e impares hasta que la longitud de cada subdivisión sea igual o inferior a 128. Esto significa que, al superar los 128 fotogramas, la secuencia se divide en dos volúmenes que deben ser procesados por separado. Esta división será suficiente hasta que se superen las

256 imágenes, momento en el cual el enfoque volverá a subdividir en 4 volúmenes. Este proceso continúa hasta superar las 512 imágenes, momento en el que se crean 8 volúmenes, cada uno procesado de manera individual.

El incremento de tiempo para el método recursivo es exponencial, pero este incremento no ocurre al mismo ritmo que con el método secuencial. ya que el método secuencial debe incrementar la cantidad de volúmenes cada vez que la cantidad de imágenes supere un **múltiplo de 128** (128, 256, 384, etc). Mientras que el método recursivo solo lo debe hacer cada **potencia de 2** a partir de que se supere el valor 128 (128, 256, 512, etc). Con esto se desprende que, el método secuencial incrementa su tiempo de procesamiento con cada **múltiplo de L**, mientras que el recursivo lo hace cada **potencia de 2 mayor o igual a L**.

En la práctica, la cantidad de imágenes en una secuencia 2D PC-MRI rara vez supera la cantidad de 200, lo que resulta en tiempos de segmentación bastante bajos. En promedio, el modelo tarda alrededor de 1000 [ms] al procesar con CPU y aproximadamente 300 [ms] con GPU. Estos tiempos dependerán de la dimensión con el cual el modelo 3D Unet se haya entrenado, en especial si se varía el valor **L** correspondiente a la cantidad de imágenes.

CAPÍTULO 5: CONCLUSIONES

En el campo de la inteligencia artificial aplicada a la medicina, la segmentación de imágenes médicas se ha establecido como una herramienta crucial para optimizar el diagnóstico, tratamiento y seguimiento de diversas patologías. Esta técnica permite identificar y delimitar áreas de interés en las imágenes, lo que facilita a los profesionales de la salud tomar decisiones más informadas y precisas. La capacidad de la inteligencia artificial para segmentar imágenes médicas, como las resonancias magnéticas, no solo supera las capacidades humanas en términos de velocidad, sino que, en muchos casos, también ofrece una mayor precisión, lo que es fundamental en contextos clínicos donde cada segundo importa.

Si bien existen métodos alternativos para la segmentación de imágenes médicas que no dependen de redes neuronales, como los métodos de contorno activo, estos presentan limitaciones significativas que pueden afectar su aplicabilidad en situaciones críticas. Estos enfoques dependen en gran medida de la calidad de la imagen, cuando esta es deficiente, como ocurre frecuentemente en muchos casos clínicos, la segmentación resultante es de baja calidad y requiere intervención manual para garantizar su precisión. Este problema se ve incrementado en imágenes con bajo contraste, una situación común en ciertas fases del ciclo cardíaco, donde la definición de la imagen puede verse comprometida, afectando así la efectividad de la segmentación mediante contornos activos. Por lo tanto, la integración de la inteligencia artificial en la segmentación de imágenes médicas representa un avance significativo que puede transformar la práctica clínica, permitiendo diagnósticos más rápidos y precisos y, en última instancia, mejorando los resultados para los pacientes.

En este trabajo, el problema se centró en la segmentación de secuencias 2D PC-MRI, las cuales corresponden a una serie de imágenes que capturan el movimiento dinámico del flujo sanguíneo o de fluidos a lo largo del tiempo. A partir de las secuencias 2D PC-MRI, se pueden obtener dos tipos de imágenes: de Magnitud y de Fase. La imagen de magnitud refleja la intensidad de la señal de resonancia magnética, generada por la cantidad de protones presentes en los tejidos, y es útil para visualizar la anatomía. La imagen de fase, por otro lado, contiene información sobre el movimiento en los tejidos, lo que permite medir la velocidad y dirección del flujo de fluidos, como la sangre.

El principal desafío abordado es la segmentación de los vasos sanguíneos a lo largo de una secuencia temporal. Este problema se complica debido a que, en muchas ocasiones, los contornos de los vasos no son claramente visibles o se pierden en ciertas fases del ciclo cardíaco. Esta falta de definición dificulta la aplicación de métodos de segmentación, como el de contorno activo, los cuales dependen de la claridad de los bordes para mantener una segmentación precisa y coherente.

En este trabajo, con el objetivo de mitigar la pérdida de contraste en las imágenes, se desarrolló un *framework* para la segmentación de imágenes de resonancia magnética con contraste de fase. Las bases de la implementación consistieron en implementar una CNN

que permita aprovechar la información espacial y temporal de las secuencias 2D PC-MRI para mitigar el problema de pérdida de contraste que pueden ocurrir a lo largo del ciclo cardiaco. Bajo esta premisa, se implementó un modelo 3D UNet, el cual recibe de entrada volúmenes tridimensionales. La principal innovación de este enfoque consiste en tratar las secuencias de imágenes 2D PC-MRI como volúmenes, agrupando las imágenes de la secuencia en una tercera dimensión espacial. Como resultado, la información temporal de la secuencia de imágenes se convierte en información espacial, que se organiza en volúmenes, permitiendo así que el modelo 3D UNet aprenda las características relevantes.

Al segmentar la región de interés en estos volúmenes, se segmentan simultáneamente todas las imágenes de la secuencia original, lo que permite que cada imagen sea influenciada por la información contenida en todo el volumen. Esto ayudó a mitigar el problema de la pérdida de contorno, ya que no solo se considera la información de la imagen individual a segmentar, sino también un contexto más amplio que mejora la predicción de la estructura en aquellos casos donde los contornos no son claros.

El modelo 3D Unet fue entrenado utilizando un conjunto de secuencias PC-MRI de distintos estudios, enfocados en analizar el flujo sanguíneo en los principales vasos durante las diferentes fases del ciclo cardíaco, como la sístole y la diástole. Durante la sístole, el corazón se contrae para expulsar la sangre hacia el cuerpo a través de la arteria aorta, lo que genera un mayor contraste en las imágenes PC-MRI. En la diástole, cuando el corazón se relaja después de bombear sangre, el contraste de los vasos tiende a disminuir. Las secuencias se procesaron bajo el enfoque de crear volúmenes que incorporaran la información estructural del vaso sanguíneo en las distintas fases registradas. El entrenamiento se centró en la segmentación de los principales vasos, y para ello se recortó el área específica del vaso sanguíneo en cada secuencia, en lugar de procesar la imagen completa.

La conversión de la segmentación de imágenes individuales a volúmenes planteó una serie de problemas dimensionales, ya que los volúmenes procesables por un modelo 3D Unet deben tener dimensiones fijas. Estas dimensiones se definieron mediante las variables **L**, **W** y **H**, donde **L (Length)** representa la cantidad de fotogramas en la secuencia, y **W (Width)** y **H (Height)** corresponden a la resolución de cada imagen. Se impusieron restricciones a estas variables, estableciendo que sus valores debían ser potencias de **2**, con un mínimo de **16**. Con esta última restricción siendo variable con respecto a la implementación el modelo 3D Unet, donde el valor mínimo viene dado por 2^n , donde **n** es la cantidad de niveles de *pooling* que tenga la implementación.

Debido a las limitaciones dimensionales, las secuencias de imágenes no podían convertirse directamente en volúmenes, ya que podían tener una cantidad arbitraria de fotogramas y resoluciones. Para adaptar las secuencias a las dimensiones requeridas, se implementaron varias técnicas. En cuanto a los valores de **W** y **H**, correspondientes a la resolución de las imágenes, se aplicó un reescalado lineal que ajusta la resolución original a la permitida por el modelo 3D Unet. Para el valor de **L**, que representa la cantidad de imágenes, se desarrollaron dos enfoques: uno **secuencial** y otro **recursivo**. Estos métodos permitieron

crear los volúmenes necesarios para abarcar todas las imágenes de la secuencia, asegurando que cada imagen fuera segmentada.

El enfoque **secuencial** es el más sencillo de los dos y consiste en extraer la cantidad L requerida por el modelo, continuando hasta abarcar todas las imágenes de la secuencia. En caso de que faltaran imágenes, se completaba el volumen con imágenes iniciales de la misma secuencia. Sin embargo, este enfoque presenta un problema: si la secuencia contiene un número significativamente alto de imágenes en comparación con el valor de L , existe el riesgo de que los volúmenes creados incluyan imágenes de una fase del ciclo cardíaco compuesta principalmente por imágenes con poca nitidez de la estructura. Esto podría anular por completo la premisa del *framework*, que busca aprovechar la información temporal de las imágenes en las diferentes fases registradas. Por ello, se implementó el enfoque **recursivo**, que divide la secuencia de manera recursiva entre las imágenes pares e impares. El objetivo de este método es crear subsecuencias que conserven la información de las imágenes a lo largo de todas las fases registradas.

Con base en estas ideas, se implementó el *framework* de segmentación utilizando el modelo 3D Unet, que consta de cinco pasos: 1. Procesamiento de imágenes; 2. Extracción de la región de interés; 3. Creación de los volúmenes; 4. Segmentación con 3D Unet; y 5. Procesamiento de la segmentación. Este método se estructuró para la segmentación de secuencias 2D PC-MRI, empleando tanto el enfoque secuencial como el recursivo. Pero, el *framework* también podría adaptarse para segmentar otras estructuras en las que el factor temporal sea un aspecto importante a considerar.

Para evaluar el desempeño del método propuesto, se examinó la calidad de la segmentación tanto a nivel cualitativo como cuantitativo, comparándolo con el resultado obtenido mediante un método de contorno activo y utilizando la segmentación manual como referencia. El modelo 3D Unet empleado en las pruebas fue entrenado para aceptar volúmenes con dimensiones de **128x32x32**, donde **128** es el valor L y representa la cantidad de imágenes, y **32x32** son los valores $W \times H$ correspondientes a la resolución de cada imagen. Se analizaron dos casos específicos en los que se sabe que el método de contorno activo presenta dificultades para segmentar a lo largo de la secuencia, así como una serie de casos de diferentes estudios, con el objetivo de evaluar el desempeño general del método.

En el caso 1, el método propuesto logró mantener una segmentación muy cercana a la manual a lo largo de toda la secuencia, especialmente durante la fase de diástole, donde el método de contorno activo perdía el contorno de la aorta debido a la pérdida de contraste. El modelo alcanzó un coeficiente de *Dice* de **0,94 ± 0,02** para el enfoque recursivo y de **0,92 ± 0,02** para el enfoque secuencial, en contraste con un valor de **0,82 ± 0,13** obtenido mediante el método de contorno activo.

En el caso 2, el método propuesto logró mantener una segmentación estable durante las fases de sístole y diástole, a diferencia del método de contorno activo, que presenta fallas en la fase de sístole. La segmentación obtenida fue muy cercana a la manual, alcanzando

un coeficiente de *Dice* de **0,95 ± 0,01** para el enfoque recursivo y de **0,96 ± 0,01** para el enfoque secuencial. En contraste, el método de contorno activo obtuvo un valor de **0,86 ± 0,11**.

En los casos de uso general, el método propuesto demostró una superioridad en la segmentación en comparación con el método de contorno activo. Obteniendo un coeficiente de *Dice* superior en todos los casos estudiados. Además, todos los valores obtenidos superaron el valor de **0,90**, con una variabilidad inferior a **0,05** tanto para los enfoques recursivo como secuencial.

En cuanto a los tiempos de ejecución del método propuesto, se observó un incremento escalonado a medida que aumentaba la cantidad de fotogramas en las secuencias de imágenes. Esto debido a la implementación del método, que requiere la creación de suficientes volúmenes para cubrir todas las imágenes de la secuencia, siendo cada volumen procesado de manera individual. Según las mediciones, el procesamiento de un volumen tarda aproximadamente 100 [ms] utilizando GPU y alrededor de 400 [ms] con CPU únicamente.

Debido a esto, los tiempos de ejecución del enfoque recursivo y del secuencial presentan diferencias significativas. En el caso del enfoque **secuencial**, el tiempo de procesamiento aumenta en aproximadamente 100 [ms] con GPU y 400 [ms] con CPU, por cada **múltiplo de 128** imágenes.

Por otro lado, el enfoque **recursivo** muestra un incremento exponencial en el tiempo de procesamiento. Pero, a diferencia del método secuencial, el incremento ocurre cada **potencia de 2** mayor o igual a **128**. Concluyendo que, para casos generalizados, el enfoque secuencial incrementa el tiempo de procesamiento cada **múltiplo de L**, y el enfoque recursivo cada **potencia de 2 mayor a L**.

Considerando un caso promedio, las secuencias 2D PC-MRI rara vez superan las 200 imágenes. Por lo tanto, al utilizar el modelo 3D Unet para procesar volúmenes de **128x32x32**, se estima que el tiempo de procesamiento será de aproximadamente 300 [ms] con GPU y 1000 [ms] utilizando solo CPU, para ambos enfoques. Estos tiempos son muy bajos. Sin embargo, los tiempos de procesamiento pueden variar dependiendo de las dimensiones con las que se haya entrenado el modelo 3D UNet.

5.1 TRABAJOS FUTUROS

Como recomendaciones para trabajos futuros, se sugiere explorar nuevas formas de implementar la arquitectura 3D Unet. En este estudio, se utilizó una versión estándar con 4 niveles de *pooling*; modificar esta arquitectura podría proporcionar resultados interesantes.

También, sería valioso investigar nuevas funciones de pérdida para el entrenamiento del modelo 3D Unet. En este trabajo, se empleó únicamente la función de pérdida *Dice Loss*, que se basa en el coeficiente de *Dice* y mide la similitud entre dos conjuntos. Explorar

funciones de pérdida alternativas que puedan captar detalles más relevantes en las imágenes PC-MRI podría generar resultados significativos.

Por otro lado, mejorar el *framework* añadiendo nuevos pasos u optimizando algunos existentes también es una estrategia interesante para mejorar los resultados y reducir los tiempos de procesamiento. La inclusión de etapas de preprocesamiento más sofisticadas, como la normalización de imágenes o la eliminación de ruido, podría ayudar a mejorar la calidad de entrada al modelo. Además, implementar métodos de post-procesamiento, como la refinación de bordes o la eliminación de pequeñas segmentaciones erróneas, podría aumentar la precisión de la segmentación final.

Finalmente, sería interesante utilizar el modelo para la segmentación de diferentes estructuras anatómicas. En este trabajo, se empleó para la segmentación de los principales vasos sanguíneos. Sería interesante ver el desempeño del modelo en estructuras más complejas, como órganos internos, lesiones, tumores, entre otros.

REFERENCIAS BIBLIOGRÁFICAS

- Berry, M. W., Mohamed, A., & Yap, B. W. (Edits.). (2019). Supervised and unsupervised learning for data science. *Springer Nature*.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation*. Athens, Greece: Springer International Publishing. doi:arXiv:1606.06650
- DiPietro, R., & D. Hager, G. (2020). Chapter 21 - Deep Learning: RNNs and LSTM. En R. DiPietro, & G. D. Hager, *Handbook of Medical Image Computing and Computer Assisted Intervention* (págs. 503-519). Academic Press.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). *Snakes: Active contour models*. International journal of computer vision. doi:10.1007/BF00133570
- kenhub. (30 de Octubre de 2023). *Ciclo cardíaco*. (C. Serrano, Editor) Obtenido de <https://www.kenhub.com/es/library/fisiologia/ciclo-cardiaco-es>
- Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1-18.
- Mahapatra, D. (2013). Cardiac Image Segmentation from Cine Cardiac MRI Using Graph Cuts and Shape Priors. *Journal of digital imaging*, 26, 721-730.
- Mayo Clinic. (2023). *Resonancia magnética*. Obtenido de <https://www.mayoclinic.org/es/tests-procedures/mri/about/pac-20384768>
- medlineplus. (28 de Febrero de 2023). *Electrocardiograma*. Obtenido de <https://medlineplus.gov/spanish/pruebas-de-laboratorio/electrocardiograma/>
- mriquestions. (2024). *Phase-Contrast MRA*. Obtenido de <https://mriquestions.com/phase-contrast-mra>
- Nubimed. (2021). *DICOM: Qué es este formato de imágenes médicas*. Obtenido de <https://www.nubimed.com/blog/dicom-que-es-este-formato-de-imagenes-medicas/>
- O'Shea, K., & Nash, R. (2015). AN INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS. *ARXIV PREPRINT*. doi:arXiv:1511.08458
- Patil, D., & Deore, S. (2013). Medical Image Segmentation: A Review. *International Journal of Computer Science and Mobile Computing*, 2(1), 22-27.
- radiopaedia. (28 de Agosto de 2024). *Cine imaging (MRI)*. (A. Campos, Editor) Obtenido de <https://radiopaedia.org/articles/cine-imaging-mri>.

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. En O. Ronneberger, P. Fischer, & T. Brox, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* (págs. 234-241). Springer International Publishing.
- Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.
- Shruti, J. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (págs. 1-7). IEEE.
- Sotelo, J. A., Valverde, I., Beerbaum, P. B., Greil, G. F., Schaeffter, T., Razavi, R., . . . Figueroa, C. A. (2015). Pressure gradient prediction in aortic coarctation using a computational-fluid-dynamics model: validation against invasive pressure catheterization at rest and pharmacological stress. *Journal of Cardiovascular Magnetic Resonance*.
- Wang, Q., Ma, Y., Zhao, K., & Tian, Y. (2020). A comprehensive survey of loss functions in machine learning. *Annals of Data Science*, 1-26.
- Wang, S.-C. (2003). Artificial Neural Network. En S.-C. Wang, *Interdisciplinary Computing in Java Programming* (págs. 81–100). Boston, MA: Springer.
- Westbrook, C., Roth, K. C., & Talbot, J. (2011). *MRI in Practice, 4th Edition*. Wiley-Blackwell.
- World Health Organization. (2019). *Enfermedades cardiovasculares*. Obtenido de <https://www.who.int/es/health-topics/cardiovascular-diseases>
- Wymer, D. T., Patel, K. P., 3rd, W. F., & Bhatia, V. K. (2020). Phase-contrast MRI: physics, techniques, and clinical applications. *Radiographics*, 40(1), 122-140.
- Yin, X.-X., Sun, L., Fu, Y., Lu, R., & Zhang, Y. (2022). U-Net-Based Medical Image Segmentation. *Journal of Healthcare Engineering*, 16.

ANEXOS

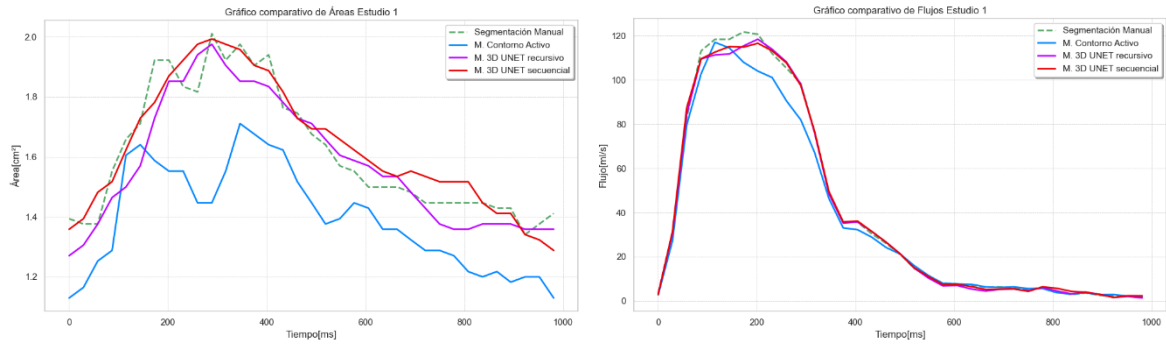


Figura 49: Gráficos de Área y Flujo para el estudio 1. Fuente: Elaboración propia.

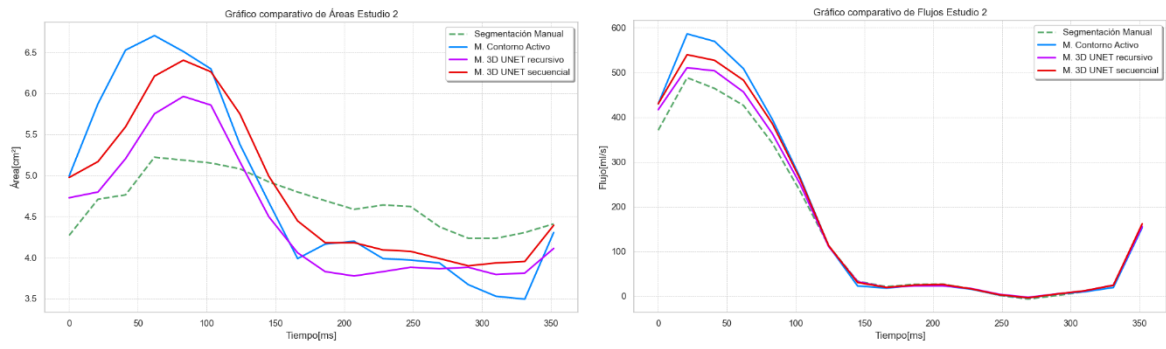


Figura 50: Gráficos de Área y Flujo para el estudio 2. Fuente: Elaboración propia.

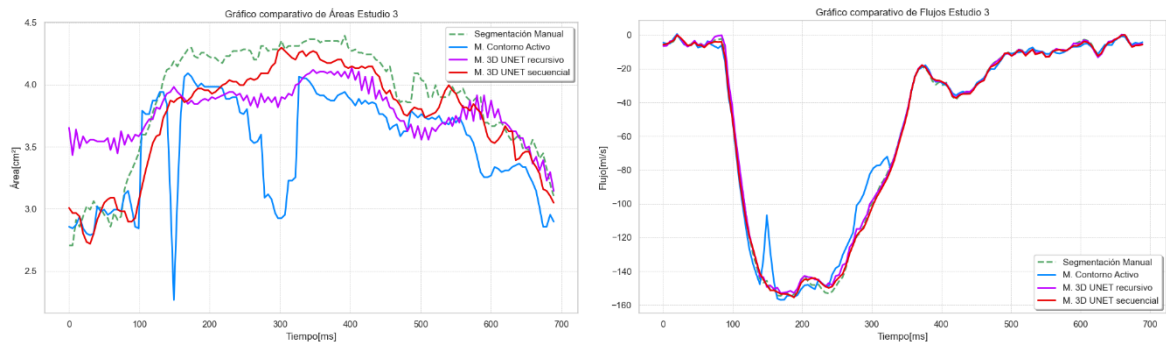


Figura 51: Gráficos de Área y Flujo para el estudio 3. Fuente: Elaboración propia.

FRAMEWORK DE SEGMENTACIÓN EFICIENTE DE IMÁGENES DE CONTRASTE DE FASE 2D OBTENIDOS POR RESONANCIA MAGNÉTICA

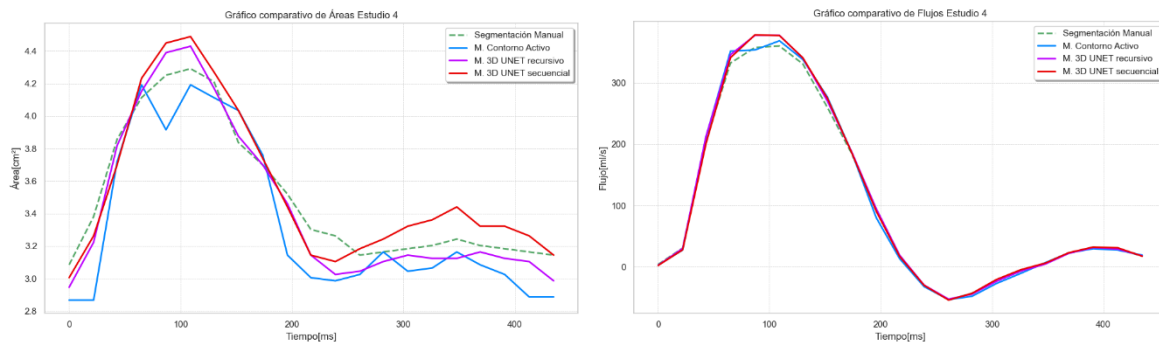


Figura 52: Gráficos de Área y Flujo para el estudio 4. Fuente: Elaboración propia.

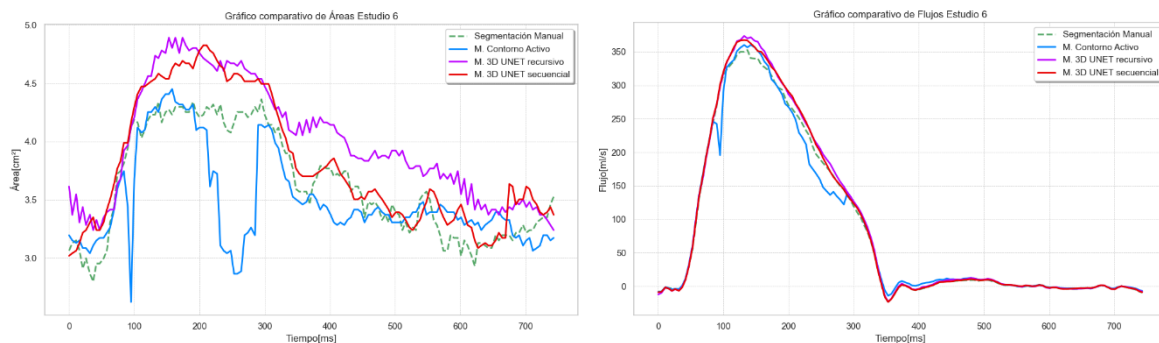


Figura 53: Gráficos de Área y Flujo para el estudio 6. Fuente: Elaboración propia.

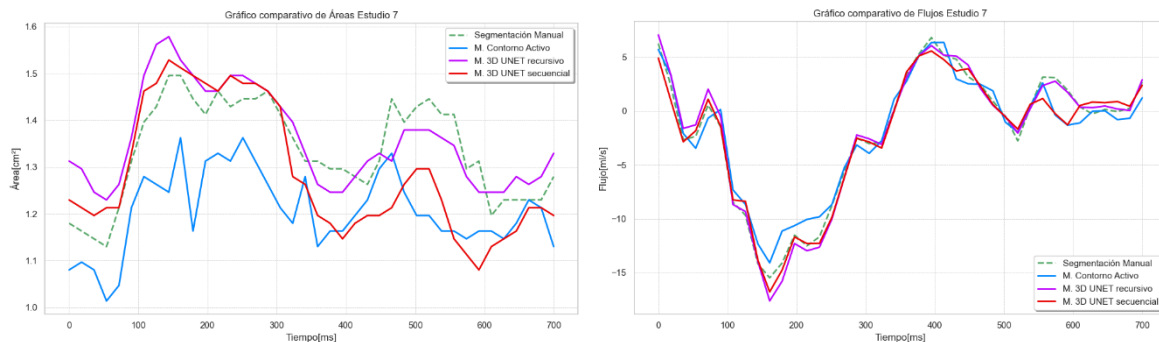


Figura 54: Gráficos de Área y Flujo para el estudio 7. Fuente: Elaboración propia.

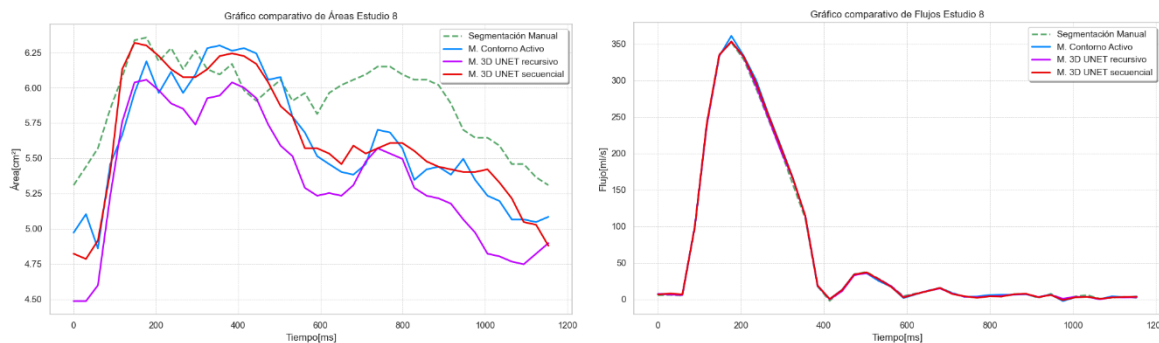


Figura 55: Gráficos de Área y Flujo para el estudio 8. Fuente: Elaboración propia.

FRAMEWORK DE SEGMENTACIÓN EFICIENTE DE IMÁGENES DE CONTRASTE DE FASE 2D OBTENIDOS POR RESONANCIA MAGNÉTICA

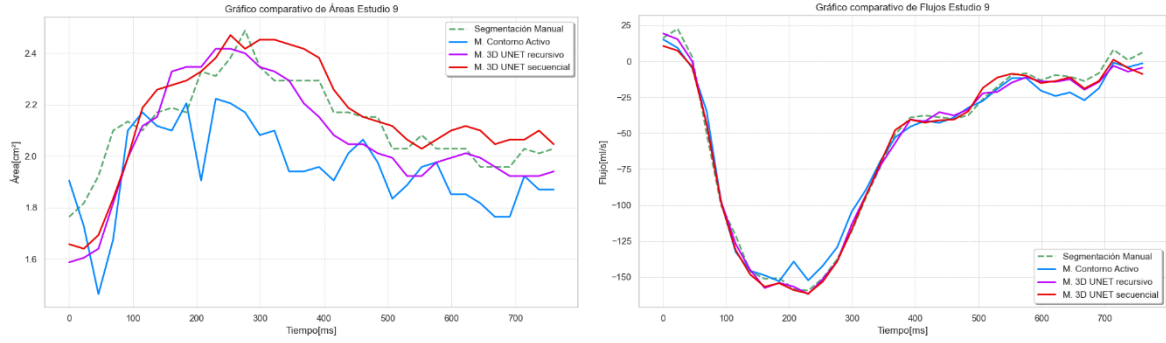


Figura 56: Gráficos de Área y Flujo para el estudio 9. Fuente: Elaboración propia.

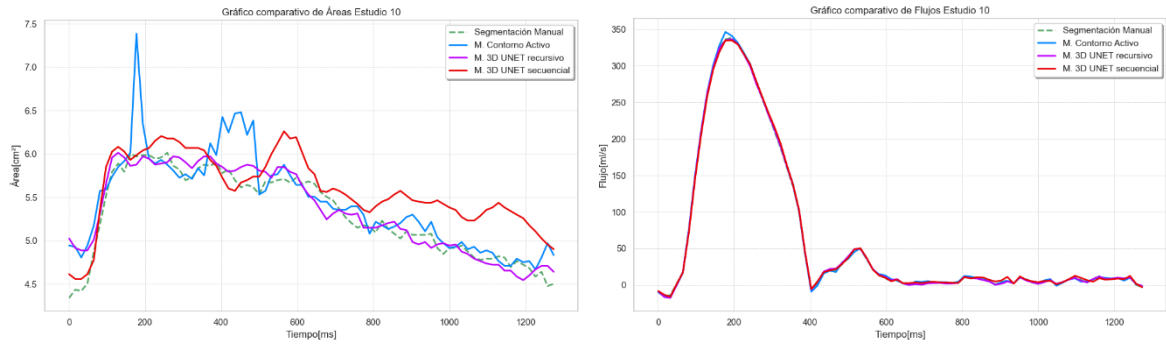


Figura 57: Gráficos de Área y Flujo para el estudio 10. Fuente: Elaboración propia.

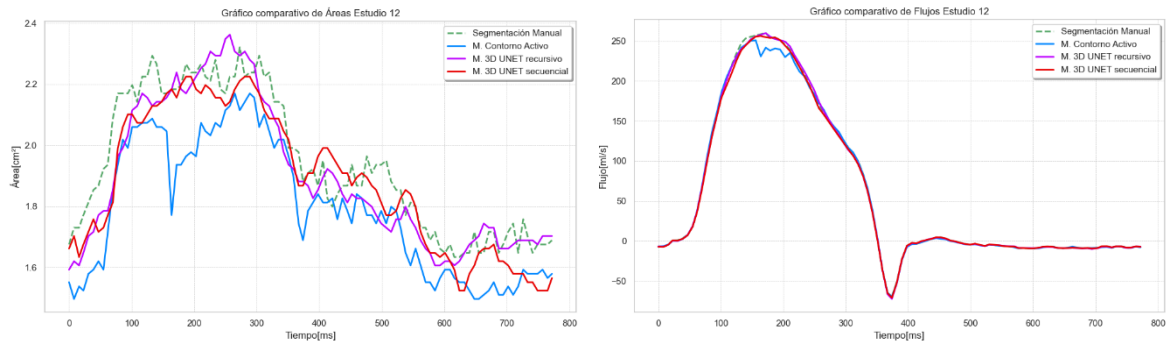


Figura 58: Gráficos de Área y Flujo para el estudio 12. Fuente: Elaboración propia.

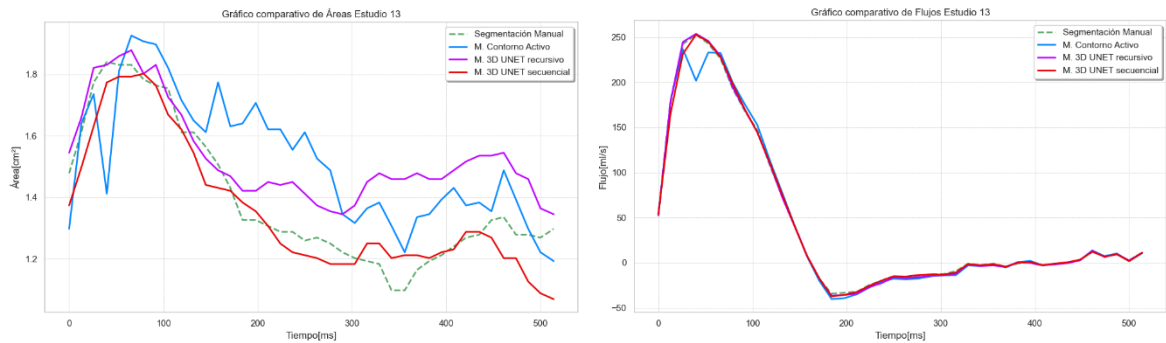


Figura 59: Gráficos de Área y Flujo para el estudio 13. Fuente: Elaboración propia.

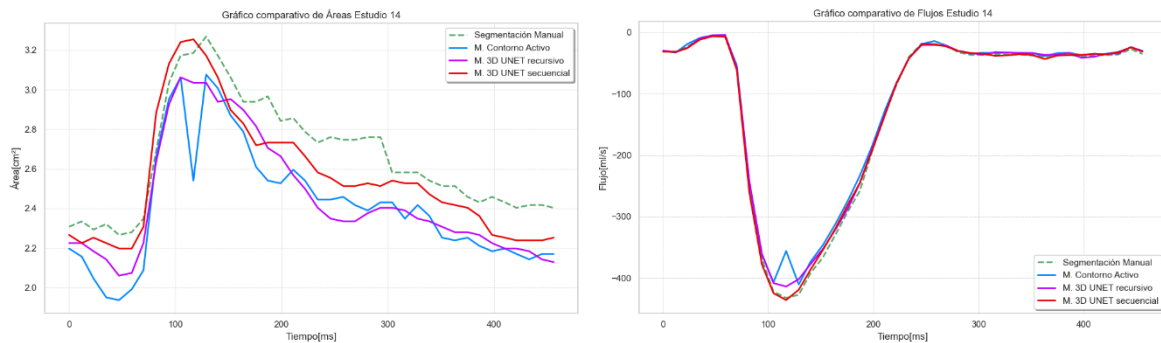


Figura 60: Gráficos de Área y Flujo para el estudio 14. Fuente: Elaboración propia.



Figura 61: Gráficos de Área y Flujo para el estudio 15. Fuente: Elaboración propia.

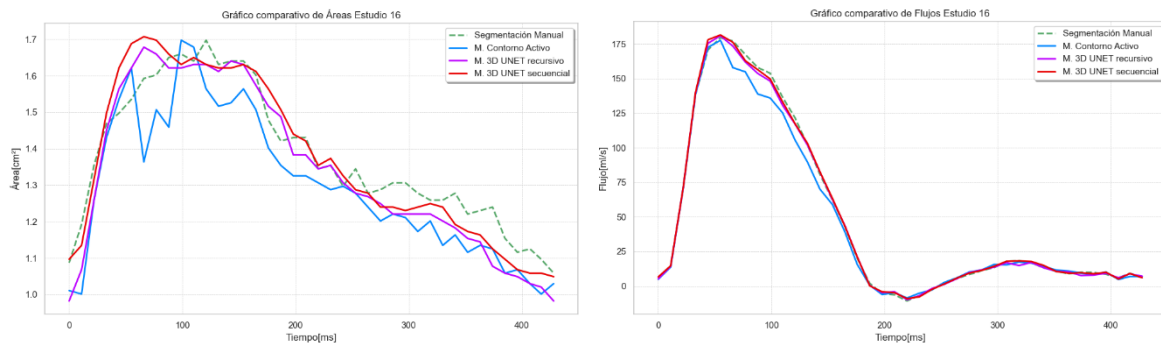


Figura 62: Gráficos de Área y Flujo para el estudio 16. Fuente: Elaboración propia.

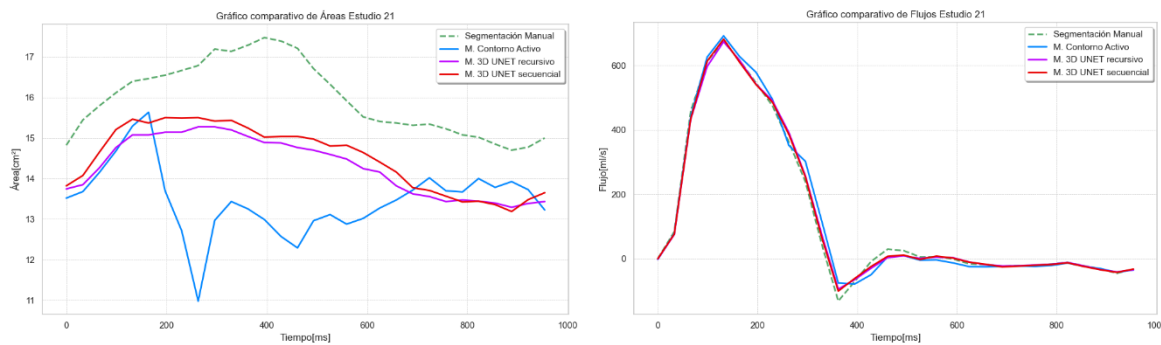


Figura 63: Gráficos de Área y Flujo para el estudio 21. Fuente: Elaboración propia.

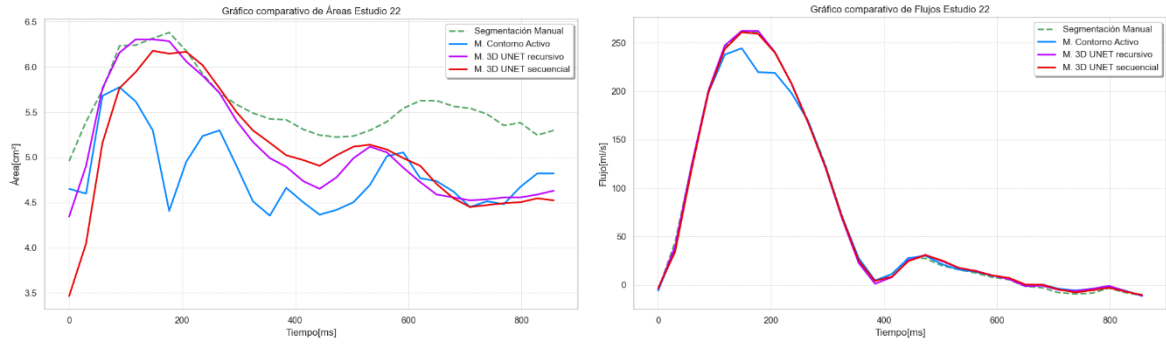


Figura 64: Gráficos de Área y Flujo para el estudio 22. Fuente: Elaboración propia.