



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

**ALGORITMO AVANZADO DE SEGMENTACIÓN Y CUANTIFICACIÓN DE
CALCIFICACIONES ARTERIALES EN TOMOSÍNTESIS MAMARIAS USANDO DEEP
LEARNING Y VISION TRANSFORMER**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

BASTIÁN FRANCISCO CASTELLI PINO

PROFESOR GUÍA:
VÍCTOR CASTAÑEDA ZEMAN

MIEMBROS DE LA COMISIÓN:
DENISSE KARL SÁEZ
CARLOS NAVARRO CLAVERÍA

SANTIAGO DE CHILE

2025

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: BASTIÁN FRANCISCO CASTELLI PINO
FECHA: 2025
PROF. GUÍA: VÍCTOR CASTAÑEDA ZEMAN

ALGORITMO AVANZADO DE SEGMENTACIÓN Y CUANTIFICACIÓN DE CALCIFICACIONES ARTERIALES EN TOMOSÍNTESIS MAMARIAS USANDO DEEP LEARNING Y VISION TRANSFORMER

Las imágenes médicas mamarias no solo permiten detectar hallazgos sugestivos de cáncer, sino también identificar calcificaciones arteriales mamarias (BAC), las cuales se han asociado con factores de riesgo cardiovascular. Actualmente, la presencia de BAC se reporta sólo de manera cualitativa y subjetiva, sin un estándar cuantitativo establecido, lo que genera la necesidad de automatizar su detección y cuantificación. La tomosíntesis mamaria, técnica que genera reconstrucciones cuasi-tridimensionales de la mama, ha surgido como una alternativa a la mamografía digital al reducir limitaciones asociadas a la densidad mamaria. Sin embargo, los estudios que abordan la segmentación automática de BAC en este tipo de imágenes mediante inteligencia artificial son aún escasos. Este trabajo de título propone desarrollar un algoritmo para segmentar y cuantificar BAC en tomosíntesis mamaria, con el objetivo de contribuir a la estandarización de su evaluación y favorecer la identificación temprana de riesgo cardiovascular en mujeres.

Se trabajó con una base de datos tridimensional de 49 volúmenes mamarios, a partir de la cual se construyó una nueva base bidimensional para permitir el uso de redes 2D y reducir los costos computacionales. Se implementaron y evaluaron modelos U-Net, Attention U-Net y Swin UNetR, combinando funciones de pérdida *Dice Loss* y *Focal Loss* para abordar el desbalance de clases presentes en la segmentación de BACs. Adicionalmente, se diseñó un segmentador de sólo arterias calcificadas y sus predicciones se utilizaron como máscara para filtrar sobresegmentaciones en la predicción de BAC, mejorando así su rendimiento.

Los experimentos realizados demostraron que el modelo Attention U-Net, combinado con un filtrado por postprocesamiento, alcanzó un coeficiente de Dice promedio del 35,1%, cuadruplicando el rendimiento reportado en la tesis previa desarrollada sobre la misma base de datos. Sin embargo, este desempeño sigue siendo insuficiente para aplicaciones clínicas, identificándose como principales limitantes el reducido tamaño del conjunto de datos, las restricciones computacionales disponibles y el hecho de que la métrica de evaluación empleada no refleja adecuadamente el desempeño cualitativo del modelo. Finalmente, se desarrolló de manera preliminar un indicador de cuantificación de calcificaciones basado en el área segmentada, abriendo la posibilidad a futuros desarrollos que permitan estimar el riesgo cardiovascular de forma más precisa. El trabajo sienta así una base sólida para continuar avanzando en la mejora de modelos y métricas aplicadas a este relevante desafío clínico.

Agradecimientos

Agradezco a mi familia por apoyarme durante este recorrido académico frente a toda adversidad que se haya presentado y brindarme las herramientas para ser quien soy hoy en día y tener la gente que me rodea a diario.

Agradezco a mi grupo de amigos que conocí al entrar a la Universidad en Plan Común quienes hicieron este cambio muy ameno, en especial quiero agradecer a Ayelen, Maite y Nicole quienes fueron mis pilares durante esta carrera, apoyándome en toda situación que estuviese pasando, acompañándome de incontables risas y que no me permitieron sentirme solo en ningún momento.

Agradezco a mi grupo de amigos que tuve al regresar de la pandemia e integrarme al Departamento de Ingeniería Eléctrica, en especial agradezco a Luciano, Pedre, Donko, Pilar y Coty, por su gran apoyo y sonrisas durante mi recorrido en la especialidad, además de soportarme y calmar mis nervios y pánico durante el desarrollo de este trabajo de título.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.2.1. Objetivo general	2
1.2.2. Objetivos específicos	2
1.3. Estructura del documento	2
2. Antecedentes	4
2.1. Marco teórico	4
2.1.1. Tomosíntesis mamaria	4
2.1.2. Calcificaciones arteriales mamarias (BAC)	5
2.1.3. Machine Learning (ML)	6
2.1.4. Deep Learning	6
2.1.5. Métricas de evaluación	8
2.1.6. Función de Pérdida	9
2.1.7. Data Augmentation	10
2.1.8. Arquitectura U-Net	10
2.1.9. Segmentación de Imágenes	11
2.1.10. Arquitectura Attention U-Net	12
2.1.11. Vision Transformers	13
2.1.12. Swin Transformers	15
2.1.13. Arquitectura Swin U-Net	16
2.1.14. Arquitectura Swin UNetR	17
2.2. Estado del arte	18
3. Metodología	20
3.1. Base de datos	20
3.2. Equipamiento y Herramientas	20
3.3. Procedimiento	21
4. Resultados y Análisis	31
4.1. Modelo en Cascada	31
4.2. Modelo de Arquitectura Única	33

4.3. Filtrado de Predicciones	36
4.4. Cuantificador de Calcificaciones	43
5. Discusión	46
6. Conclusión	49
6.1. Proyección a Futuro	50
Bibliografía	51
Anexos	54
A. Resultados detallados por volumen	54
A.1. Coeficientes de Dice para U-Net bajo distintas funciones de pérdida	54
A.2. Coeficientes de Dice para distintas arquitecturas	55
A.3. Coeficientes de Dice para diferentes filtros	55
B. Predicciones del mejor modelo desarrollado	57

Índice de Tablas

2.1.	Matriz de Confusión.	8
3.1.	Experimentos realizados al entrenar una misma U-Net variando la función de pérdida.	28
4.1.	Comparación del coeficiente de Dice del modelo en cascada realizado con los resultados de la tesis previa.	31
4.2.	Coefficientes de Dice promedio de entrenamientos con la red U-Net variando su función de pérdida y sus hiperparámetros.	33
4.3.	Coefficientes de Dice promedio de entrenamientos variando la arquitectura usada y manteniendo una misma función de pérdida.	34
4.4.	Coefficientes de Dice promedio al aplicar la máscara obtenida con distintos segmentadores de arterias.	37
4.5.	Comparación de métricas del modelo segmentador de BACs (Attention U-Net con filtrado por arterias) frente a resultados de la tesis previa.	39
4.6.	Valor del índice de calcificación diseñado en las predicciones del mejor modelo filtrado creado.	44
A.1.	Coefficientes de Dice por volumen para U-Net bajo distintas funciones de pérdida.	54
A.2.	Coefficientes de Dice por volumen para distintas arquitecturas.	55
A.3.	Coefficientes de Dice para diferentes filtros	55

Índice de Ilustraciones

2.1.	Representación de mama con dos lesiones superpuestas. A) Imagen obtenida con mamografía donde una lesión oculta a la otra. B) Imagen obtenida con tomosíntesis mamaria, donde se logra detectar ambas lesiones gracias al barrido del tubo de rayos X [2].	4
2.2.	Mamografía donde se aprecian diferentes morfologías de BAC en las secciones C y D [4].	5
2.3.	Arquitectura Típica de una Red Convolutiva 2D (CNN) [12].	7
2.4.	Diferencia de la capa convolutiva entre una CNN 2D y una CNN 3D [11].	7
2.5.	Arquitectura de la red U-Net [17].	11
2.6.	Diagrama explicativo de la arquitectura Attention U-Net y sus módulos de atención [18].	13
2.7.	Arquitectura Transformers. [19]	14
2.8.	Modelo Vision Transformer. [20]	14
2.9.	Esquema general del Swin Transformer [23].	16
2.10.	Esquema general de la arquitectura Swin U-Net [24].	17
2.11.	Esquema general de la arquitectura Swin UNetR [25].	18
3.1.	Arquitectura U-Net tridimensional diseñada.	22
3.2.	Representación del paso de base de datos tridimensional a nueva base de datos bidimensional.	23
3.3.	Representación de los cortes sin fondo almacenados en la base de datos bidimensional.	24
3.4.	Errores en las etiquetas al redimensionar los cortes de mama en 256×256 píxeles.	25
3.5.	Modelo en cascada diseñado.	27
4.1.	Caso arbitrario de una predicción de arterias calcificadas en el conjunto de validación.	33
4.2.	Comparación de las curvas de pérdida de las arquitecturas Attention U-Net y Swin UNetR ante una misma función de pérdida.	35
4.3.	Ejemplo de segmentación realizada con arquitectura Attention U-Net.	36
4.4.	Representación del filtrado de detecciones fuera de arterias calcificadas. En (b), los rectángulos rojos destacan algunas predicciones positivas que no corresponden a calcificaciones asociadas a vasos, ilustrando las sobresegmentaciones que son eliminadas tras aplicar el filtrado mostrado en (c).	38

4.5.	Predicción representativa de casos con alta precisión pero baja sensibilidad. En (c) se observa que las regiones más visibles de las arterias son correctamente segmentadas, mientras que las zonas parcialmente ocultas por tejido denso no son detectadas, explicando el comportamiento de alta precisión pero baja sensibilidad.	41
4.6.	Predicción representativa de casos con baja precisión pero sensibilidad alta. En (b) se observa que la calcificación es extremadamente reducida en comparación con el tamaño total de la mama, el rectángulo rojo indica la localización aproximada de dicha calcificación para facilitar su visualización. (c) Indica que el modelo logra detectar correctamente la calcificación, pero cualquier sobresegmentación mínima impacta fuertemente las métricas cuantitativas, resultando en baja precisión.	42
4.7.	Predicción en volumen con máscara nula. En (b) se observa que la predicción es extremadamente reducida en comparación con el tamaño total de la mama, los rectángulos rojos indican la localización aproximada de dicha predicción para facilitar su visualización.	43
B.1.	Predicción del volumen DBT-P00001_MLOI con coeficiente de Dice 0,575.	57
B.2.	Predicción del volumen DBT-P00285_MLOI con coeficiente de Dice 0,261.	58
B.3.	Predicción del volumen DBT-P00303_MLOD con coeficiente de Dice 0,126.	59
B.4.	Predicción del volumen DBT-P00491_MLOD con coeficiente de Dice 0,200.	60
B.5.	Predicción del volumen DBT-P00523_MLOD, de etiqueta nula, con coeficiente de Dice 0,000.	61
B.6.	Predicción del volumen DBT-P00688_MLOD con coeficiente de Dice 0,391.	62
B.7.	Predicción del volumen DBT-P00887_MLOD con coeficiente de Dice 0,525.	63
B.8.	Predicción del volumen DBT-P01488_MLOD con coeficiente de Dice 0,262.	64
B.9.	Predicción del volumen DBT-P01762_MLOI con coeficiente de Dice 0,481.	65
B.10.	Predicción del volumen DBT-P01778_MLOI con coeficiente de Dice 0,269.	66
B.11.	Predicción del volumen DBT-P01957_MLOD con coeficiente de Dice 0,336.	67
B.12.	Predicción del volumen DBT-P02244_MLOD con coeficiente de Dice 0,451.	68
B.13.	Predicción del volumen DBT-P02510_MLOD con coeficiente de Dice 0,722.	69
B.14.	Predicción del volumen DBT-P04368_MLOD con coeficiente de Dice 0,221.	70
B.15.	Predicción del volumen DBT-P04396_MLOD con coeficiente de Dice 0,436.	71

Capítulo 1

Introducción

1.1. Motivación

Las imágenes médicas mamarias se utilizan para buscar hallazgos sugerentes de cáncer. Sin embargo, estas imágenes pueden presentar más hallazgos, tales como las calcificaciones arteriales mamarias (abreviadas como BAC por su nombre en inglés *Breast Arterial Calcification*), debido a que la comunidad científica ha observado una asociación de estas con factores de riesgo de patologías cardiovasculares [1].

En la actualidad, en el contexto clínico sólo existe una notificación cualitativa al identificar visualmente las BAC, es decir, simplemente se indica si hay presencia o no presencia de arterias calcificadas, pero dado a la subjetividad de la interpretación del observador, no existe aún un estándar que permita cuantificar estas BAC de forma certera, por ende, existe una necesidad de automatizar el proceso de detección y cuantificación de las BAC para lograr estandarizar las evaluaciones. De esta manera se podría ayudar a los radiólogos a tener un indicador de riesgo de enfermedades cardiovasculares.

Por otra parte, hoy en día se está implementando la técnica de tomosíntesis mamaria, la cual captura diversas imágenes bidimensionales de la mama desde múltiples proyecciones, ajustando el ángulo con que se toma cada una, obteniendo una representación cuasi-tridimensional de la mama, lo que mejora la visualización de estructuras internas y reduce las complicaciones asociadas a la superposición de tejidos, especialmente en mamas densas, en comparación con las imágenes obtenidas mediante mamografía digital convencional [2]. Sin embargo, existen pocos estudios de detección y segmentación de BAC con inteligencia artificial en imágenes médicas obtenidas con tomosíntesis mamarias, la mayoría de estudios utilizan imágenes obtenidas con la segunda técnica mencionada.

Debido a la pseudo-tridimensionalidad de las imágenes de tomosíntesis y la relación de la aparición de BACs con enfermedades cardiovasculares es que este trabajo pretende crear un algoritmo para la segmentación y cuantificación de calcificaciones arteriales mamarias. Con el propósito de

contribuir a estandarizar la evaluación de las BAC, permitiendo identificar de manera temprana posibles factores de riesgo cardiovascular y apoyar el diagnóstico clínico y el seguimiento oportuno de las pacientes que se someten periódicamente a estos exámenes.

1.2. Objetivos

1.2.1. Objetivo general

Desarrollar un algoritmo avanzado de deep learning y vision transformer para la detección y cuantificación de calcificaciones vasculares en imágenes de tomosíntesis mamaria, abordando los desafíos de la cuasi-tridimensionalidad de la imagen y las características morfológicas de las calcificaciones.

1.2.2. Objetivos específicos

Para lograr cumplir con el objetivo general descrito, se plantean los siguientes objetivos específicos:

1. Desarrollar un algoritmo de Deep Learning y Vision Transformer que segmente las calcificaciones arteriales para automatizar su cuantificación.
2. Validar el algoritmo de segmentación desarrollado para medir su precisión según su coeficiente de Dice.
3. Creación de un indicador que cuantifique las calcificaciones vasculares para un potencial uso como indicador de riesgos de enfermedades cardiovasculares.

1.3. Estructura del documento

En el Capítulo 2 se indican los antecedentes necesarios para la comprensión de lo desarrollado en capítulos posteriores. Particularmente, se detallan los términos médicos necesarios para el entendimiento del problema a resolver, además de los conceptos de inteligencia computacional que serán de utilidad para comprender los modelos a desarrollar en este estudio. En este capítulo también se señalan los últimos avances que se logran encontrar en la literatura sobre soluciones propuestas a este problema.

En el Capítulo 3 se describe la metodología utilizada en este trabajo de título, comenzando por detallar la base de datos a partir de la cual se obtuvieron las imágenes de tomosíntesis mamaria. A continuación, se presentan el equipamiento disponible y las herramientas utilizadas para el desarrollo del proyecto. Posteriormente, se explica el procedimiento llevado a cabo para cumplir con cada uno de los objetivos planteados, junto con los fundamentos que respaldan las decisiones tomadas en cada etapa.

El Capítulo 4 presenta los resultados obtenidos a partir del procedimiento descrito en el capítulo anterior, junto con el análisis correspondiente de cada uno de ellos.

Posteriormente, en el Capítulo 5 se discuten los resultados obtenidos, comparándolos con la tesis previa a este trabajo y destacando tanto las fortalezas como las limitaciones del algoritmo diseñado.

Por último, en el Capítulo 6 se presentan las conclusiones del trabajo realizado, destacando cómo se cumplieron los objetivos específicos planteados. Además, se proponen posibles mejoras al algoritmo, derivadas de las limitaciones identificadas en este estudio.

Capítulo 2

Antecedentes

2.1. Marco teórico

2.1.1. Tomosíntesis mamaria

La mamografía es una técnica de imagen mamaria bidimensional utilizada para detectar cáncer en etapas tempranas. Sin embargo, su sensibilidad se ve considerablemente reducida en el caso de mamas densas [2]. Para superar esta limitación, en la actualidad se ha comenzado a implementar la tomosíntesis mamaria, una técnica que adquiere múltiples imágenes bidimensionales de baja dosis de radiación desde diferentes ángulos, con la mama comprimida, mediante el desplazamiento del tubo de rayos X a lo largo de un arco predeterminado. Este procedimiento permite reconstruir un volumen cuasi tridimensional de la mama en cortes paralelos al detector, típicamente de 1 *mm* de espesor.

En la Figura 2.1 se puede apreciar cómo las imágenes obtenidas utilizando tomosíntesis mamaria brindan información adicional al rotar el tubo de rayos X, logrando detectar ambas lesiones, a diferencia de la imagen obtenida utilizando una mamografía.

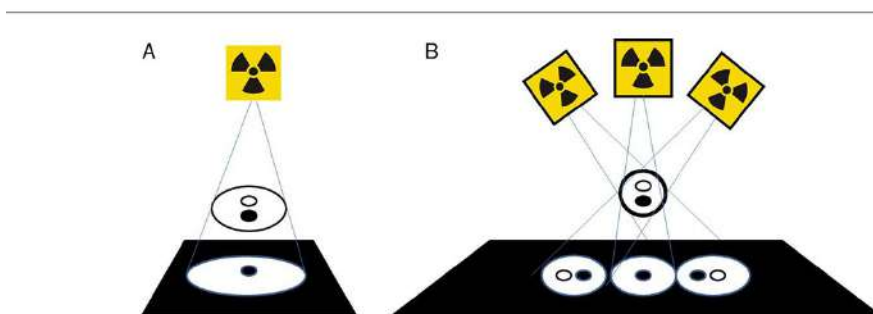


Figura 2.1: Representación de mama con dos lesiones superpuestas. A) Imagen obtenida con mamografía donde una lesión oculta a la otra. B) Imagen obtenida con tomosíntesis mamaria, donde se logra detectar ambas lesiones gracias al barrido del tubo de rayos X [2]

2.1.2. Calcificaciones arteriales mamarias (BAC)

La calcificación vascular fue considerada durante mucho tiempo como un proceso pasivo y degenerativo, asociado principalmente al envejecimiento, la aterosclerosis, alteraciones metabólicas como la diabetes mellitus o la enfermedad renal crónica avanzada, y a ciertas enfermedades genéticas poco comunes. Sin embargo, estudios más recientes han establecido que se trata de un proceso activo y regulado, comparable al de la mineralización ósea y el metabolismo del hueso [3].

Las BAC corresponden a depósitos de calcio mediales no oclusivos y circunferenciales en las arterias de tamaño pequeño o mediano [4], las cuales se pueden observar en imágenes médicas de mamas como opacidades lineales, tubulares o paralelas, a menudo descritas como “vías de tren”, diferenciándose de microcalcificaciones malignas por su morfología y distribución, presentando patrones de complejidad topológica y superposición de vasos y/o tejidos en proyecciones bidimensionales, dificultando así la estandarización de la identificación y cuantificación de las BAC [5], como es apreciable en la Figura 2.2.

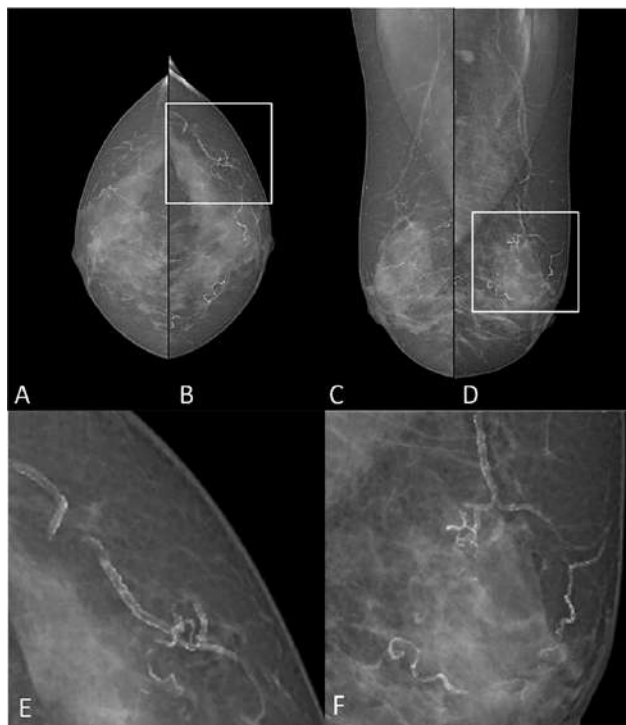


Figura 2.2: Mamografía donde se aprecian diferentes morfologías de BAC en las secciones C y D [4].

Recientemente, las BAC han sido descritas dentro de “los cinco principales problemas de salud de las mujeres en cardiología preventiva, en la vanguardia de la investigación reciente y en curso”, además, se ha reportado una relación entre las BAC y enfermedades cardiovasculares, como la enfermedad arterial coronaria (CAD) [5]. Por ende, las BACs podrían ser utilizadas como un biomarcador en una imagen mamaria para determinar el riesgo cardiovascular y funcionar como un

factor predictivo para eventos cardiovasculares.

2.1.3. Machine Learning (ML)

El Machine Learning es un subcampo de la inteligencia artificial que busca replicar la forma en que los humanos aprenden, utilizando datos y algoritmos y gradualmente mejorando su precisión [6]. El Machine Learning se puede desglosar en tres principales etapas: representación, evaluación y optimización [7].

- Representación: Es la forma en que el sistema modela los datos y organiza el espacio de hipótesis para la tarea de clasificación o predicción.
- Evaluación: Una función de evaluación, también conocida como *scoring function*, que mide la calidad de las hipótesis generadas.
- Optimización: Un método para buscar la mejor solución posible dentro del espacio de hipótesis, comúnmente empleando técnicas de optimización.

2.1.4. Deep Learning

Las redes neuronales son un enfoque del machine learning que toma decisiones de forma similar al cerebro humano, intentando imitar la manera en que las neuronas del cerebro trabajan en conjunto para identificar sucesos, evalúan las opciones y llegan a una conclusión para dicho suceso.

Están compuestas por distintas capas de nodos (también llamados neuronas) interconectadas, una capa de entrada en la cual ingresan los datos para el procesamiento, una o más capas ocultas y una capa de salida donde ocurre la clasificación. Cada nodo tiene su respectivo peso y umbral, en caso de superar este último valor, la neurona se activa y envía información a la siguiente capa [8].

Las redes neuronales son el núcleo de los modelos Deep Learning, pues consisten en redes neuronales que poseen tres o más capas de nodos interconectados, además de su respectiva capa de entrada y de salida. Se diferencia con los modelos de machine learning al no requerir que los datos de entrada estén estructurados y etiquetados para obtener resultados precisos, pues pueden extraer características, relaciones y rasgos necesarios a partir de datos no estructurados [9].

En específico, existe un tipo red neuronal llamadas redes neuronales convolucionales (CNN), que se utilizan en clasificación y segmentación de imágenes, intentando imitar las capacidades de la visión humana de detectar las características de una imagen [10], cuya arquitectura típica es observable en la Figura 2.3. Debido a esto, este es el tipo de modelo deep learning que es utilizado para cumplir el objetivo de esta memoria, destacando que estas redes neuronales convolucionales pueden ser bidimensionales o tridimensionales, dependiendo de la imagen con que se quiera trabajar. Las CNN 3D permiten el procesamiento y análisis con datos volumétricos, como imágenes médicas obtenidas mediante tomosíntesis mamaria. [11].

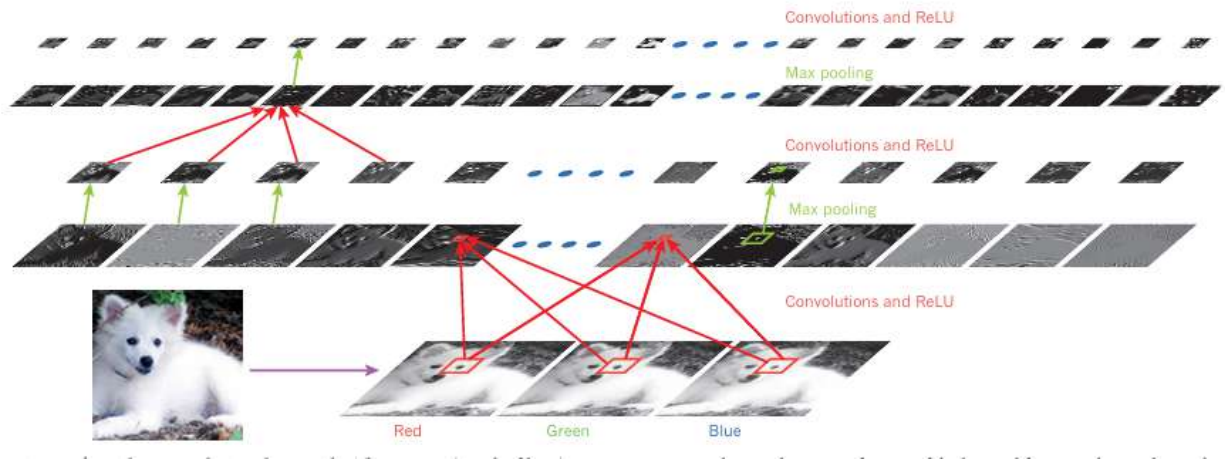


Figura 2.3: Arquitectura Típica de una Red Convolutiva 2D (CNN) [12].

La red neuronal convolutiva 3D se encuentra compuesta por las siguientes capas [11]:

1. Capa de entrada: Recibe los datos de entrada volumétricos donde cada elemento representa un punto de datos con su correspondiente valor.
2. Capas de convolución: Se filtra el volumen de entrada o las características extraídas de una capa anterior. Los filtros son cubos que se deslizan a través del volumen realizando una convolución en las tres dimensiones, permitiendo extraer características como textura, patrones de movimiento y otros detalles espaciales o temporales. En la figura 2.4 se puede observar la diferencia de esta capa con la vista en la arquitectura de una CNN 2D observada en la Figura 2.3.

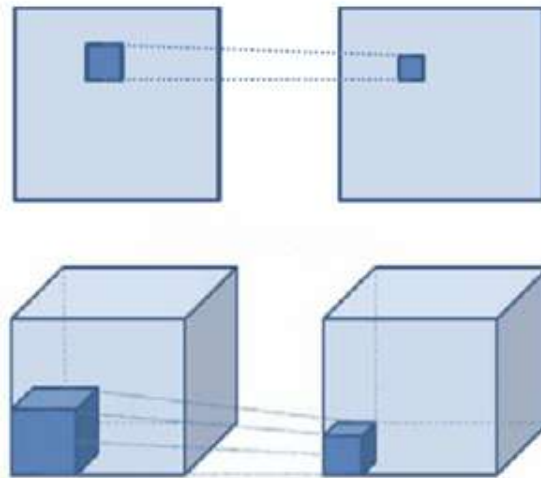


Figura 2.4: Diferencia de la capa convolutiva entre una CNN 2D y una CNN 3D [11].

3. Función de activación: Se aplica una función no lineal a los resultados con el fin de aprender representaciones más complejas.

4. Capa de pooling: Disminuye el tamaño del volumen reduciendo la resolución espacial y temporal de los datos, esto provoca que a su vez disminuya el número de parámetros y cálculos necesarios. Permite evitar el sobreajuste de los datos y reduce el tiempo de cómputo.
5. Capa completamente conectada: Recibe las características aplanadas o compactas y representativas de las capas anteriores y las conecta a todas las neuronas de las siguientes capas.
6. Capa de salida: Produce la salida del modelo, generalmente se utiliza una función de activación softmax, la cual normaliza la salida, para calcular las probabilidades de que el volumen pertenezca a cada una de las clases objetivo.

El entrenamiento de modelos de Deep Learning sigue un ciclo estándar, donde primero, se inicializan los parámetros de la red, como los pesos y umbrales de las neuronas, junto con el optimizador que ajustará estos parámetros. A continuación, durante cada iteración, los datos de entrada se propagan hacia adelante a través de la red para generar predicciones, calculándose la pérdida como medida del error entre estas predicciones y los valores esperados. Este error luego se retropropaga por la red para actualizar los pesos y minimizar la pérdida, utilizando algoritmos de optimización. Durante el proceso, se monitorea el rendimiento del modelo en un conjunto de validación independiente para prevenir sobreajuste al conjunto de entrenamiento. Finalmente, se suelen almacenar puntos de control que permiten guardar el estado del modelo con el mejor rendimiento observado en base al conjunto de validación, asegurando su recuperación para evaluaciones futuras [13].

2.1.5. Métricas de evaluación

Existen diversos métodos para evaluar los modelos de Deep Learning a partir de las predicciones que realizan. Estas predicciones pueden clasificarse en cuatro categorías, tal como se muestra en la Tabla 2.1. A partir de esta matriz de confusión, es posible calcular distintas métricas de desempeño, como precisión, sensibilidad y coeficiente de Dice, las cuales se detallan a continuación [14].

Tabla 2.1: Matriz de Confusión.

		Predicción	
		Positivo	Negativo
Valor Real	Positivo	Verdadero Positivo (VP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadero Negativo (VN)

- **Precisión:** Mide la proporción de verdaderos positivos predichos con respecto al total de predicciones positivas realizadas, es decir, de todos los casos que el modelo identificó como positivos, cuántos realmente lo son. Se representa matemáticamente por la Ecuación 2.1. Una alta precisión indica que el modelo comete pocos falsos positivos, es decir, cuando predice la presencia de la clase de positiva, generalmente acierta.

$$\text{Precisión} = \frac{VP}{VP + FP} \quad (2.1)$$

- **Sensibilidad:** Corresponde a una métrica que mide la proporción entre verdaderos positivos precedidos con respecto al total de casos positivos en la muestra, se representa matemáticamente por la Ecuación 2.2. Una alta sensibilidad significa que el modelo es capaz de detectar la mayoría de los casos positivos reales, minimizando la cantidad de falsos negativos.

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (2.2)$$

- **Coefficiente de Dice:** Es la métrica de evaluación más usada para validar segmentaciones de imágenes médicas en dos o tres dimensiones [15], ya que indica la similitud entre dos conjuntos, como se representa en la Ecuación 2.3, donde los valores se encuentran en el rango entre 0 y 1, significando este último que los conjuntos son idénticos. Cabe destacar que matemáticamente es equivalente al F1-score, pero aplicado en contexto de segmentación a nivel de píxel.

$$DICE(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} = \frac{2 \cdot VP}{2 \cdot VP + FP + FN} \quad (2.3)$$

2.1.6. Función de Pérdida

Durante el entrenamiento de redes neuronales, las funciones de pérdida tienen como objetivo cuantificar la discrepancia entre las predicciones del modelo y los valores esperados. Esta medida permite optimizar los parámetros internos del modelo, guiando al proceso de aprendizaje para minimizar dicho error y mejorar progresivamente el rendimiento del sistema [14]. Algunas de estas funciones de pérdida usadas en tareas de segmentación son:

- **Dice Loss:** Se basa en el coeficiente de Dice previamente mencionado, midiendo la superposición entre las predicciones generadas por el modelo y sus respectivas etiquetas, de acuerdo con la Ecuación 2.4.

$$L_{Dice} = 1 - DICE(A, B) \quad (2.4)$$

Donde los conjuntos A y B representan las predicciones y las etiquetas, respectivamente. Esta función puede tomar valores entre 0 y 1, donde valores cercanos a cero indican un mejor desempeño, ya que reflejan un coeficiente de Dice cercano a uno, es decir, una superposición casi completa entre ambos conjuntos.

- **Focal Loss:** Fue introducida para abordar el desbalance de clases en problemas de detección, atenuando el impacto de los ejemplos fáciles y poniendo foco en aquellos difíciles de clasificar, visualizada en la expresión de la ecuación 2.5 [16].

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (2.5)$$

Donde p_t corresponde a la probabilidad predicha para la clase verdadera, α balancea los pesos de las clases, afectando la importancia que se le otorga a la clase positiva sobre la negativa, mientras que γ determina cuánto se centra el algoritmo en los ejemplos difíciles, reduciendo progresivamente la contribución de aquellos ejemplos que son clasificados correctamente con

alta confianza. Esto permite que el modelo enfoque su aprendizaje en las muestras donde tiene más incertidumbre o donde tiende a cometer errores.

2.1.7. Data Augmentation

El *data augmentation* es una técnica de regularización ampliamente utilizada en deep learning, cuyo objetivo es aumentar la cantidad y diversidad del conjunto de datos de entrenamiento mediante transformaciones artificiales aplicadas a las muestras originales [13]. Estas transformaciones buscan generar variaciones realistas que preserven la clase o etiqueta asociada a cada muestra, permitiendo al modelo aprender representaciones más generales y evitando que dependa de patrones específicos del conjunto de datos con el cual se entrena.

Entre las transformaciones comúnmente aplicadas en imágenes se encuentran las rotaciones, traslaciones, simetrías (horizontal o vertical), escalados, recortes aleatorios, adición de ruido (como ruido gaussiano), cambios de brillo o contraste, entre otros. Al exponer al modelo a estas variaciones durante el entrenamiento, se incentiva a que aprenda características robustas e invariantes a pequeñas modificaciones en la entrada.

En este contexto, el término *robustez* hace referencia a la capacidad del modelo para mantener un rendimiento estable frente a perturbaciones o variaciones no vistas durante el entrenamiento, generalizando adecuadamente a nuevos datos en los que se evalúe el modelo.

2.1.8. Arquitectura U-Net

Esta red posee su nombre debido a la forma de “U” que tiene su arquitectura, como se observa en la Figura 2.5, en la cual la vía izquierda es de contracción, donde se captura el contexto global de la imagen y sigue la estructura típica de una red neuronal convolucional, la cual consiste en la aplicación repetida de dos convoluciones 3×3 , seguidas de una unidad lineal rectificadora (ReLU), finalizando con una operación “max pooling” 2×2 , duplicando la cantidad de canales de características en cada paso.

Por otra parte, la vía derecha es de expansión y se encarga de localizar de forma precisa en la imagen original la zona correspondiente al contexto obtenido por la ruta anterior. En esta vía, cada paso consiste en primero realizar un “up-sampling” del mapa de características seguido de una convolución 2×2 ; a este proceso se le llama “up-convolution”, lo que divide en la mitad los canales de características. Luego, se realizan dos convoluciones 3×3 seguidas de una ReLU. En la capa final se realiza una convolución 1×1 con el fin de asignar cada vector de características al número de clases deseadas.

Una característica distintiva de la arquitectura U-Net es el uso de conexiones de salto (*skip connections*) entre las capas correspondientes del encoder y el decoder. Estas conexiones permiten transmitir directamente características de baja y media escala del encoder hacia el decoder, preservando detalles espaciales finos que suelen perderse durante el proceso de downsampling. Al

concatenar las características de alta resolución provenientes del encoder con las representaciones de mayor nivel generadas en el decoder, se facilita la localización precisa de las estructuras segmentadas [17].

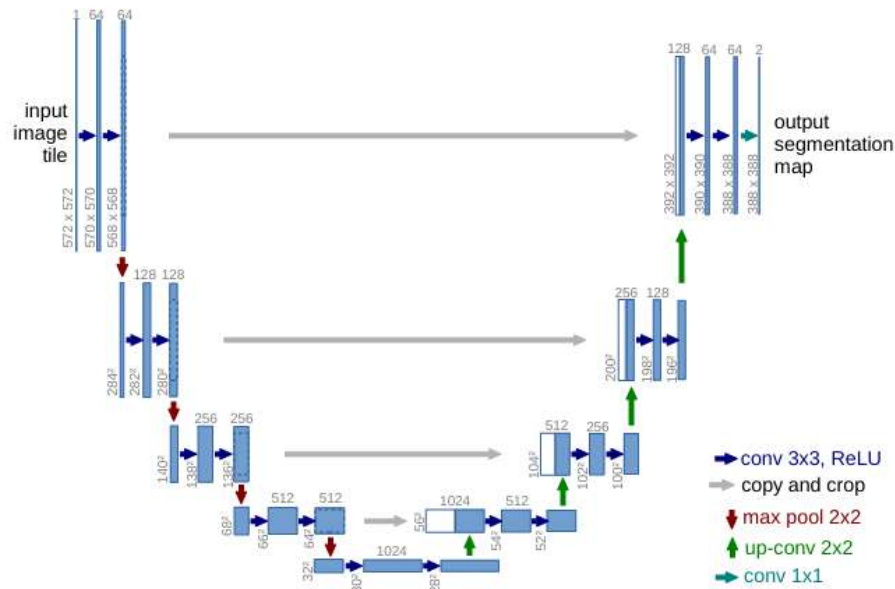


Figura 2.5: Arquitectura de la red U-Net [17].

2.1.9. Segmentación de Imágenes

Una de las tareas fundamentales en el procesamiento de imágenes es la segmentación, que consiste en asignar una etiqueta a cada píxel, delimitando regiones que comparten características similares [14]. En este contexto, las redes neuronales convolucionales tipo U-Net han demostrado un desempeño destacado, principalmente gracias a su arquitectura en forma de ‘U’, que permite combinar de manera eficiente la información de contexto global (extraída en el encoder) con los detalles espaciales locales (recuperados en el decoder mediante las *skip connections*) [17].

Existen tres enfoques principales de segmentación de imágenes:

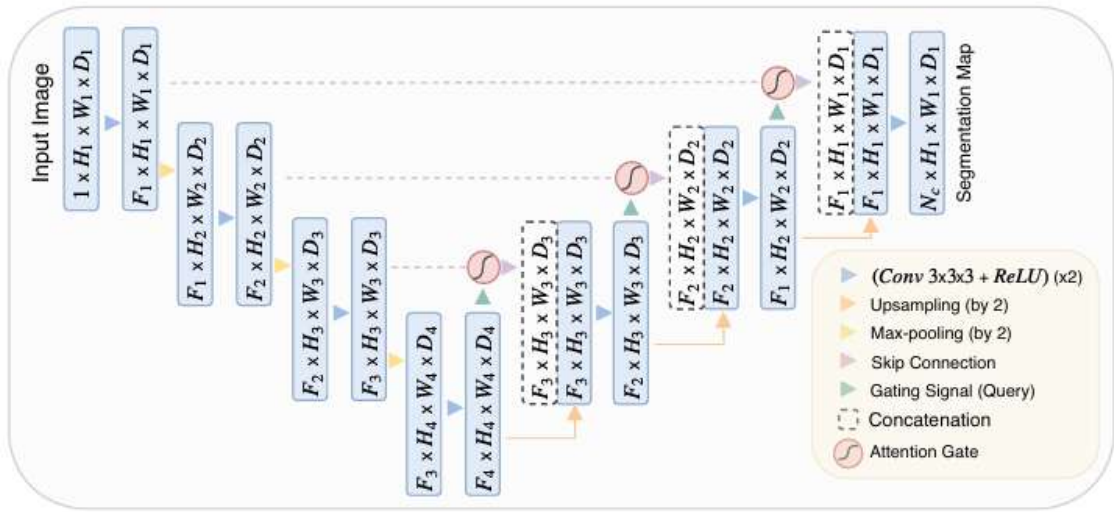
- Segmentación semántica: Asigna a cada píxel la etiqueta de la clase a la que pertenece, sin distinguir entre distintas instancias de la misma clase.
- Segmentación por instancia: No solo identifica la clase, sino que diferencia entre instancias individuales de la misma clase.
- Segmentación panóptica: Combina ambos enfoques anteriores, asignando una clase a cada píxel e identificando simultáneamente cada instancia individual.

2.1.10. Arquitectura Attention U-Net

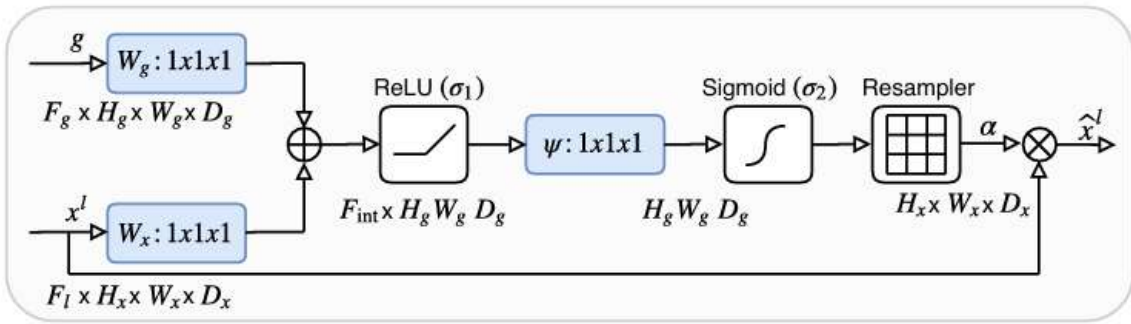
La Attention U-Net es una extensión de la arquitectura U-Net que incorpora módulos de atención (*attention gates, AGs*) para mejorar la segmentación en imágenes médicas, especialmente cuando los objetos de interés presentan gran variabilidad en forma y tamaño. Estos módulos permiten a la red aprender de forma automática a enfocar su atención en regiones relevantes de la imagen, suprimiendo activaciones en áreas irrelevantes.

Cada *attention gate* recibe como entrada una señal de activación local proveniente del encoder y una señal de *gating* global proveniente de capas más profundas. Internamente, ambos tensores son alineados mediante transformaciones lineales y combinados, pasando por una función no lineal (como ReLU) y una activación sigmoide para generar un mapa de atención espacial. Este mapa asigna pesos entre 0 y 1 a cada ubicación, modulando la activación original mediante una multiplicación punto a punto. Así, el modelo amplifica las regiones relevantes y suprime las irrelevantes antes de transmitir las al decoder, lo que permite focalizar los recursos de la red en estructuras anatómicas clave sin introducir una carga computacional significativa.

Los módulos de atención se integran en las conexiones tipo *skip connections* de la U-Net original, filtrando las características propagadas desde las capas encoder hacia las capas decoder, permitiendo conservar únicamente la información relevante para la tarea de segmentación. El diagrama explicativo de esta arquitectura se puede observar en la Figura 2.6. Este enfoque no introduce un aumento significativo en la carga computacional y puede entrenarse de manera similar a una U-Net estándar. Experimentalmente, se ha demostrado que Attention U-Net mejora la sensibilidad y precisión de las predicciones, logrando resultados consistentes en diversos conjuntos de datos, incluso cuando el tamaño del conjunto de entrenamiento es limitado [18].



(a) Diagrama de la arquitectura Attention U-Net



(b) Diagrama de los módulos de atención.

Figura 2.6: Diagrama explicativo de la arquitectura Attention U-Net y sus módulos de atención [18].

2.1.11. Vision Transformers

La arquitectura Transformer, visible en la Figura 2.7 fue introducida originalmente para tareas de procesamiento de lenguaje natural [19], donde la rama de la derecha, el decodificador, es utilizado principalmente para generación de secuencias, como generación de texto. Sin embargo, el enfoque de este estudio es en imágenes médicas, donde se utiliza un modelo denominado Vision Transformer que hace uso sólo la rama izquierda de la arquitectura presente en la Figura 2.7, el codificador, dicho modelo se observa en más detalle en la Figura 2.8.

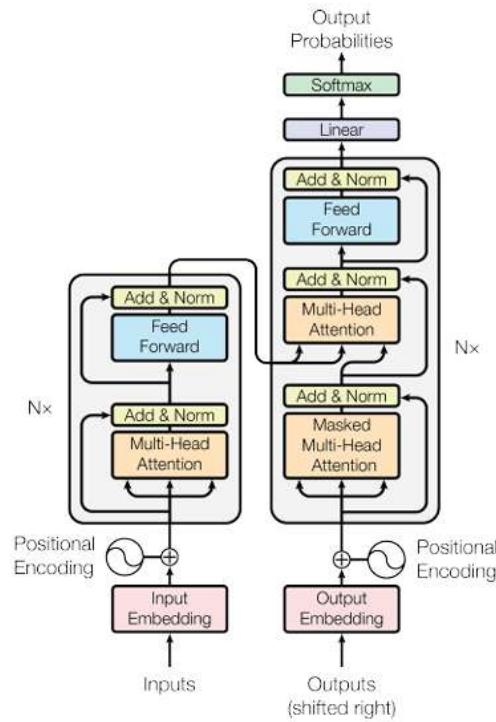


Figura 2.7: Arquitectura Transformers. [19]

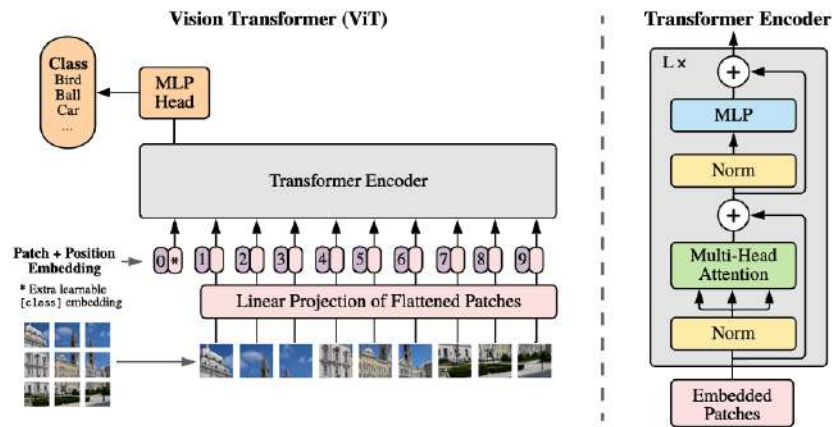


Figura 2.8: Modelo Vision Transformer. [20]

En el modelo Vision Transformer, la información de entrada que se utiliza es de una imagen, sin embargo, dicha información debe ser ingresada como una secuencia, es por esto que la imagen se divide en parches y se representa cada uno como vectores, agregándole un “token” que irá recibiendo información a partir de los demás elementos del vector, además se agrega información sobre la posición que posee en la imagen cada parche [20] debido a que el codificador no tiene una comprensión inherente del orden espacial de estos.

El codificador está compuesto por distintos bloques, un bloque de normalización que se aplica previo a los demás bloques en el cual se normalizan las funciones de activación. Además, existe el bloque de atención multi-cabeza, donde el mecanismo de auto-atención de los transformer corresponde a uno de sus puntos más importantes [21], el cual indica las interacciones que tiene un elemento de la secuencia con el resto, agregándoles esta información y se calcula de la forma que se aprecia en la Ecuación 2.6 [19], donde Q corresponde a la consulta que se está realizando, K son las claves, es decir el elemento que se va a comparar con la consulta, mientras que V es la información que posee dicha clave y finalmente d corresponde a su dimensión. Este cálculo es realizado múltiples veces en el bloque de atención multi-cabeza de forma paralela para así capturar relaciones más complejas [21], donde los pesos que obtiene cada cabeza se concatenan obteniendo así el resultado final.

$$Atencion(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}V\right) \quad (2.6)$$

Por último, el codificador tiene un bloque que contiene un perceptrón multicapa, el cual es una red neuronal completamente conectada con funciones de activación no lineales capaces de aprender relaciones más complejas de información [22]. Posterior a aplicar el codificador, se utiliza el “token” mencionado, que posee información que representa a la imagen gracias a todo el proceso detallado y se ingresa a un perceptrón multicapa para realizar la clasificación.

2.1.12. Swin Transformers

El Swin Transformer (*Shifted Window Transformer*) es una arquitectura de tipo Transformer jerárquica, propuesta para tareas de visión por computador, que extiende a los Vision Transformers (ViT) tradicionales, incorporando mecanismos de atención local eficientes, preservando al mismo tiempo la capacidad de capturar dependencias globales en la imagen. A diferencia del ViT, que opera sobre toda la imagen de forma plana, el Swin Transformer construye representaciones a múltiples escalas, lo que permite un procesamiento más eficiente y efectivo en imágenes de alta resolución.

La arquitectura se basa en la división de la imagen en pequeñas ventanas no superpuestas (windows), sobre las cuales se aplica atención auto-regresiva de forma local. Posteriormente, se introduce un desplazamiento (shift) en la disposición de las ventanas entre capas consecutivas, permitiendo así que la red capture relaciones entre ventanas adyacentes. Este mecanismo, denominado *Shifted Window Attention*, reduce significativamente el costo computacional con respecto a aplicar atención global, escalando linealmente con el tamaño de la imagen en lugar de cuadráticamente.

El Swin Transformer está organizado en etapas jerárquicas, donde la resolución espacial de las características se reduce progresivamente mediante operaciones de *patch merging*, aumentando simultáneamente el número de canales. Esto permite construir representaciones enriquecidas a distintas escalas, siguiendo un esquema similar a arquitecturas convolucionales clásicas (como U-Net), pero empleando bloques basados en atención en lugar de convoluciones.

Cada bloque de la arquitectura contiene capas de normalización (LayerNorm), bloques de atención multi-cabeza sobre ventanas desplazadas, y módulos feed-forward tipo MLP (perceptrones multicapa) que permiten modelar relaciones no lineales. El esquema general de la arquitectura Swin Transformer se muestra en la Figura 2.9, donde se aprecia tanto la estructura jerárquica como el detalle interno de los bloques con atención por ventanas desplazadas [23].

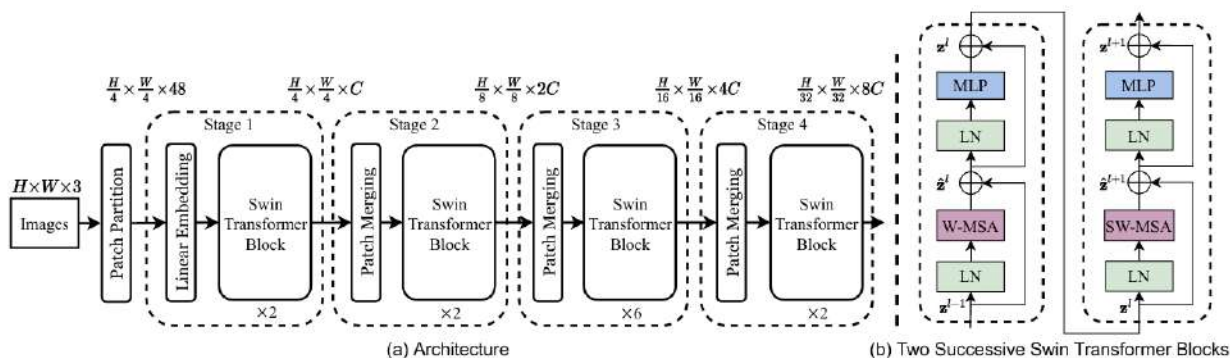


Figura 2.9: Esquema general del Swin Transformer [23].

2.1.13. Arquitectura Swin U-Net

Swin U-Net es un modelo de segmentación diseñado específicamente para imágenes médicas, que combina la estructura en forma de “U” característica de U-Net con los bloques jerárquicos basados en Swin Transformers. Esta arquitectura reemplaza los bloques convolucionales tradicionales por módulos de atención jerárquica, permitiendo capturar dependencias a largo alcance y detalles finos de las imágenes de forma eficiente [24].

El encoder de Swin U-Net sigue un esquema jerárquico, compuesto por múltiples etapas que aplican atención por ventanas desplazadas (*shifted window attention*) y operaciones de *patch merging*, lo que permite reducir progresivamente la resolución espacial mientras aumenta la dimensionalidad de las características. Gracias a esto, se obtiene una representación multiescala enriquecida de la imagen, similar al comportamiento de arquitecturas como U-Net, pero con el beneficio adicional de la auto-atención multi-cabeza.

Por su parte, el decoder utiliza bloques simétricos que aplican operaciones de *patch expanding* para recuperar la resolución original, acompañado de *skip connections* que transmiten características de las etapas encoder correspondientes. Estas conexiones permiten integrar información de contexto global y detalles locales, favoreciendo una segmentación precisa. El esquema general descrito se observa en la Figura 2.10.

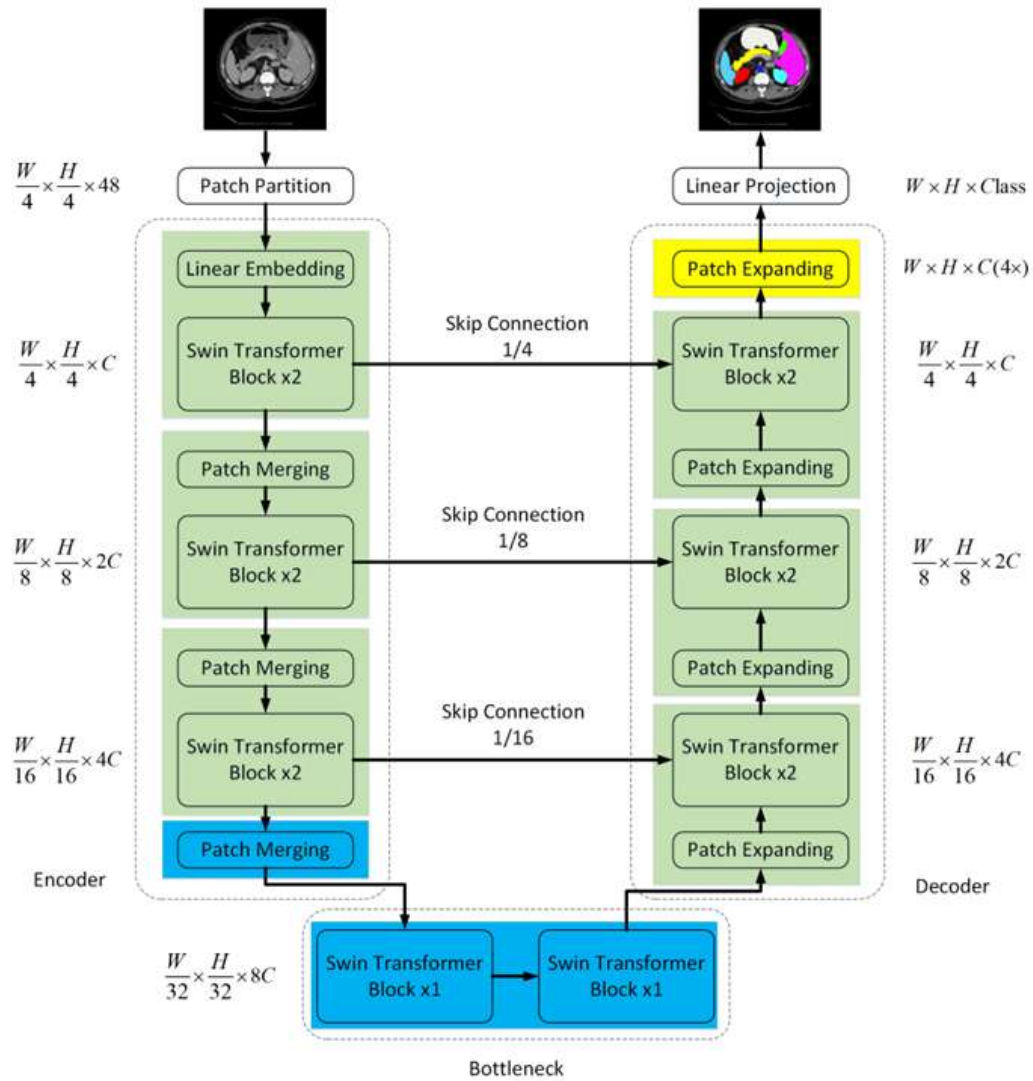


Figura 2.10: Esquema general de la arquitectura Swin U-Net [24].

2.1.14. Arquitectura Swin UNetR

La Swin UNetR es una arquitectura de segmentación semántica desarrollada para imágenes médicas, que combina los Swin Transformers con la estructura tipo U-Net, pero utilizando un diseño inspirado en UNetR (*UNet Transformer*) [25]. A diferencia de Swin U-Net, que tiene una arquitectura jerárquica simétrica, Swin UNetR integra un extractor de características basado en Swin Transformer únicamente en el encoder, manteniendo un decoder clásico con bloques convolucionales.

El encoder basado en Swin Transformer divide la imagen de entrada en *patches* y extrae representaciones jerárquicas aplicando bloques de atención por ventanas desplazadas, capturando así dependencias de largo alcance y estructuras multiescala. Las representaciones intermedias del encoder se conectan al decoder mediante múltiples *skip connections*, lo que permite integrar infor-

mación de bajo y alto nivel para una segmentación precisa.

El decoder consiste en bloques convolucionales y operaciones de *upsampling* progresivo, que recuperan la resolución espacial original. La combinación de un encoder puramente basado en Transformers y un decoder convolucional permite a Swin UNetR beneficiarse tanto de la capacidad de atención global como de la eficiencia y precisión espacial de las convoluciones locales.

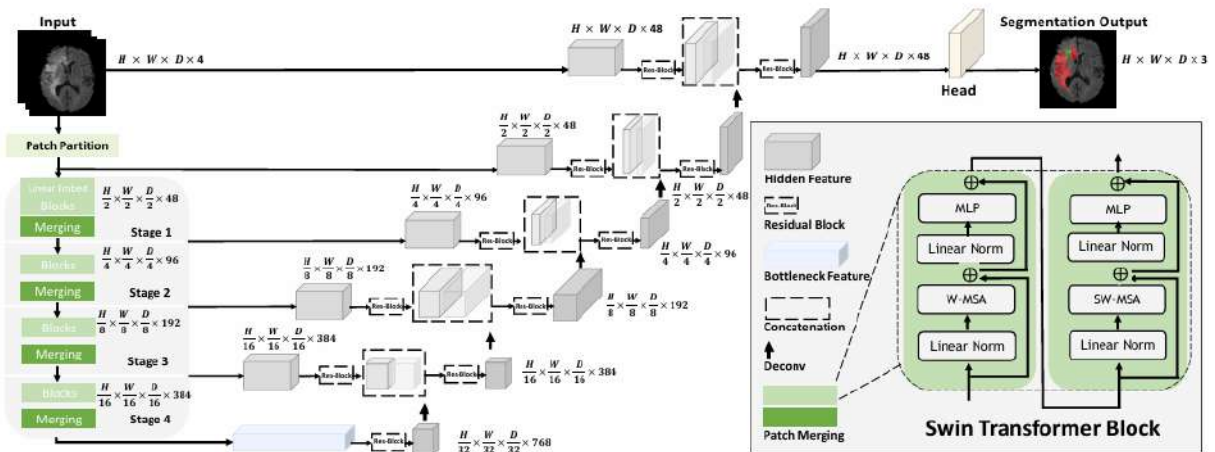


Figura 2.11: Esquema general de la arquitectura Swin UNetR [25].

2.2. Estado del arte

El presente trabajo de título nace a partir del trabajo previo realizado por M. Flores y C. Ramírez en su tesis profesional [26], en donde desarrollan un algoritmo que cuantifica las BAC utilizando deep learning en imágenes médicas de tomosíntesis mamarias obtenidas a partir de una muestra de la base de datos *Breast Cancer Screening - Digital Breast Tomosynthesis (BCS-DBT)* [27]. En esta tesis se desarrolla un modelo basado en DenseNet121 para clasificar los volúmenes con presencia de BACs, y dos modelos de segmentación basados en la arquitectura U-Net, uno para BAC y otro para vasos sanguíneos, obteniendo resultados bajo lo óptimo deseado, por lo que se consideran como modelos no confiables en los cuales se pueden realizar diversas mejoras para aumentar su rendimiento.

Los estudios previos encontrados en la literatura sobre el uso de inteligencia artificial para segmentar y/o cuantificar BACs en imágenes médicas de tomosíntesis mamaria son escasos, ya que la mayoría de los trabajos en este tipo de imágenes se centran en la detección de cáncer de mama. Sin embargo, existe un estudio reciente de Wang *et al.* [28] que se enfoca en la segmentación de BACs en mamografías digitales y tomosíntesis mamaria, y desarrolla además un modelo de corrección automática de etiquetas para diferenciar las BACs reales de otras estructuras como tejidos fibroglandulares densos. No obstante, este estudio se centra exclusivamente en la mejora de

la segmentación y no aborda la cuantificación clínica de BACs, que es uno de los objetivos de esta memoria. Además, reporta métricas de desempeño basadas en el coeficiente de Dice ($\approx 0,48$), estos resultados fueron obtenidos utilizando una base de datos propia y privada, por lo que no son directamente comparables con los resultados obtenidos en este trabajo al ser bases de datos distintas.

Por otra parte, al considerar imágenes médicas de mamografías se logran encontrar diversos estudios de segmentación y/o cuantificación de calcificaciones arteriales mamarias en los que se utilizan diversos modelos y métricas de evaluación. En [29] se utiliza un modelo VGG16 preentrenado como punto de partida para desarrollar una CNN que es evaluada con las métricas de exactitud (95%) y área bajo la curva de ROC (0,94), lo cual se puede notar que son métricas bastante altas, sin embargo propensas a errores debido a que las mamografías disminuyen su sensibilidad en mamas densas, como ha sido mencionado anteriormente, esto se busca evitar usando las imágenes de tomosíntesis. Casos similares ocurren con [30] en que divide las mamografías en parches más pequeños para mantener la resolución y así mejorar la detección de calcificaciones, luego compara los modelos CNN ResNet18, ResNet50 y ResNet101 para realizar una clasificación binaria en cada parche, siendo ResNet50 el modelo con mejor rendimiento para las métricas de evaluación utilizadas, que fueron exactitud (96,4%), sensibilidad (76,2%), especificidad (98,1%) y área bajo la curva ROC (0,96), nuevamente se ve un gran rendimiento, pero es propenso a disminuir debido a las densidades mamarias y su efecto en las mamografías, lo cual se puede evitar utilizando tomosíntesis mamaria, pues disminuye la superposición de estructuras.

Cabe mencionar que durante las etapas finales de este trabajo se identificó el estudio de *Maguire et al.* (2025) [31], que propone un enfoque de segmentación para calcificaciones arteriales mamarias utilizando la arquitectura DeepLabv3+ aplicado a mamografías de tipo FFDM (Full-Field Digital Mammography). Una de las principales particularidades de este estudio es la incorporación del length-based Dice score, una métrica diseñada para evaluar la coincidencia entre líneas centrales de las BAC segmentadas, en lugar de comparar únicamente las áreas segmentadas como lo hace el coeficiente de Dice tradicional. Esto permite capturar de forma más precisa errores relacionados con la longitud y continuidad de las calcificaciones, penalizando especialmente las discrepancias en su extensión a lo largo de los vasos. Sin embargo, dado que este estudio es reciente y fue identificado en etapas muy avanzadas del desarrollo de esta memoria, no alcanzó a ser considerado en la creación de la metodología ni influyó en las decisiones experimentales tomadas.

Por último, acotar que no se encuentra en la literatura un modelo que utilice vision transformers para segmentar o cuantificar BACs en imágenes médicas de tomosíntesis mamaria.

Capítulo 3

Metodología

3.1. Base de datos

Para el desarrollo del presente trabajo se utiliza una muestra de la base de datos *Breast Cancer Screening - Digital Breast Tomosynthesis (BCS-DBT)* [27], la cual contiene 5060 conjuntos de imágenes de tomosíntesis mamaria, correspondientes a pacientes con una edad promedio de 50 años y una desviación estándar del 11 %. Esta base se encuentra disponible en el repositorio *The Cancer Imaging Archive (TCIA)*.

A partir de esta base de datos, se construyó un conjunto de entrenamiento compuesto por 49 estudios con presencia de BAC, según lo descrito en [26], el cual se mantiene para el desarrollo de este trabajo. Cada estudio incluye su respectivo volumen mamario y tres máscaras de segmentación diferentes:

- BAC: corresponde a la etiqueta de las calcificaciones arteriales dentro de la mama.
- Arteria: representa el trazado completo de la arteria en la que se encuentran las calcificaciones.
- Calcificación: contiene las calcificaciones no arteriales que pueden estar presentes en la mama.

Estas etiquetas fueron realizadas manualmente por estudiantes del Departamento de Tecnología Médica de la Universidad de Chile. El mismo formato se utiliza para el conjunto de validación, el cual está compuesto por 15 estudios.

3.2. Equipamiento y Herramientas

Para el desarrollo de este trabajo se cuenta con un computador servidor con procesador Intel Core i7-4930K de 6 núcleos, 64 GB de memoria RAM. Este cuenta con la disponibilidad de trabajar con una tarjeta GPU GeForce GTX 980 4 GB y una tarjeta GPU GeForce RTX 4060 Ti 8 GB. Todo el equipamiento está ubicado en el programa de Biología Integrativa del ICBM de la Facultad de Medicina de la Universidad de Chile.

Parte de los experimentos fueron ejecutados en el Laboratorio Nacional de Computación de Alto Rendimiento (NLHPC), específicamente en la partición v100 del clúster Guacolda. Esta partición cuenta con nodos Dell PowerEdge R740, cada uno equipado con dos procesadores Intel Xeon Gold 6152 a 2,10GHz (22 núcleos cada uno), 187GB de memoria RAM y dos GPUs NVIDIA Tesla V100, diseñadas para aplicaciones de alto rendimiento en cómputo científico e inteligencia artificial. El uso de este servidor fue necesario debido a la alta demanda computacional de algunos modelos, que superaba las capacidades del entorno descrito anteriormente.

Durante el desarrollo de los algoritmos se utilizan diversas librerías de Python, siendo MONAI la principal (versión 1.5.0), ya que está específicamente diseñada para el procesamiento de imágenes médicas [32]. Esta librería permite trabajar directamente con archivos en formato .nii.gz, que contienen los volúmenes mamarios y sus respectivas etiquetas, además de proveer implementaciones de múltiples arquitecturas de redes neuronales comúnmente empleadas en tareas de segmentación semántica.

Adicionalmente, se emplea PyTorch 2.6.0 para el entrenamiento de las redes neuronales convolucionales, NumPy 1.26.4 para la creación de una nueva base de datos en forma de arreglos comprimidos, lo que permite agilizar el procesamiento al cargar las imágenes, Matplotlib 3.10.1 para la visualización de resultados, Scikit-Learn 1.6.1 para evaluar los modelos cuantitativamente, y Pandas 2.2.3 para la tabulación y análisis de métricas de evaluación.

3.3. Procedimiento

Inicialmente se trabaja con las imágenes tridimensionales provenientes de la base de datos, las cuales representan volúmenes de mama compuestos por ancho, alto y número de cortes axiales. A estos volúmenes se les aplican transformaciones que normalizan su rango de intensidades, de manera consistente con el trabajo previo de tesis [26]. Posteriormente, cada volumen es reescalado a una resolución fija de (256, 256, 32), ya que trabajar con el volumen completo (aproximadamente (1900, 2600, 67)) supera la capacidad de memoria disponible durante el entrenamiento. Destacar también que, exclusivamente en el conjunto de entrenamiento, se aplican técnicas de *data augmentation* de forma aleatoria sobre algunos volúmenes. Estas incluyen simetría axial, adición de ruido gaussiano, y rotaciones bidimensionales aplicadas corte a corte, con el fin de aumentar la robustez del modelo.

Una vez generados los volúmenes preprocesados, estos se utilizan como conjunto de entrenamiento para una red U-Net tridimensional, cuyo esquema se presenta en la Figura 3.1. La arquitectura implementada consta de cuatro niveles de profundidad, con canales progresivos de 32, 64, 128 y 256, y utiliza bloques residuales con normalización por lotes. Las operaciones de downsampling se realizan con strides de 2 en cada dimensión, y la red trabaja con entradas y salidas de un solo canal. El entrenamiento se configura para un total de 100 épocas, siguiendo la estructura típica descrita en el Capítulo 2, utilizando el conjunto de validación para monitorear el desempeño del

modelo y guardar automáticamente aquel que alcance la menor pérdida en validación durante el proceso. La función de pérdida seleccionada es *Dice Loss*, dado que esta métrica es ampliamente utilizada en tareas de segmentación en imágenes médicas, al estar basada en el coeficiente de Dice, como se ha mencionado previamente, y se utiliza el optimizador Adam, el cual se mantiene a lo largo de todas las redes diseñadas.

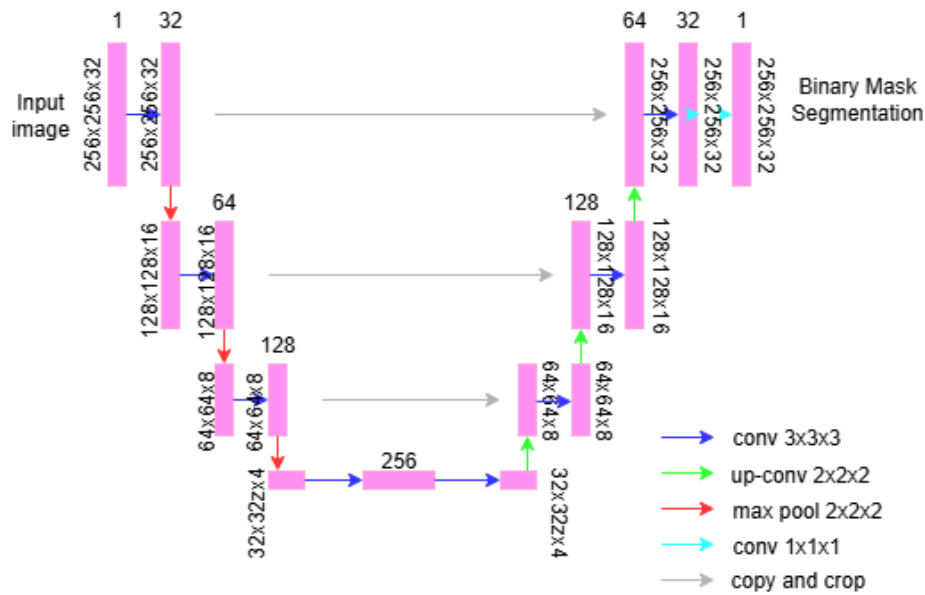


Figura 3.1: Arquitectura U-Net tridimensional diseñada.

Sin embargo, al intentar entrenar este modelo tridimensional, se presenta la dificultad de no poder completar el proceso debido a limitaciones de memoria en el entorno de trabajo. En particular, el tamaño del modelo excede la memoria disponible cuando esta no puede ser utilizada en su totalidad, incluso tras aplicar una reducción considerable en la resolución de las imágenes. Para abordar esta complejidad, se considera que en la tesis anterior relacionada a este trabajo se obtuvieron resultados insatisfactorios al segmentar directamente los volúmenes tridimensionales, alcanzando un coeficiente de Dice del 8% en el conjunto de validación [26].

Además, se recuerda que las etiquetas de segmentación fueron generadas manualmente corte por corte. En consecuencia, se decide construir una nueva base de datos en la que cada corte se trata como una imagen independiente, permitiendo aplicar redes de segmentación bidimensionales. Estas redes requieren menos recursos computacionales, ya que sólo necesitan cargar un corte bidimensional a la vez para realizar las operaciones durante el entrenamiento, en contraste con las redes tridimensionales, que deben cargar y procesar el volumen mamario completo de manera simultánea. A su vez, al replicar de forma más directa la forma en que fue realizada el etiquetado original, se espera que se obtenga un mejor desempeño en la segmentación. En la Figura 3.2 se puede observar una representación de la nueva base de datos en comparación a la original. Se pasa

de una base de datos con imágenes de dimensiones alto por ancho por cantidad de cortes (H, W, C), a un conjunto de imágenes bidimensionales de alto por ancho (H, W). De este modo, se obtiene un total de C imágenes bidimensionales por cada volumen mamario de la base tridimensional, resultando en $N \cdot C$ imágenes en la nueva base a partir de N volúmenes originales.

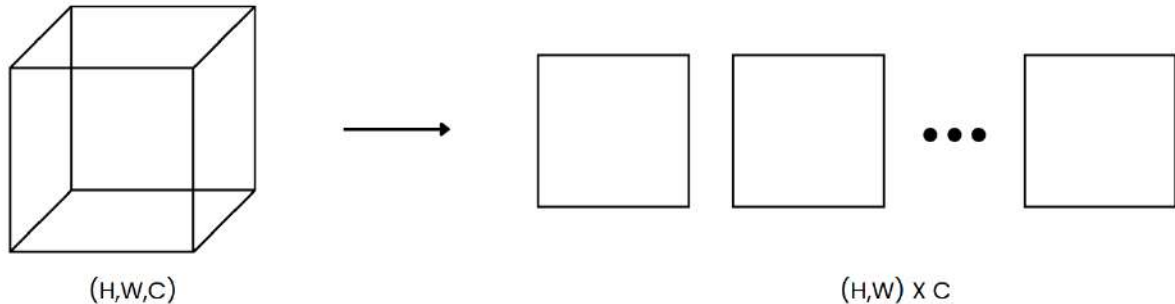
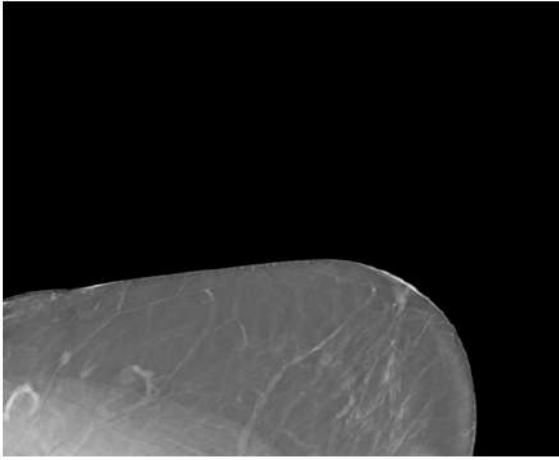


Figura 3.2: Representación del paso de base de datos tridimensional a nueva base de datos bidimensional.

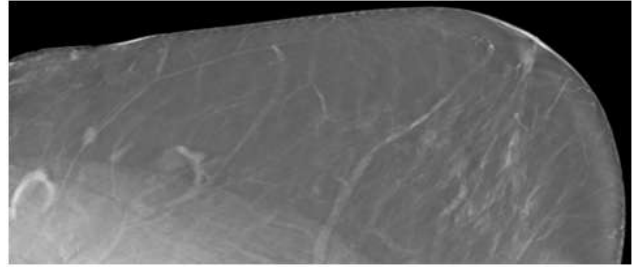
El procesamiento independiente de los cortes, cargándolos directamente desde los archivos .nii.gz, resulta en una lectura y manipulación considerablemente lenta durante la ejecución (3,087 segundos por carga aproximadamente). Por esta razón, se opta por almacenar cada corte oblicuo como un archivo .npz, el cual permite guardar arreglos de NumPy en formato comprimido. Esta decisión implica un mayor uso de espacio en disco, pero mejora significativamente la eficiencia en el acceso y procesamiento de los datos en ejecuciones posteriores (0,015 segundos por carga aproximadamente).

Además, las imágenes originales contienen una gran cantidad de fondo negro, el cual no aporta información relevante para la tarea de segmentación dentro de la región mamaria. Por ello, al construir esta nueva base de datos, se remueve dicho fondo y se almacena únicamente la región que corresponde anatómicamente a la mama. Este recorte se aplica tanto a las imágenes como a las etiquetas asociadas. En particular, se almacenan dos tipos de etiquetas: una que representa las calcificaciones arteriales mamarias y otra correspondiente a las arterias que las contienen.

Un ejemplo de estos cortes se muestra en la Figura 3.3, donde se observa cómo la eliminación del fondo negro permite reducir considerablemente las dimensiones de la imagen, optimizando el uso de almacenamiento y disminuyendo los tiempos de cómputo en etapas posteriores.



(a) Corte original de la base de datos tridimensional.

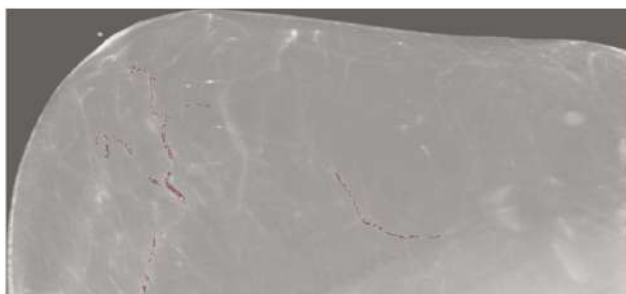


(b) Corte sin fondo almacenado en la base de datos bidimensional.

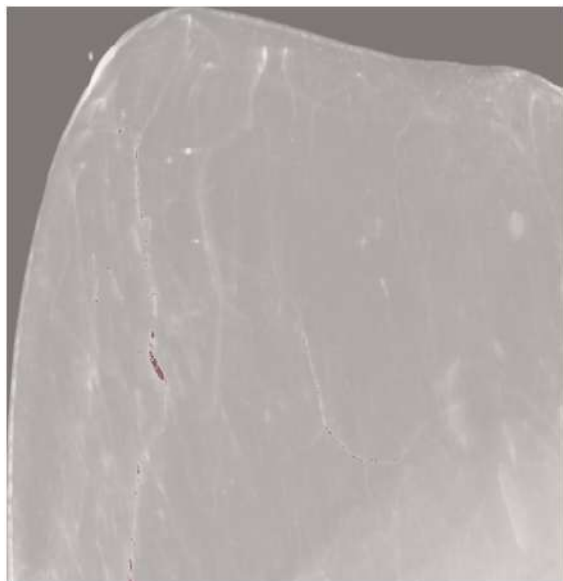
Figura 3.3: Representación de los cortes sin fondo almacenados en la base de datos bidimensional.

Dado que ahora se procesan cada uno de los cortes de forma independiente para su posterior entrenamiento, se decide no aplicar transformaciones de *data augmentation*. Esta decisión busca preservar una estructura espacial coherente entre los distintos cortes de una misma mama. A diferencia del enfoque anterior, donde las transformaciones se aplicaban al volumen tridimensional completo, en este caso se privilegia la consistencia anatómica entre cortes adyacentes.

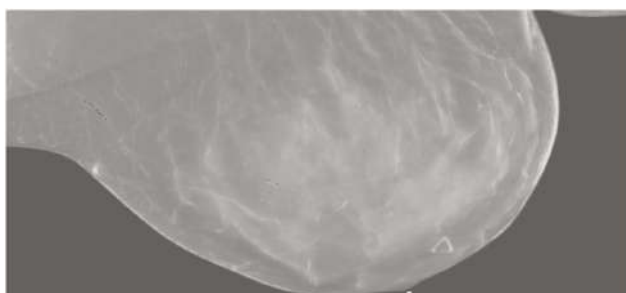
Con el objetivo de reducir el uso de memoria durante la ejecución del proceso, inicialmente se decide mantener el reescalado de las imágenes a una resolución de 256×256 píxeles. Sin embargo, se observa que esta reducción provoca la pérdida significativa de información en las etiquetas de las calcificaciones arteriales, e incluso su completa eliminación en algunos casos. Esta situación se ejemplifica en la Figura 3.4, donde en la imagen 3.4.b se aprecia como parte importante de la etiqueta es truncada, mientras que en la imagen de la Figura 3.4.d, la etiqueta desaparece por completo tras el reescalado.



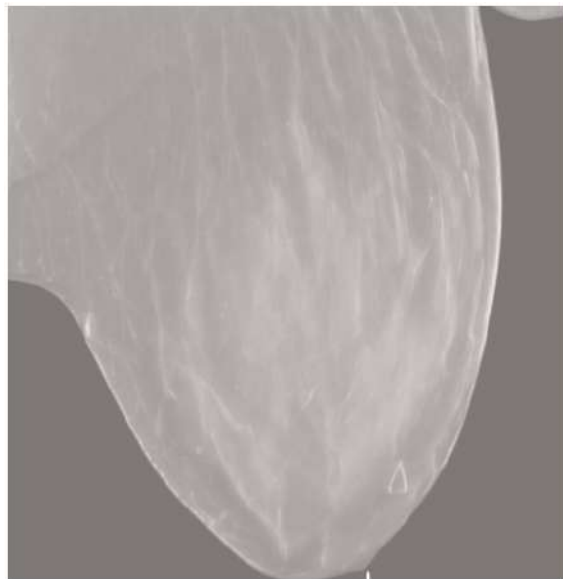
(a) Corte de mama con mucha calcificación.



(b) Etiqueta de calcificación truncada al redimensionar.



(c) Corte de mama con poca calcificación.



(d) Etiqueta de calcificación nula al redimensionar.

Figura 3.4: Errores en las etiquetas al redimensionar los cortes de mama en 256×256 píxeles.

Dado que estos ejemplos pertenecen al conjunto de validación, la pérdida de etiqueta afecta negativamente la evaluación del modelo, impidiendo una medición fiel de su desempeño real. Además, este mismo reescalado se aplica al conjunto de entrenamiento, lo que conlleva que el modelo aprenda a partir de etiquetas incompletas o, en algunos casos, inexistentes. Como consecuencia, se decide aumentar la resolución de entrada a 1024×1024 píxeles, con el fin de preservar la información anatómica relevante y asegurar una segmentación precisa durante el entrenamiento y la evaluación, pese a consumir más recursos durante su ejecución.

Una vez definido el preprocesamiento de los conjuntos utilizados para el entrenamiento, se pro-

pone aprovechar las dos etiquetas disponibles para diseñar un modelo de segmentación en cascada. En este enfoque, la imagen de entrada se procesa inicialmente mediante una red U-Net, análoga a la versión tridimensional previamente descrita pero adaptada para trabajar en dos dimensiones, cuya tarea es segmentar las arterias que contienen calcificaciones.

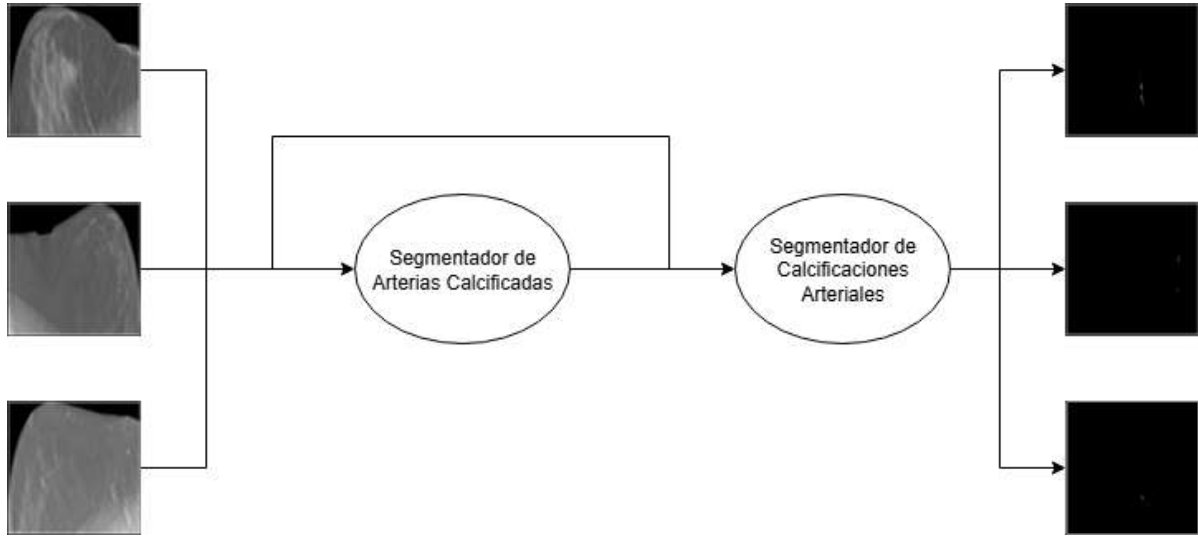
Posteriormente, las segmentaciones generadas por esta primera red se utilizan como entrada adicional para una segunda red, en este caso una Attention U-Net bidimensional, configurada con los mismos canales por capa (32, 64, 128 y 256) y diseñada para recibir dos canales de entrada, la imagen bidimensional original y la predicción de arterias calcificadas. Los mecanismos de atención en las conexiones laterales entre el encoder y el decoder de esta arquitectura, permiten que la red enfoque su atención en las regiones de mayor relevancia anatómica para la segmentación de calcificaciones arteriales, abarcando los problemas que genera el desbalance entre BACs y fondo en el volumen mamario. De esta forma, se busca guiar la predicción de BACs hacia las zonas donde se ha identificado su posible existencia. Además, incluye regularización mediante *dropout*, con tal de evitar su sobreajuste al conjunto de entrenamiento y mejorando la generalización del modelo.

El esquema general del modelo en cascada propuesto, junto con la arquitectura de la Attention U-Net, se muestra en la Figura 3.5, donde se aprecia que el segmentador de calcificaciones arteriales retorna máscaras binarias a partir de los cortes bidimensionales de mama y de las predicciones realizadas por el segmentador de arterias calcificadas.

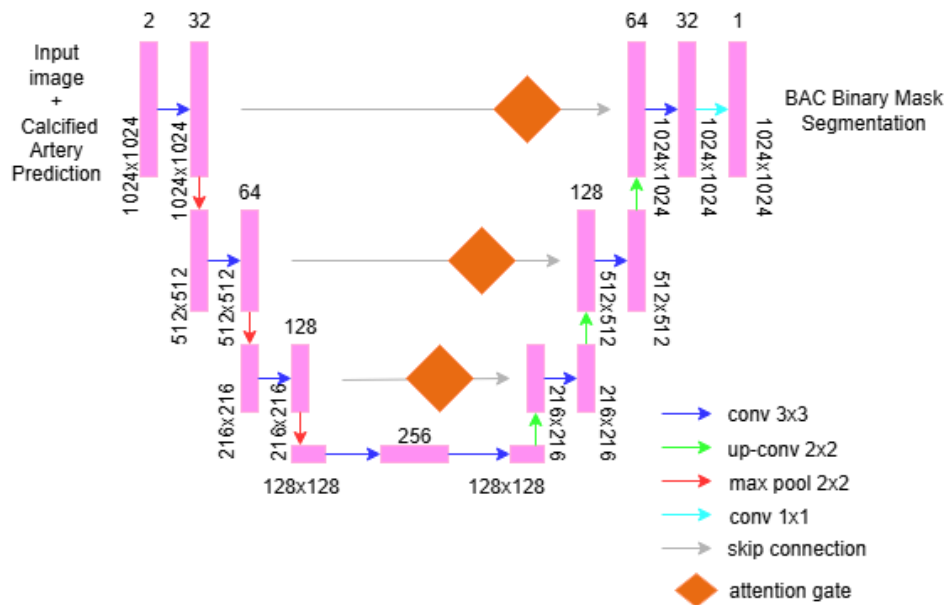
Aunque el modelo se entrena de forma independiente sobre cada corte bidimensional, se propone una estrategia de evaluación global que considera todos los cortes de una misma mama en conjunto exclusivamente para la etapa de validación. Para ello, se combinan todas las segmentaciones generadas durante la validación en una única imagen de referencia, donde se marca la presencia de estructuras a lo largo del eje *Z*, sin considerar su profundidad específica. De igual manera, se construye una máscara compuesta a partir de las etiquetas reales. Luego, se calcula un único coeficiente de Dice por mama, lo que permite evaluar de forma más robusta la capacidad del modelo para detectar las estructuras de interés en general, sin penalizar por omisiones en cortes individuales. Este procedimiento se utiliza únicamente para seleccionar el mejor modelo durante el entrenamiento, mientras que el aprendizaje del modelo se realiza corte a corte de manera independiente.

Ambas redes del modelo en cascada se entrenan durante 200 épocas. Además de diferenciarse en su arquitectura, una basada en U-Net y la otra en Attention U-Net, también emplean funciones de pérdida distintas. En el caso del segmentador de arterias con calcificaciones, se utiliza la función de pérdida *Dice Loss*, por las mismas razones expuestas anteriormente en la red tridimensional. En cambio, para el segmentador de calcificaciones arteriales, se opta por una combinación de *Dice Loss* y *Focal Loss* (*Dice Focal Loss*), ya que esta última se utiliza en situaciones de fuerte desbalance de clases, como ocurre en este caso, donde la cantidad de píxeles correspondientes a calcificaciones es muy pequeña en relación con el tamaño total de la mama, los parámetros α y γ de esta función de pérdida se eligen en base a su interpretación teórica. Se utiliza $\alpha = 0,8$ con el

objetivo de asignar un mayor peso a la clase positiva (calcificación), de modo que el modelo penalice con mayor severidad los errores en los que se predice fondo en lugar de calcificación. Por otro lado, se establece $\gamma = 2,0$, ya que un valor más alto de este parámetro permite enfocar la pérdida en los ejemplos más difíciles, reduciendo la influencia de los casos que el modelo predice con alta confianza.



(a) Diagrama modelo en cascada



(b) Arquitectura Attention U-Net bidimensional diseñada.

Figura 3.5: Modelo en cascada diseñado.

Una vez tabulados los resultados del modelo en cascada, se decide evaluar la segmentación de calcificaciones arteriales utilizando una única red, sin el paso intermedio de segmentar las arterias calcificadas. Si bien el enfoque en cascada tenía un fundamento lógico, guiar la segmentación de

calcificaciones a partir de las arterias que las contienen, en la práctica las segmentaciones arteriales obtenidas no fueron suficientemente precisas, lo que probablemente afecta negativamente el rendimiento del segundo segmento del modelo. Por ello, se plantea que un modelo que reciba directamente el contexto completo de la imagen podría obtener mejores resultados. Para esta prueba se entrena una red U-Net análoga a las ya descritas, utilizando los mismos parámetros en la función de pérdida *Dice Focal Loss*. No se utiliza una Attention U-Net en este experimento, debido a que presenta un mayor tiempo de entrenamiento en comparación con la U-Net tradicional, por lo que se prioriza la eficiencia computacional en esta evaluación.

Al observar que el rendimiento mejora al utilizar una única red para la segmentación de calcificaciones arteriales, se decide realizar una serie de experimentos manteniendo fija la arquitectura U-Net, pero variando la función de pérdida. En particular, se exploran distintas configuraciones de la función *Dice Focal Loss*, modificando los valores de los parámetros α y γ , con el objetivo de determinar empíricamente la combinación que entrega el mejor desempeño. Esta estrategia busca complementar el razonamiento teórico con evidencia experimental, identificando la configuración más adecuada para el problema abordado. Los distintos experimentos realizados se resumen en la Tabla 3.1, donde se compara el uso de la función *Dice Loss* frente a *Dice Focal Loss* con diferentes valores de sus hiperparámetros, evaluando si un ajuste fino de estos permite mejorar el rendimiento del modelo.

Tabla 3.1: Experimentos realizados al entrenar una misma U-Net variando la función de pérdida.

	Función de Pérdida	α	γ
Experimento 1	<i>Dice Loss</i>	-	-
Experimento 2	<i>Dice Focal Loss</i>	0,8	2,0
Experimento 3		0,25	2,0
Experimento 4		0,25	1,0

Una vez identificada la configuración de función de pérdida que entrega el mejor desempeño en la arquitectura U-Net, se decide aplicar dicha configuración a la red Attention U-Net previamente definida. El objetivo es comparar el rendimiento de ambas arquitecturas bajo condiciones equivalentes, evaluando si la incorporación de mecanismos de atención ofrece ventajas adicionales en la tarea de segmentación.

Posteriormente, se decide incorporar una arquitectura distinta con el objetivo de comparar su rendimiento con el mejor modelo obtenido hasta el momento. La arquitectura seleccionada corresponde a una Swin UNetR, implementada previamente en MONAI. Debido a su alta demanda computacional, esta red no pudo ser entrenada en el mismo servidor utilizado para los modelos anteriores, por lo que se entrena en el servidor NLHPC previamente descrito.

El encoder está compuesto por múltiples niveles jerárquicos con profundidades (2, 2, 2, 2), don-

de cada nivel incluye bloques de atención con múltiples cabezas ($num_heads = (3, 6, 12, 24)$), lo que permite incrementar progresivamente la capacidad de representación del modelo capturando patrones tanto locales como globales en la imagen. El tamaño de las ventanas de atención se fija en 7×7 píxeles, y se utiliza una dimensión base de características de 96 filtros ($feature_size$). Como técnica de regularización, se aplica una tasa de *dropout* de 0,1 tanto a la salida del perceptrón multicapa (MLP) como a los pesos de atención.

Una vez definidos todos los modelos destinados a segmentar las calcificaciones arteriales en mamas, se decide visitar el modelo de segmentación de arterias calcificadas con el objetivo de utilizar sus predicciones como parte de un proceso de postprocesamiento. En particular, se realiza una intersección entre las máscaras de calcificaciones predichas y las segmentaciones de arterias generadas por dicho modelo, apreciable en la expresión de la ecuación 3.1, con el fin de eliminar posibles falsos positivos que se encuentren fuera del contexto vascular. Esta estrategia asume que las verdaderas calcificaciones arteriales no se verán afectadas negativamente, ya que el modelo de segmentación de arterias muestra una tendencia a la sobresegmentación, lo que otorga un margen de cobertura suficiente para conservar las regiones relevantes.

$$\text{Pixel}(i, j) \in \text{BAC} \wedge \text{Arteria Calcificada} \rightarrow \text{Pixel}(i, j) = \text{BAC} \quad (3.1)$$

Como fue mencionado previamente, el principal problema del modelo segmentador de arterias calcificadas es su tendencia a la sobresegmentación, generando máscaras que incluyen arterias en general, en lugar de limitarse exclusivamente a aquellas que contienen calcificaciones. Para abordar esta limitación, se incorporan nuevas etiquetas correspondientes a arterias sin calcificaciones, las cuales fueron generadas manualmente por estudiantes del Departamento de Tecnología Médica de la Universidad de Chile. No obstante, estas etiquetas adicionales están disponibles únicamente para 16 pacientes del conjunto de entrenamiento y 6 del conjunto de validación. A fin de integrarlas al flujo de trabajo existente, se les aplica el mismo proceso de preprocesamiento descrito anteriormente, generando sus respectivos archivos .npz, lo que permite utilizarlas en un contexto bidimensional, procesando cada corte por separado.

A partir de estas nuevas etiquetas se desarrollan dos modelos distintos. El primero tiene como objetivo segmentar todas las arterias, tanto calcificadas como no calcificadas, utilizando como etiqueta la suma de ambas máscaras. El segundo modelo, en cambio, está orientado a la segmentación específica de arterias calcificadas, pero incorporando como entrada adicional la etiqueta de arterias no calcificadas, con el propósito de que el modelo aprenda no solo a identificar regiones con calcificaciones, sino también a descartar aquellas donde se ha verificado la ausencia de estas.

Sin embargo, dado que menos de la mitad de las imágenes del conjunto de datos poseen esta nueva anotación, se opta por entrenar este segundo modelo únicamente para obtener pesos iniciales que incorporen dicha información contextual. Posteriormente, estos pesos se utilizan para iniciali-

zar el segmentador de arterias calcificadas original, que se entrena sobre el conjunto completo. En ambos casos, se mantiene la arquitectura U-Net previamente utilizada para la tarea de segmentación de arterias.

Una vez desarrollados todos los modelos, se procede a diseñar el cuantificador de calcificaciones arteriales. Dado que el trabajo se realiza en imágenes bidimensionales y la evaluación se lleva a cabo unificando las predicciones de todos los cortes en una sola proyección, se reutilizan los modelos de segmentación de arterias calcificadas y de calcificaciones por separado. A partir de sus predicciones, se calcula el área (en número de píxeles) correspondiente a cada una de estas segmentaciones. El indicador propuesto corresponde a la razón entre el área de intersección entre calcificaciones y arterias calcificadas, y el área total de las arterias calcificadas. Este valor, que varía entre 0 y 1, representa la proporción del trayecto arterial afectado por calcificaciones. Mientras más cercano al valor máximo, mayor área de la arteria que se encuentra calcificada, lo que se asocia a un mayor riesgo de desarrollar enfermedades cardiovasculares.

$$\text{Indice de calcificación} = \frac{\text{Area de (BACs} \cap \text{arterias calcificadas)}}{\text{Area total de arterias calcificadas}} \quad (3.2)$$

Capítulo 4

Resultados y Análisis

En este capítulo se presentan los resultados obtenidos a partir de las distintas redes neuronales convolucionales bidimensionales desarrolladas para la segmentación de calcificaciones arteriales en tomosíntesis mamaria. Además, se realiza un análisis cuantitativo y cualitativo de su desempeño, considerando métricas de segmentación relevantes, comparaciones entre funciones de pérdida y arquitecturas, junto con visualizaciones representativas de los casos evaluados.

4.1. Modelo en Cascada

El primer modelo implementado sobre la base de datos creada con los cortes bidimensionales de cada mama corresponde al modelo en cascada, es decir, aquel modelo que en una primera etapa se segmentan las arterias calcificadas y, posteriormente, se utilizan estas predicciones como entrada para la segmentación de las calcificaciones arteriales. En la Tabla 4.1 se presenta una comparación de los coeficientes de Dice obtenidos en el conjunto de validación. Cabe destacar que estas evaluaciones consideran las máscaras combinadas de todos los cortes de cada mama en una única proyección bidimensional mediante proyección de máxima intensidad (*Maximum Intensity Projection*, MIP), técnica que calcula, para cada píxel en la imagen 2D resultante, el valor máximo entre los voxeles correspondientes a lo largo del eje Z, es decir, no se toma en cuenta la posición específica de las predicciones a lo largo de dicho eje [33].

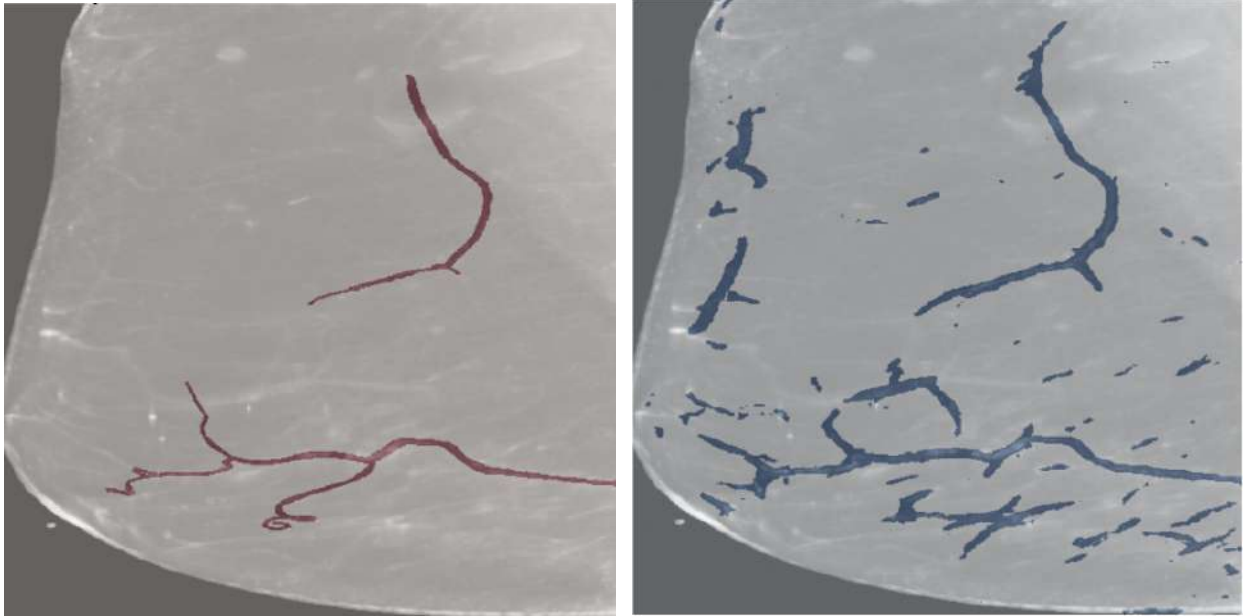
Tabla 4.1: Comparación del coeficiente de Dice del modelo en cascada realizado con los resultados de la tesis previa.

Volumen	Dice Arteria Tesis	Dice Arteria Modelo	Dice BAC Tesis	Dice BAC Modelo
DBT-P00001_MLOI	0,199	0,413	0,316	0,424
DBT-P00285_MLOI	0,132	0,162	0,112	0,135
DBT-P00303_MLOD	0,209	0,233	0,060	0,004

DBT-P00491_MLOD	0,217	0,188	0,078	0,129
DBT-P00523_MLOD	0,016	0,074	0,000	0,000
DBT-P00688_MLOD	0,194	0,165	0,013	0,068
DBT-P00887_MLOD	0,051	0,127	0,355	0,372
DBT-P01488_MLOD	0,000	0,091	0,000	0,009
DBT-P01762_MLOI	0,088	0,177	0,000	0,442
DBT-P01778_MLOI	0,000	0,069	0,000	0,023
DBT-P01957_MLOD	0,088	0,231	0,000	0,027
DBT-P02244_MLOD	0,046	0,149	0,077	0,132
DBT-P02510_MLOD	0,000	0,150	0,000	0,506
DBT-P04368_MLOD	0,000	0,402	0,000	0,203
DBT-P04396_MLOD	0,249	0,220	0,146	0,034
Promedio \pm Desviación Estándar	0,094 \pm 0,093	0,190 \pm 0,102	0,080 \pm 0,115	0,167 \pm 0,179

Al analizar la Tabla 4.1, se observa que para ambos segmentadores aumenta el coeficiente de Dice promedio obtenido, duplicando aproximadamente su rendimiento en ambos casos. Si bien esta mejora promedio es significativa, se evidencia una alta variabilidad entre los diferentes estudios analizados. En algunos casos, como DBT-P02510_MLOD, el aumento en la segmentación de BACs es notable (de 0,000 a 0,506), mientras que en otros, como DBT-P04396_MLOD, se observa una disminución en el desempeño. No obstante, como tendencia general, se aprecia una mejora consistente al trabajar con redes neuronales convolucionales bidimensionales en lugar de tridimensionales.

Sin embargo, pese a la mejora en el coeficiente de Dice promedio obtenido por el segmentador de calcificaciones arteriales, este resultado sigue considerándose insuficiente. Por ello, se analizan las posibles causas que podrían explicar este rendimiento limitado, llegando a lo observado en la Figura 4.1. En ella se muestra un ejemplo representativo de las predicciones realizadas por el segmentador de arterias calcificadas, donde se aprecia una clara sobresegmentación en comparación con la etiqueta original, fenómeno que se repite sistemáticamente en las predicciones. Esta deficiencia en la segmentación de arterias impacta directamente en el desempeño del segmentador de BACs, al utilizar como entrada una máscara de arterias poco precisa. Para confirmar esta hipótesis, se decide explorar modelos de arquitectura única, que reciben como entrada solamente los cortes oblicuos bidimensionales de mama, sin depender de la guía proporcionada por las predicciones de arterias calcificadas.



(a) Etiqueta arteria calcificada.

(b) Predicción arteria calcificada.

Figura 4.1: Caso arbitrario de una predicción de arterias calcificadas en el conjunto de validación.

4.2. Modelo de Arquitectura Única

Con el objetivo de identificar el mejor modelo basado en una arquitectura única, se realizaron diversas pruebas orientadas a determinar la función de pérdida más adecuada, así como sus respectivos hiperparámetros. En la Tabla 4.2 se presentan los coeficientes de Dice promedio obtenidos en cuatro experimentos realizados bajo distintas configuraciones, todos ellos implementados sobre una misma arquitectura U-Net, sus valores por volumen se pueden observar en el Anexo A.1.

Tabla 4.2: Coeficientes de Dice promedio de entrenamientos con la red U-Net variando su función de pérdida y sus hiperparámetros.

Función de Pérdida	Hiperparámetro α	Hiperparámetro γ	Coeficiente de Dice Promedio \pm Desviación Estándar
<i>Dice Loss</i>	-	-	0,194 \pm 0,158
<i>Dice Focal Loss</i>	0,8	2,0	0,195 \pm 0,164
	0,25	2,0	0,214 \pm 0,196
	0,25	1,0	0,213 \pm 0,198

Se observa que el peor desempeño corresponde al uso exclusivo de la función de pérdida *Dice Loss*, lo que refuerza la importancia de entrenar la red utilizando una función combinada que incluya tanto *Dice Loss* como *Focal Loss*. Respecto a los hiperparámetros evaluados, se inició probando un valor de $\alpha = 0,8$, buscando asignar un mayor peso a las calcificaciones arteriales y penalizar con más fuerza los falsos negativos. Sin embargo, los resultados muestran que disminuir α mejora el desempeño de la red, mientras que reducir γ no genera mejoras significativas, por lo que este

parámetro se mantiene en 2,0, valor asociado a un mayor enfoque en los casos difíciles. No se realizaron más pruebas, dado que cada entrenamiento requiere aproximadamente un día, por lo que se decidió limitar estos experimentos a cuatro, considerando el alcance de este trabajo y la necesidad de destinar tiempo a pruebas adicionales planificadas.

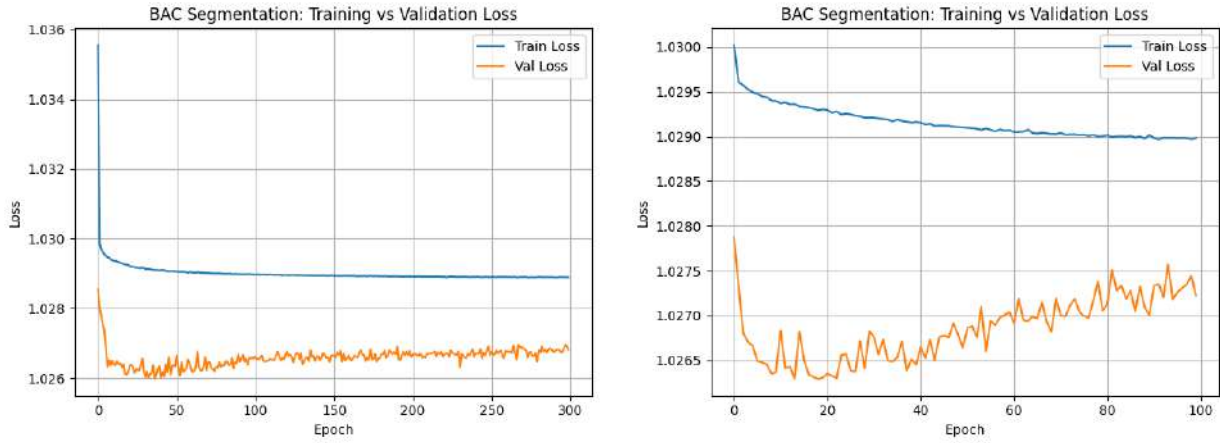
Una vez identificada la función de pérdida a utilizar, se realizaron experimentos adicionales variando la arquitectura de la red, comparando su rendimiento según el coeficiente de Dice promedio obtenido en tres arquitecturas distintas. La primera prueba se llevó a cabo utilizando la arquitectura U-Net, seguida de una nueva evaluación con la arquitectura Attention U-Net. Finalmente, se realizó una prueba adicional con la arquitectura Swin UNetR, ejecutada en el servidor NLHPC debido a su mayor demanda computacional. Los coeficientes de Dice promedio obtenidos en estas comparaciones se presentan en la Tabla 4.3, mientras que los coeficientes de Dice por volumen se encuentran en el Anexo A.2.

Tabla 4.3: Coeficientes de Dice promedio de entrenamientos variando la arquitectura usada y manteniendo una misma función de pérdida.

Arquitectura	Coeficiente de Dice Promedio \pm Desviación Estándar
U-Net	0,214 \pm 0,196
Attention U-Net	0,345 \pm 0,175
Swin UNetR	0,057 \pm 0,045

Se observa que el mejor resultado se obtiene utilizando la arquitectura Attention U-Net, superando tanto a la U-Net convencional como a la Swin UNetR en términos de coeficiente de Dice promedio. Esta mejora se atribuye a la incorporación de mecanismos de atención dentro de la arquitectura, los cuales permiten a la red enfocarse en las regiones más relevantes de la imagen durante el proceso de segmentación, mejorando así su capacidad para distinguir las calcificaciones arteriales del fondo. En contraste, la arquitectura U-Net convencional carece de este mecanismo explícito de enfoque, lo que limita su desempeño en escenarios de alta complejidad o con fuerte desequilibrio entre clases, como es este caso, donde las calcificaciones arteriales representan una proporción muy pequeña en comparación con el volumen total de la mama.

En cuanto a la Swin UNetR, se observa que presenta el peor desempeño entre las arquitecturas evaluadas hasta el momento. Esto podría deberse a que la mayor complejidad de la red la hace más propensa al sobreajuste, especialmente considerando que el conjunto de entrenamiento está compuesto únicamente por 49 volúmenes de mama. Esta situación queda evidenciada en la Figura 4.2, donde se aprecia que la curva de pérdida de validación de la arquitectura Swin UNetR aumenta rápidamente, alcanzando valores superiores en tan solo 100 épocas de entrenamiento en comparación con la curva de pérdida de validación de la arquitectura Attention U-Net, que muestra un comportamiento mucho más estable a lo largo de las 300 épocas.

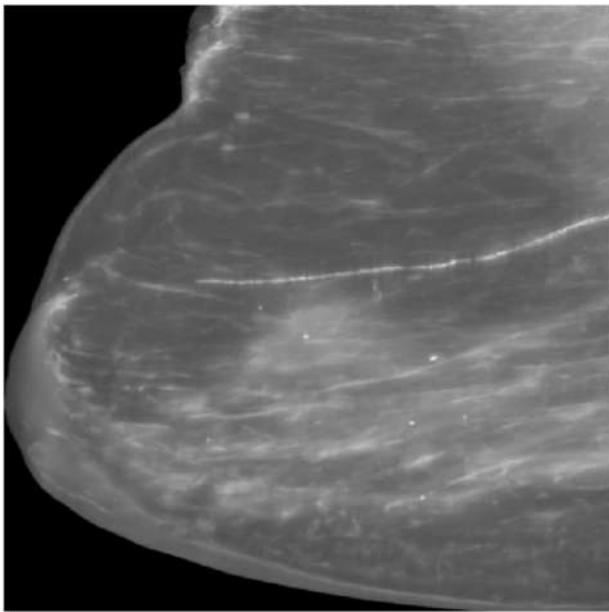


(a) Curvas de pérdida de la arquitectura Attention U-Net. (b) Curvas de pérdida de la arquitectura Swin UNetR.

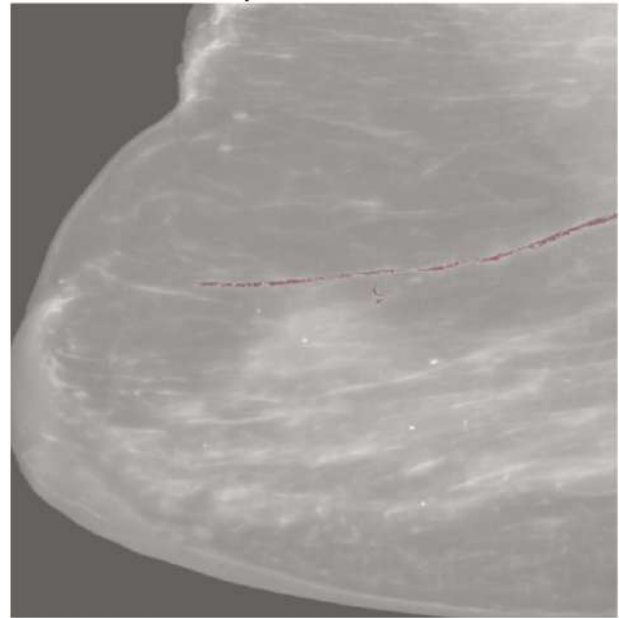
Figura 4.2: Comparación de las curvas de pérdida de las arquitecturas Attention U-Net y Swin UNetR ante una misma función de pérdida.

Si bien existe la posibilidad de mejorar su rendimiento ajustando algunos de sus hiperparámetros, como el *dropout*, que actúa como mecanismo de regularización para reducir el sobreajuste, llevar a cabo dichas optimizaciones implica desafíos prácticos. Dado que el entrenamiento se realiza en el servidor NLHPC, es necesario esperar una cola de ejecución, la cual en esta ocasión fue de aproximadamente cinco días, a lo que se suma el tiempo requerido para completar cada entrenamiento. Por lo tanto, encontrar la configuración óptima de hiperparámetros para esta arquitectura supondría un sacrificio considerable del tiempo disponible. Considerando el buen rendimiento alcanzado por la Attention U-Net y el desempeño deficiente del primer intento con Swin UNetR, se opta por priorizar el trabajo con la primera, en lugar de destinar recursos a la optimización de la arquitectura Swin UNetR.

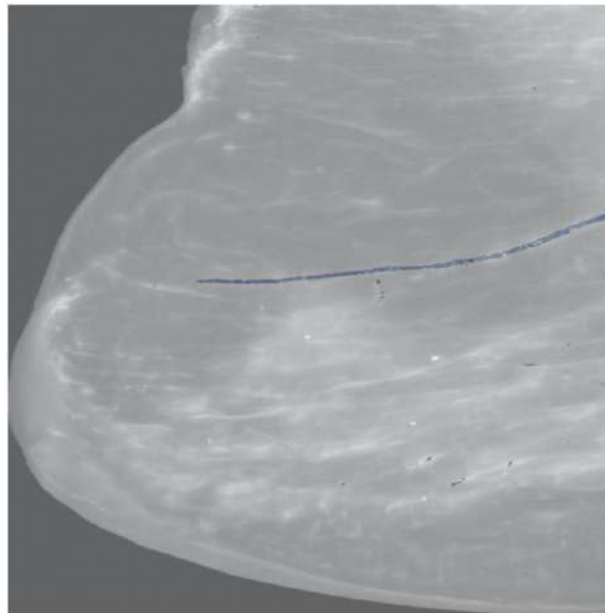
Al analizar las segmentaciones realizadas por el mejor modelo desarrollado hasta el momento, que utiliza la arquitectura Attention U-Net y la función de pérdida *Dice Focal Loss* con hiperparámetros $\alpha = 0,25$ y $\gamma = 2,0$, se observa que, a pesar de mejorar aproximadamente tres veces el coeficiente de Dice promedio en comparación con los resultados obtenidos en la tesis previa [26], dicho valor sigue siendo relativamente bajo y aún no resulta aceptable para un uso clínico. Por ello, se procede a analizar cualitativamente las segmentaciones generadas. En la Figura 4.3 se muestra un ejemplo representativo, en donde se aprecia una proyección de máxima identidad de la mama en la Figura 4.3.a, la etiqueta de la calcificación arterial en la mama en la Figura 4.3.b, y su respectiva predicción en la Figura 4.3.c, donde si bien se aprecia una segmentación generalmente adecuada de las BAC, también se evidencia una sobresegmentación, tanto dentro de la arteria calcificada como en regiones fuera de ella.



(a) MIP de la mama.



(b) Etiqueta de las BACs de la mama



(c) Predicción de las BACs de la mama

Figura 4.3: Ejemplo de segmentación realizada con arquitectura Attention U-Net.

4.3. Filtrado de Predicciones

Con el objetivo de filtrar las sobresegmentaciones fuera de las arterias calcificadas representadas en la Figura 4.3, se decide utilizar el segmentador de arterias desarrollado previamente para el modelo en cascada. Si bien esta red presentaba un problema importante de sobresegmentación, como se ejemplificó en la Figura 4.1, lo que lo hacía poco adecuado como guía de entrada para la segmentación de BACs, aún puede ser aprovechado como una máscara de postprocesamiento. Específicamente, se emplea para eliminar las predicciones de BACs que se encuentran fuera de las

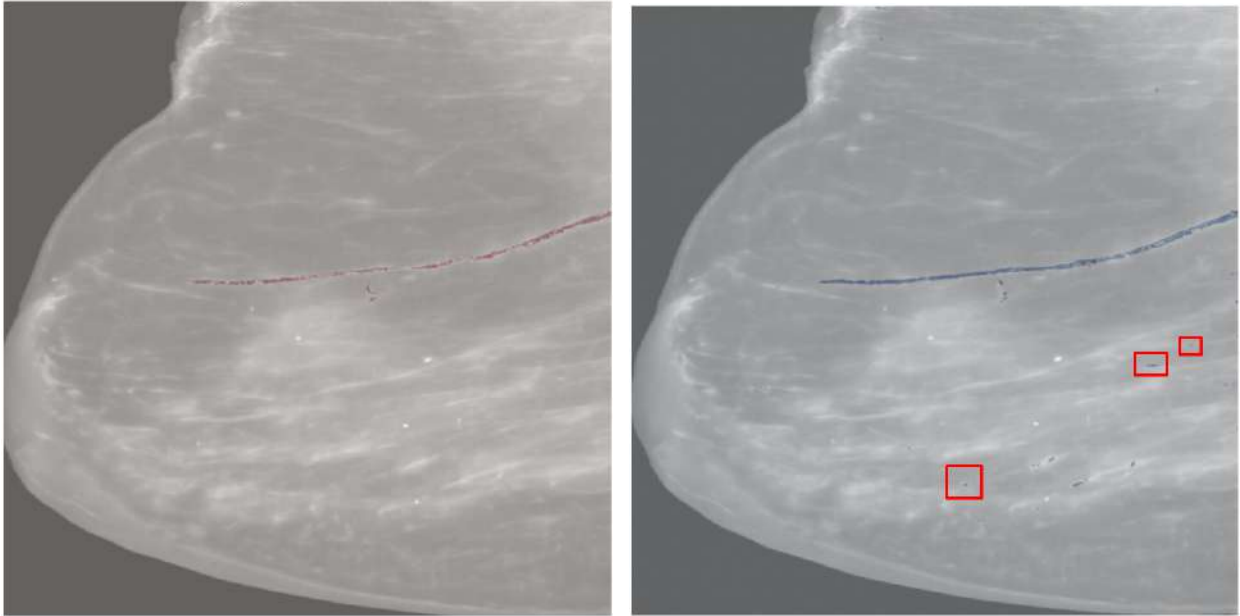
regiones arteriales, sin el riesgo de suprimir segmentaciones correctas de calcificaciones arteriales, ya que, como se mencionó previamente, su principal deficiencia radica en la sobresegmentación. Con esta estrategia se busca mejorar el rendimiento final del modelo.

Con tal de mejorar el desempeño del segmentador de BACs, se realizaron nuevos experimentos entrenando modelos que incorporan la nueva etiqueta de arterias no calcificadas, además de utilizar el segmentador de arterias calcificadas previamente desarrollado para el modelo en cascada. La Tabla 4.4 presenta los coeficientes de Dice promedio obtenidos para el segmentador de BACs tras aplicar el filtro en postprocesamiento generado por cada uno de estos tres enfoques, un segmentador de arterias calcificadas entrenado con las etiquetas de arterias calcificadas y no calcificadas, un segmentador de arterias general entrenado combinando ambas etiquetas, y el segmentador de arterias calcificadas utilizado previamente en el modelo en cascada. Los coeficientes de Dice obtenidos en cada volumen del conjunto de validación se observan en el Anexo A.3.

Tabla 4.4: Coeficientes de Dice promedio al aplicar la máscara obtenida con distintos segmentadores de arterias.

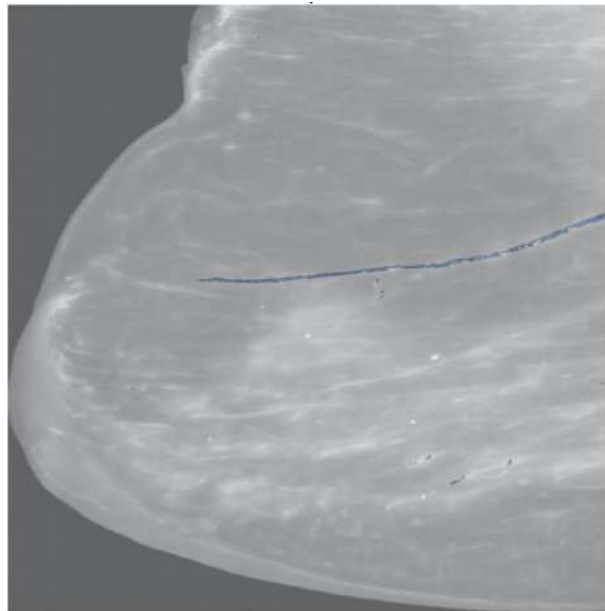
Experimento	Coeficiente de Dice Promedio de BACs post filtrado \pm Desviación Estándar
Segmentador de arterias calcificadas usando sólo su etiqueta	0,351 \pm 0,187
Segmentador de arterias general usando etiquetas de arterias calcificadas y no calcificadas	0,315 \pm 0,196
Segmentador de arterias calcificadas usando ambas etiquetas de arterias	0,337 \pm 0,197

Se observa que el mejor coeficiente de Dice se obtiene al utilizar como máscara el segmentador de arterias calcificadas creado originalmente, entrenado únicamente con las etiquetas de arterias calcificadas, una representación de su mejora de desempeño se puede observar en la Figura 4.4. Se plantea como hipótesis que este resultado se debe a que la etiqueta de arterias no calcificadas está disponible solo para 16 volúmenes del conjunto de entrenamiento y 6 del conjunto de validación, lo que reduce considerablemente la cantidad de datos utilizada para entrenar los modelos que incorporan esta información adicional. Esta limitación podría haber afectado negativamente la capacidad de generalización de las predicciones frente a casos que no disponen de la nueva etiqueta, a pesar de contar con mayor información en los casos que sí la poseen.



(a) Etiqueta de BACs.

(b) Predicción sin filtrado, coeficiente de Dice = 0,689.



(c) Predicción filtrada, coeficiente de Dice = 0,722.

Figura 4.4: Representación del filtrado de detecciones fuera de arterias calcificadas. En (b), los rectángulos rojos destacan algunas predicciones positivas que no corresponden a calcificaciones asociadas a vasos, ilustrando las sobresegmentaciones que son eliminadas tras aplicar el filtrado mostrado en (c).

Una vez identificado el mejor modelo para la segmentación de BACs, se procedió a tabular sus resultados por volumen en la Tabla 4.5, incorporando además las métricas de precisión y sensibilidad, con el objetivo de realizar un análisis más detallado del desempeño obtenido.

Tabla 4.5: Comparación de métricas del modelo segmentador de BACs (Attention U-Net con filtrado por arterias) frente a resultados de la tesis previa.

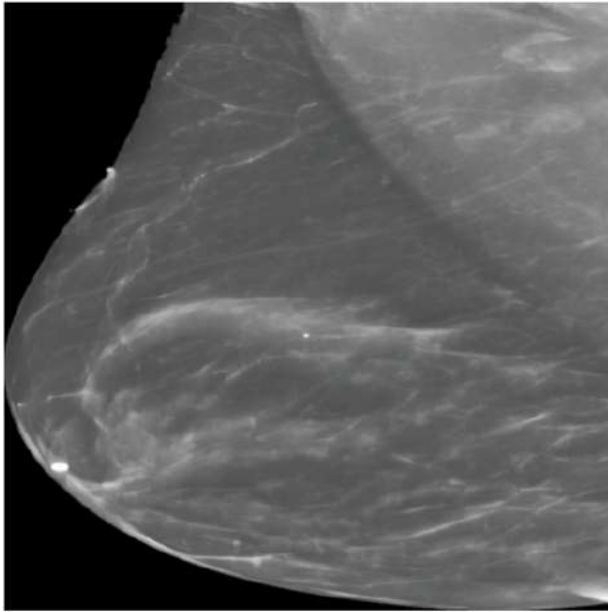
Volumen	Dice BAC Tesis	Dice BAC Modelo Attention U-Net Filtrado con Segmentación de Arterias	Precisión	Sensibilidad
DBT-P00001_MLOI	0,316	0,575	0,456	0,778
DBT-P00285_MLOI	0,112	0,261	0,752	0,158
DBT-P00303_MLOD	0,060	0,126	0,394	0,075
DBT-P00491_MLOD	0,078	0,200	0,871	0,113
DBT-P00523_MLOD	0,000	0,000	0,000	0,000
DBT-P00688_MLOD	0,013	0,391	0,447	0,348
DBT-P00887_MLOD	0,355	0,525	0,399	0,764
DBT-P01488_MLOD	0,000	0,262	0,188	0,435
DBT-P01762_MLOI	0,000	0,481	0,366	0,700
DBT-P01778_MLOI	0,000	0,269	0,163	0,771
DBT-P01957_MLOD	0,000	0,336	0,274	0,433
DBT-P02244_MLOD	0,077	0,451	0,304	0,870
DBT-P02510_MLOD	0,000	0,722	0,595	0,918
DBT-P04368_MLOD	0,000	0,221	0,738	0,130
DBT-P04396_MLOD	0,146	0,436	0,411	0,464
Promedio \pm Desviación Estándar	0,080 \pm 0,115	0,351 \pm 0,187	0,424 \pm 0,236	0,464 \pm 0,318

Al analizar cuantitativamente los resultados presentados en la Tabla 4.5, se observa una mejora significativa en el desempeño del modelo segmentador de BACs propuesto, alcanzando un coeficiente de Dice promedio de 0,351, en comparación con el 0,080 obtenido en la tesis previa, lo que representa más de cuatro veces de incremento. Además, se reportan valores promedio de precisión y sensibilidad de 0,424 y 0,464 respectivamente, indicando un balance intermedio entre la capacidad del modelo para evitar falsos positivos y su habilidad para detectar la mayor cantidad posible de calcificaciones arteriales. No obstante, se evidencia una considerable variabilidad entre volúmenes, por lo que se decide revisar cualitativamente algunos casos representativos del desempeño del algoritmo, mientras que las predicciones obtenidas en todos los volúmenes del conjunto de validación se pueden observar en el Anexo B.

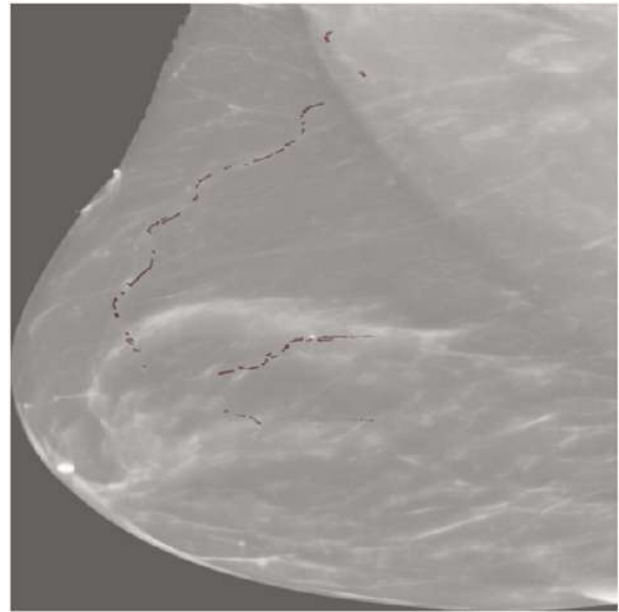
- Volúmenes con coeficiente de Dice, precisión y sensibilidad altos: como es el caso del volumen DBT-P02510_MLOD, cuya predicción se muestra en la Figura 4.4. En este caso, la segmentación generada se asemeja visualmente en gran medida a la etiqueta, presentando únicamente pequeñas predicciones fuera de la arteria calcificada. Sin embargo, a pesar de ser el volumen con el coeficiente de Dice más alto, este valor sigue sin alcanzar métricas consideradas buenas (por ejemplo, superiores a 0,85 o 0,9). Una posible razón de esta limitación es lo señalado previamente al introducir la Figura 3.4, donde se evidencia que parte de la etiqueta

puede perderse durante el redimensionamiento de la imagen. Aunque actualmente el redimensionamiento aplicado no es tan drástico como en ese ejemplo, es posible que aún afecte la integridad de las etiquetas a menor escala, lo que podría explicar que algunas predicciones correctas no sean reconocidas como verdaderos positivos al momento de calcular las métricas y al mismo tiempo se aprecie un buen resultado visualmente.

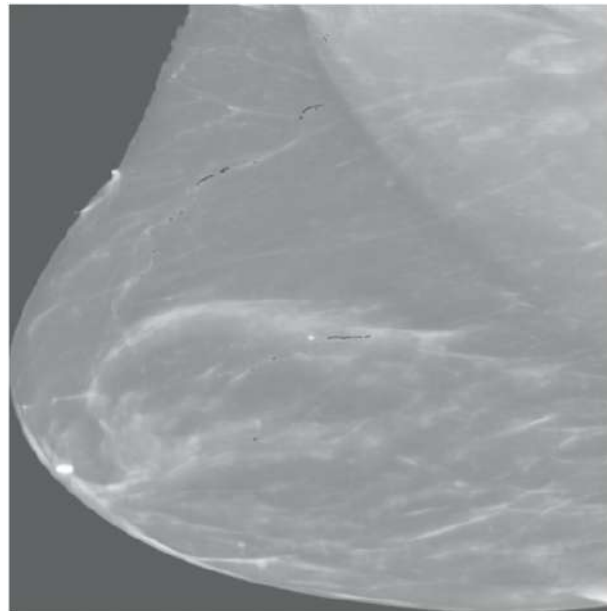
- Volúmenes con valores altos de precisión pero bajos de sensibilidad: como es el caso del volumen DBT-P00285_MLOI, tienden a repetirse en pacientes con mayor densidad mamaria, es decir, mamas que presentan una proporción más elevada de tejido fibroglandular (denso) en comparación con tejido adiposo (graso) [34]. Esta característica dificulta que el algoritmo segmente correctamente en regiones donde las arterias están parcialmente cubiertas por dicho tejido denso, limitándose a segmentar únicamente en zonas donde las arterias son visualmente más claras. Esto se observa en la Figura 4.5, donde la precisión es alta debido a que el modelo identifica correctamente las regiones claramente visibles, pero la sensibilidad es baja porque omite segmentaciones en las áreas donde las arterias se van ocultando por el tejido fibroglandular, a pesar de que efectivamente contienen calcificaciones.



(a) MIP de la mama.



(b) Etiqueta de BACs.

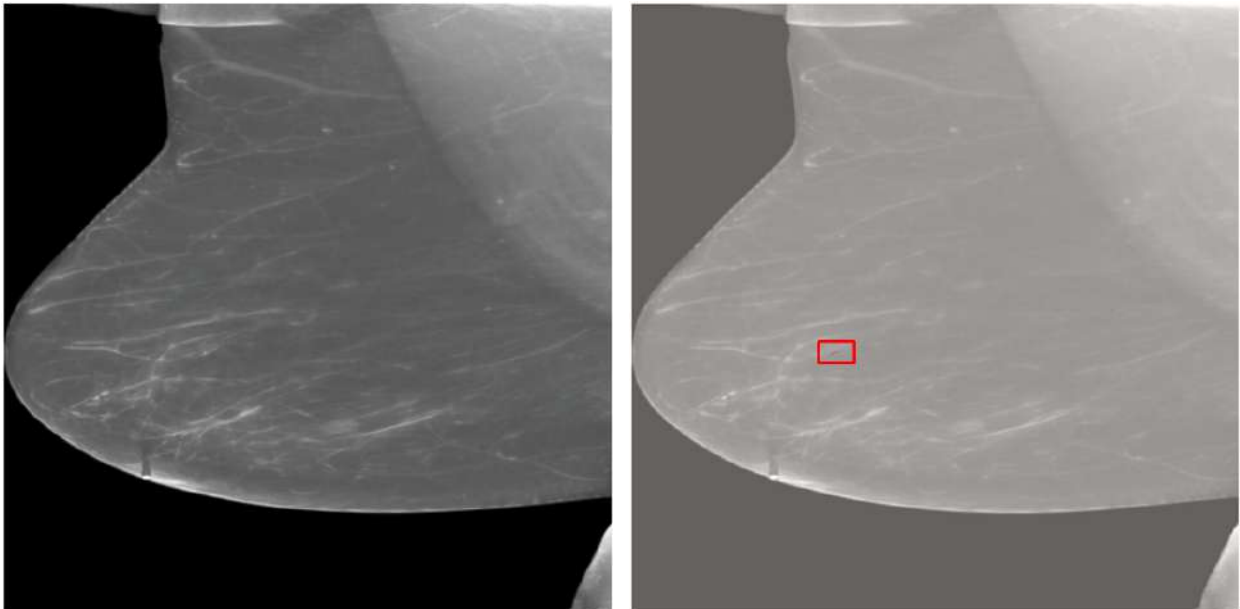


(c) Predicción filtrada.

Figura 4.5: Predicción representativa de casos con alta precisión pero baja sensibilidad. En (c) se observa que las regiones más visibles de las arterias son correctamente segmentadas, mientras que las zonas parcialmente ocultas por tejido denso no son detectadas, explicando el comportamiento de alta precisión pero baja sensibilidad.

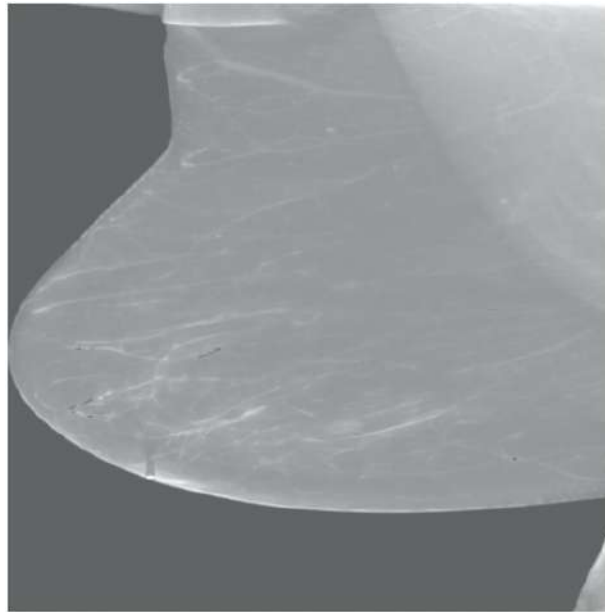
- Volúmenes con precisión baja y sensibilidad alta: como es el caso del volumen DBT-P01778_MLOI, suelen corresponder a mamas con una cantidad muy reducida de calcificaciones arteriales. Como se observa en la Figura 4.6, la etiqueta de BAC en este volumen es mínima en comparación con el tamaño total de la mama. Esta situación provoca que, aunque el modelo logre detectar correctamente la calcificación arterial presente, cualquier sobresegmentación, por pequeña

que sea, penalice fuertemente el valor del coeficiente de Dice y la precisión. En consecuencia, incluso cuando cualitativamente se aprecia que la calcificación ha sido identificada correctamente sin sobresegmentar en gran medida, el modelo presenta métricas cuantitativas bajas debido a estas predicciones adicionales.



(a) MIP de la mama.

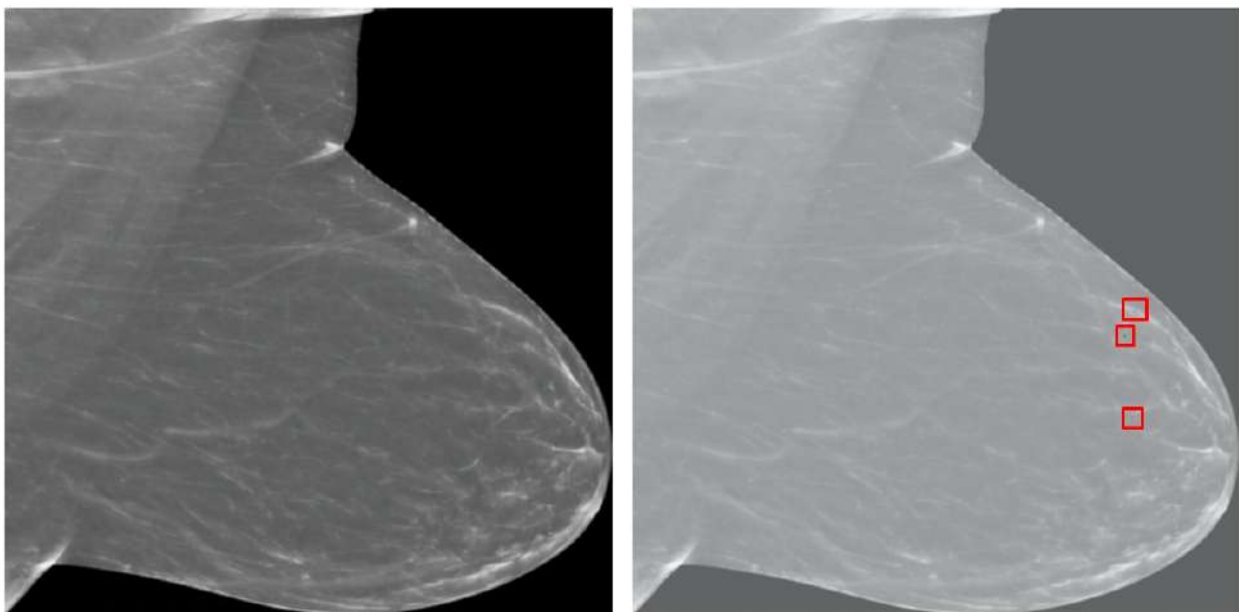
(b) Etiqueta de BACs.



(c) Predicción filtrada.

Figura 4.6: Predicción representativa de casos con baja precisión pero sensibilidad alta. En (b) se observa que la calcificación es extremadamente reducida en comparación con el tamaño total de la mama, el rectángulo rojo indica la localización aproximada de dicha calcificación para facilitar su visualización. (c) Indica que el modelo logra detectar correctamente la calcificación, pero cualquier sobresegmentación mínima impacta fuertemente las métricas cuantitativas, resultando en baja precisión.

- Volumen con máscara nula: se observa que el volumen DBT-P00523_MLOD presenta el valor mínimo en todas las métricas evaluadas. Esto se debe a que no contiene calcificaciones arteriales, por lo que al aplicar la Ecuación 2.3, el resultado solo puede tomar dos valores extremos: cero si se segmenta alguna región (ya que no hay solapamiento real), y uno si no se segmenta nada (coincidiendo perfectamente con la ausencia de calcificación). Lo mismo ocurre para las métricas precisión y sensibilidad. Como se aprecia en la Figura 4.7, pequeñas predicciones erróneas reducen al mínimo el rendimiento cuantitativo del modelo, aunque cualitativamente se puede considerar que está realizando un buen trabajo al casi no detectar calcificaciones. Estas pequeñas segmentaciones que el modelo aún tiende a predecir corresponden a regiones de arterias cubiertas por tejido fibroglandular, fenómeno también perceptible en las demás predicciones realizadas, apreciable en la Figura 4.6 por ejemplo, donde al existir arterias reales en esas zonas, la máscara aplicada no logra filtrar completamente dichas predicciones.



(a) MIP de la mama.

(b) Etiqueta de BACs.

Figura 4.7: Predicción en volumen con máscara nula. En (b) se observa que la predicción es extremadamente reducida en comparación con el tamaño total de la mama, los rectángulos rojos indican la localización aproximada de dicha predicción para facilitar su visualización.

4.4. Cuantificador de Calcificaciones

Finalmente, se aplica el índice de cuantificación diseñado sobre las predicciones de calcificaciones arteriales, utilizando las segmentaciones predichas de arterias calcificadas para estimar el área total de referencia. Este índice, descrito en la Ecuación 3.2 del Capítulo 3, corresponde a la razón entre el área de intersección entre BACs y arterias calcificadas, y el área total de arterias calcificadas. Sin embargo, como se observa en la Figura 4.1 y en la Tabla 4.1, este segmentador no logró buenos resultados, por lo que se incluye adicionalmente el cálculo del índice empleando

la etiqueta de arterias calcificadas, con el objetivo de estimar el valor que podría obtenerse en un escenario ideal con un segmentador de arterias calcificadas preciso. Además, se comparan estos resultados con los índices obtenidos usando la etiqueta real de BACs, lo cual se presenta resumido en la Tabla 4.6.

Tabla 4.6: Valor del índice de calcificación diseñado en las predicciones del mejor modelo filtrado creado.

Volumen	Etiqueta de BAC dentro de arteria calcificada predicha	Predicción de BAC dentro de arteria calcificada predicha	Etiqueta de BAC dentro de etiqueta de arteria calcificada	Predicción de BAC dentro de etiqueta de arteria calcificada
DBT-P00001_MLOI	0,044	0,079	0,174	0,293
DBT-P00285_MLOI	0,023	0,007	0,104	0,024
DBT-P00303_MLOD	0,018	0,004	0,109	0,013
DBT-P00491_MLOD	0,040	0,006	0,253	0,035
DBT-P00523_MLOD	0,000	0,001	0,000	0,000
DBT-P00688_MLOD	0,003	0,003	0,029	0,016
DBT-P00887_MLOD	0,009	0,019	0,131	0,198
DBT-P01488_MLOD	0,003	0,007	0,069	0,072
DBT-P01762_MLOI	0,025	0,057	0,132	0,273
DBT-P01778_MLOI	0,001	0,003	0,009	0,016
DBT-P01957_MLOD	0,002	0,003	0,005	0,006
DBT-P02244_MLOD	0,005	0,016	0,062	0,167
DBT-P02510_MLOD	0,033	0,051	0,490	0,736
DBT-P04368_MLOD	0,041	0,018	0,209	0,060
DBT-P04396_MLOD	0,008	0,011	0,061	0,067
Promedio ± Desviación Estándar	0,017 ± 0,016	0,019 ± 0,024	0,132 ± 0,126	0,132 ± 0,194

De la Tabla 4.6 se observa que, al calcular el índice utilizando las segmentaciones predichas de arterias calcificadas, se obtienen valores consistentemente inferiores al 8%. Esto se debe a que dichas predicciones presentan una marcada sobresegmentación, lo que incrementa artificialmente el área total considerada como arteria calcificada. Como resultado, al dividir el área de BACs detectadas por esta área sobresegmentada, el índice final resulta muy bajo, reflejando que sólo una mínima proporción de la arteria identificada presenta calcificaciones.

Con el fin de realizar un análisis más profundo de la segmentación de BACs, se calcula el índice considerando como referencia el área total de las etiquetas de arterias calcificadas. De forma general, se observa que las predicciones tienden a indicar que el área ocupada por BACs es mayor que la indicada por las propias etiquetas. Este fenómeno puede explicarse, en parte, por lo descrito

en la Figura 3.4, donde se muestra que al redimensionar las imágenes, se pierde parte de la etiqueta real, mientras que las predicciones no sufren este recorte al generarse directamente sobre la imagen redimensionada. Además, se considera preferible que el algoritmo tienda a sobresegmentar las calcificaciones arteriales dentro de las arterias, pues es más deseable alertar sobre un posible riesgo (incluso en presencia de algunos falsos positivos) que incurrir en falsos negativos y no identificar casos que realmente representen un riesgo cardiovascular.

Sin embargo, se identifican cuatro volúmenes en los cuales el índice calculado a partir de las predicciones de BACs es considerablemente menor en comparación con el índice obtenido usando las etiquetas, mostrando que menos del 10% del área de la arteria calcificada se encuentra cubierta por calcificaciones según el modelo, cuando en realidad este no es el caso. Este comportamiento refleja una limitación del algoritmo y corresponde principalmente a volúmenes representados por el escenario ilustrado en la Figura 4.5, es decir, mamas densas en las que la alta proporción de tejido fibroglandular dificulta la visualización de las arterias, afectando negativamente la capacidad del modelo para segmentar correctamente las calcificaciones arteriales en dichas regiones.

Capítulo 5

Discusión

En este capítulo se discuten los resultados obtenidos en las pruebas de segmentación y cuantificación de calcificaciones arteriales en tomosíntesis mamarias, interpretando su relevancia, comparándolos con estudios previos y abordando limitaciones.

En primer lugar, se considera la tesis previa a este trabajo como el principal punto de comparación, dado que utiliza la misma base de datos para su desarrollo y emplea el coeficiente de Dice como métrica de evaluación. Los resultados obtenidos muestran que el algoritmo desarrollado en este trabajo aumenta consistentemente su rendimiento en cada uno de los volúmenes mamarios del conjunto de validación, alcanzando un coeficiente de Dice promedio más de cuatro veces superior al reportado en la tesis anterior.

Las principales diferencias entre ambos trabajos, que explican en gran parte la mejora en el rendimiento, se pueden resumir en los siguientes puntos:

- Se construye una nueva base de datos bidimensional a partir de la base tridimensional original, separando cada corte oblicuo de mama como una imagen bidimensional independiente. El entrenamiento y las predicciones se realizan sobre cada corte por separado, lo que además asemeja el enfoque seguido durante el etiquetado manual original de las calcificaciones arteriales en los volúmenes mamarios. Los cortes se reúnen únicamente al momento de la evaluación global.
- Se incorpora la función de pérdida *Focal Loss* en combinación con *Dice Loss*, con el objetivo de abordar el desbalance existente entre las calcificaciones arteriales y el fondo.
- Se emplea la arquitectura Attention U-Net, la cual, gracias a sus mecanismos de atención, permite que la red se enfoque en las regiones de mayor relevancia de la imagen, abordando nuevamente la dificultad planteada por el desbalance entre clases.
- Se emplea un mecanismo para reducir las segmentaciones de calcificaciones arteriales fuera de las arterias calcificadas, aplicando una máscara predicha de arterias calcificadas que permite filtrar aquellas BACs mal segmentadas, eliminando predicciones fuera del contexto anatómico correcto, mejorando así su rendimiento.

Sin embargo, el coeficiente de Dice promedio alcanzado sigue siendo bajo, con un 35,1 %, y un valor máximo de 72,2 %, lo cual aún está lejos de ser considerado aceptable para un uso médico, donde idealmente se esperan valores superiores al 90 %.

Este desempeño limitado del algoritmo puede atribuirse a diversos factores presentes durante el desarrollo de este trabajo. Entre ellos destacan las limitaciones computacionales, ya que trabajar con una GPU de 8 GB de memoria impidió implementar algoritmos de mayor complejidad, como añadir más capas a las redes neuronales convolucionales diseñadas o utilizar resoluciones más cercanas a las dimensiones originales de los cortes bidimensionales. Otra limitación relevante fue la cantidad de imágenes disponibles para entrenar el modelo, únicamente 49 volúmenes mamarios, lo cual resulta insuficiente considerando tanto el desbalance de clases presente como la complejidad general del problema abordado.

Por otra parte, los resultados obtenidos en volúmenes mamarios representados en la Figura 4.6 permiten identificar ciertas limitaciones en la métrica de evaluación utilizada. Cualitativamente, se observa que el modelo realiza segmentaciones adecuadas, logrando detectar correctamente las calcificaciones presentes y generando solo pequeñas segmentaciones adicionales fuera de la arteria calcificada. Sin embargo, debido al tamaño extremadamente reducido de la etiqueta de calcificación en estos casos, cualquier predicción fuera de ella penaliza fuertemente el resultado cuantitativo, obteniéndose coeficientes de Dice tan bajos como 26,2 %, a pesar de evidenciarse visualmente una segmentación razonable.

Este problema se evidencia aún más en la Figura 4.7, donde se presenta un volumen con máscara nula como etiqueta de calcificaciones arteriales. En este caso, el modelo realiza predicciones ínfimas, lo que se traduce en un valor mínimo del coeficiente de Dice. Sin embargo, cualitativamente se observa que el algoritmo identifica correctamente que prácticamente ninguna región del volumen corresponde a una BAC. No obstante, desde un punto de vista cuantitativo, la métrica utilizada penaliza fuertemente cualquier mínima predicción en ausencia de etiqueta, reflejando erróneamente un desempeño completamente fallido.

Otra limitación del modelo se evidencia en volúmenes mamarios representados por la Figura 4.5, correspondientes a mamas densas, donde el algoritmo no logra segmentar correctamente en las zonas con mayor proporción de tejido fibroglandular. La alta densidad dificulta la visualización de las arterias, lo que repercute tanto en el desempeño cuantitativo como cualitativo del modelo, reflejando un rendimiento deficiente en estos casos específicos.

Por último, se destaca que, para obtener una estimación más fiel del indicador de calcificación diseñado, fue necesario utilizar las etiquetas manuales de arterias calcificadas en lugar de las predicciones generadas por los modelos entrenados, dado que no se logró desarrollar un algoritmo capaz de segmentar estas estructuras sin que realicen una alta sobresegmentación. Cabe destacar que el enfoque principal de este trabajo estuvo centrado en la segmentación y cuantificación de

calcificaciones arteriales, por lo que no se destinaron recursos adicionales profundos a optimizar un modelo específico para la segmentación precisa de arterias calcificadas.

Actualmente, el índice de cuantificación propuesto depende de las etiquetas de arterias calcificadas, lo que limita su aplicabilidad clínica directa, ya que en escenarios reales no se dispone de estas segmentaciones de referencia. No obstante, si en el futuro se logran superar las limitaciones actuales del modelo de segmentación diseñado, el indicador podría constituir una herramienta útil para cuantificar de manera objetiva la proporción de arteria afectada por calcificaciones, aportando una métrica cuantitativa que podría correlacionarse con el riesgo cardiovascular asociado a las BAC. Será necesario realizar estudios adicionales para validar si la información entregada por este índice es clínicamente relevante y predictiva, o si debería combinarse con otros índices propuestos para una evaluación integral del riesgo cardiovascular en pacientes.

Capítulo 6

Conclusión

En este capítulo se presentan las principales conclusiones del trabajo realizado, destacando los logros alcanzados en relación con los objetivos específicos planteados, así como las proyecciones a futuro derivadas de las limitaciones identificadas en este estudio.

En primer lugar, se destaca el desarrollo del algoritmo de deep learning propuesto, cuyo objetivo fue segmentar automáticamente las calcificaciones arteriales presentes en las imágenes de tomosíntesis mamaria. Se logró desarrollar un modelo de arquitectura Attention U-Net capaz de segmentar cualitativamente de manera adecuada las calcificaciones en mamas no densas, aunque los resultados cuantitativos no alcanzaron valores óptimos, un coeficiente de Dice del 35,1 %, en parte debido a las limitaciones asociadas a la métrica de evaluación utilizada. Sin embargo, el algoritmo presenta dificultades al generalizarse en casos de mamas densas, lo que abre espacio para mejoras futuras. Adicionalmente, se identificó que, debido a los recursos computacionales disponibles, al tamaño limitado de la base de datos y a la complejidad de su arquitectura, el uso de modelos de deep learning basados sólo en CNNs produjo un mejor rendimiento que aquellos que incorporaban Vision Transformers, resultando más adecuado prescindir de estos últimos en el contexto en que se desarrolló este trabajo.

Por otra parte, se logró validar adecuadamente los algoritmos desarrollados, utilizando los resultados obtenidos de coeficiente de Dice para tomar decisiones metodológicas fundamentadas. El análisis de los errores observados durante la validación permitió ajustar progresivamente los modelos, identificando y seleccionado así el modelo con el mejor rendimiento dentro de los evaluados.

Por último, se logró diseñar de manera preliminar un indicador que cuantifica las calcificaciones arteriales, identificando una fórmula capaz de reflejar un posible riesgo asociado. El algoritmo desarrollado mostró resultados promedio similares a los obtenidos a partir del etiquetado manual. Sin embargo, fue necesario utilizar las etiquetas de arterias calcificadas para calcular correctamente este indicador, ya que no se logró implementar un modelo capaz de segmentar dichas arterias con la precisión requerida. Por este motivo, el indicador no podría ser aplicado en un escenario real sin la disponibilidad previa de estas etiquetas, quedando como una herramienta en etapa exploratoria.

6.1. Proyección a Futuro

A partir del análisis de los resultados obtenidos y de las discusiones realizadas, se identifican diversas oportunidades de mejora que podrían implementarse para desarrollar un algoritmo con un rendimiento superior.

En caso de contar con mayores recursos computacionales, sería posible realizar pruebas con redes neuronales convolucionales de mayor complejidad que las desarrolladas en este trabajo, incorporando arquitecturas más profundas o dedicando el tiempo necesario a optimizar modelos que incluyan Vision Transformers.

Los modelos que incorporan Vision Transformers, debido a su mayor complejidad, también podrían beneficiarse de la incorporación de nuevas bases de datos que permitan aumentar el número de volúmenes mamarios con calcificaciones arteriales disponibles para entrenamiento. Considerando el número limitado de muestras utilizado en este trabajo, disponer de un conjunto de datos más amplio posiblemente contribuiría a mejorar su rendimiento.

Existe una oportunidad de mejora en el uso de nuevas métricas de evaluación que reflejen de manera más fiel los resultados cualitativos obtenidos por el algoritmo, como por ejemplo considerar el coeficiente de Dice de longitud [31]. Esta métrica podría emplearse inicialmente como herramienta de validación para evaluar la continuidad de las calcificaciones segmentadas, complementando al coeficiente de Dice tradicional centrado en área. También se podría explorar su incorporación en el entrenamiento, combinándola con otras funciones de pérdida para guiar al modelo no solo a segmentar áreas correctas, sino también a mantener la coherencia longitudinal de las BAC. A su vez, se podría explorar esquemas de evaluación en los que los volúmenes con muy pocas calcificaciones arteriales tengan un menor peso en la métrica final, permitiendo así evitar que casos marginales penalicen desproporcionadamente el desempeño global reportado.

Por otra parte, para mejorar el rendimiento en mamas densas, podría incorporarse un algoritmo previo que detecte la densidad mamaria, permitiendo adaptar el segmentador a estas situaciones específicas, como por ejemplo, mediante el ajuste de hiperparámetros, la implementación de estrategias de entrenamiento diferenciadas o el desarrollo de un modelo especializado para estos casos, mejorando así la generalización del algoritmo.

Por último, desarrollar un algoritmo robusto para la segmentación de arterias calcificadas permitiría una evaluación más precisa del indicador diseñado en este trabajo, además de posibilitar su aplicación en casos donde no se cuente previamente con las etiquetas de referencia. A su vez, presenta una oportunidad de diseño de nuevos indicadores que permitan estimar de forma más adecuada el riesgo cardiovascular, como por ejemplo, cuantificar el grado de obstrucción transversal de la arteria o determinar qué proporción de su longitud presenta calcificaciones, caracterizando de forma más completa el potencial riesgo que se presenta.

Bibliografía

- [1] Hendriks, E. J., De Jong, P. A., Van Der Graaf, Y., Mali, W. P., Van Der Schouw, Y. T., y Beulens, J. W., “Breast arterial calcifications: A systematic review and meta-analysis of their determinants and their association with cardiovascular events”, *Atherosclerosis*, vol. 239, no. 1, pp. 11–20, 2015, [doi:10.1016/j.atherosclerosis.2014.12.035](https://doi.org/10.1016/j.atherosclerosis.2014.12.035).
- [2] Elizalde Pérez, A., “Tomosíntesis mamaria: bases físicas, indicaciones y resultados”, *Revista de Senología y Patología Mamaria*, vol. 28, no. 1, pp. 39–45, 2015, [doi:10.1016/j.senol.2014.10.004](https://doi.org/10.1016/j.senol.2014.10.004).
- [3] Valdivielso, J. M., “Vascular calcification: types and mechanisms”,.
- [4] Trimboli, R. M., Codari, M., Guazzi, M., y Sardanelli, F., “Screening mammography beyond breast cancer: breast arterial calcifications as a sex-specific biomarker of cardiovascular risk”, *European Journal of Radiology*, vol. 119, p. 108636, 2019, [doi:10.1016/j.ejrad.2019.08.005](https://doi.org/10.1016/j.ejrad.2019.08.005).
- [5] Yoon, Y. E., Kim, K. M., Lee, W., Han, J. S., Chun, E. J., Ahn, S., Choi, S. I., Yun, B. L., y Suh, J.-W., “Breast Arterial Calcification is Associated with the Progression of Coronary Atherosclerosis in Asymptomatic Women: A Preliminary Retrospective Cohort Study”, *Scientific Reports*, vol. 10, no. 1, p. 2755, 2020, [doi:10.1038/s41598-020-59606-y](https://doi.org/10.1038/s41598-020-59606-y).
- [6] “What Is Machine Learning (ML)? | IBM”, 2021, <https://www.ibm.com/topics/machine-learning> (visitado el 2024-12-02).
- [7] Domingos, P., “A few useful things to know about machine learning”, *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012, [doi:10.1145/2347736.2347755](https://doi.org/10.1145/2347736.2347755).
- [8] “What is a Neural Network? | IBM”, 2021, <https://www.ibm.com/topics/neural-networks> (visitado el 2024-12-02).
- [9] “What Is Deep Learning? | IBM”, 2024, <https://www.ibm.com/topics/deep-learning> (visitado el 2024-12-02).
- [10] “What are Convolutional Neural Networks? | IBM”, 2021, <https://www.ibm.com/topics/convolutional-neural-networks> (visitado el 2024-12-02).
- [11] Cutillas, C. A., *Detección de tumores en imágenes médicas mediante transformers*. Tesis PhD, Universidad de Alicante, 2023.
- [12] LeCun, Y., Bengio, Y., y Hinton, G., “Deep learning”, *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, [doi:10.1038/nature14539](https://doi.org/10.1038/nature14539).

- [13] Goodfellow, I., Bengio, Y., y Courville, A., Deep Learning. MIT Press. <http://www.deeplearningbook.org>.
- [14] Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., Chavez-Urbiola, E. A., y Romero-Gonzalez, J. A., “Loss Functions and Metrics in Deep Learning”, 2024, [doi:10.48550/arXiv.2307.02694](https://doi.org/10.48550/arXiv.2307.02694). arXiv:2307.02694 [cs].
- [15] Taha, A. A. y Hanbury, A., “Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool”, BMC Medical Imaging, vol. 15, no. 1, p. 29, 2015, [doi:10.1186/s12880-015-0068-x](https://doi.org/10.1186/s12880-015-0068-x).
- [16] Lin, T.-Y., Goyal, P., Girshick, R., He, K., y Dollár, P., “Focal Loss for Dense Object Detection”, 2018, [doi:10.48550/arXiv.1708.02002](https://doi.org/10.48550/arXiv.1708.02002). arXiv:1708.02002 [cs].
- [17] Ronneberger, O., Fischer, P., y Brox, T., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, 2015, [doi:10.48550/arXiv.1505.04597](https://doi.org/10.48550/arXiv.1505.04597). arXiv:1505.04597 [cs].
- [18] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., y Rueckert, D., “Attention U-Net: Learning Where to Look for the Pancreas”, 2018, [doi:10.48550/arXiv.1804.03999](https://doi.org/10.48550/arXiv.1804.03999). arXiv:1804.03999 [cs].
- [19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., y Polosukhin, I., “Attention Is All You Need”, 2023, [doi:10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762). arXiv:1706.03762 [cs].
- [20] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., y Hounsby, N., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”, 2021, [doi:10.48550/arXiv.2010.11929](https://doi.org/10.48550/arXiv.2010.11929). arXiv:2010.11929 [cs].
- [21] Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., y Shah, M., “Transformers in Vision: A Survey”, ACM Computing Surveys, vol. 54, no. 10s, pp. 1–41, 2022, [doi:10.1145/3505244](https://doi.org/10.1145/3505244). arXiv:2101.01169 [cs].
- [22] “Multilayer Perceptrons in Machine Learning: A Comprehensive Guide”., <https://www.datacamp.com/tutorial/multilayer-perceptrons-in-machine-learning> (visitado el 2024-12-02).
- [23] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., y Guo, B., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows”, 2021, [doi:10.48550/arXiv.2103.14030](https://doi.org/10.48550/arXiv.2103.14030). arXiv:2103.14030 [cs].
- [24] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., y Wang, M., “Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation”, 2021, [doi:10.48550/arXiv.2105.05537](https://doi.org/10.48550/arXiv.2105.05537). arXiv:2105.05537 [eess].
- [25] Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., y Xu, D., “Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images”, 2022, [doi:10.48550/arXiv.2201.01266](https://doi.org/10.48550/arXiv.2201.01266). arXiv:2201.01266 [eess].

- [26] Flores, M. y Ramírez, C., “Método de cuantificación de calcificaciones vasculares en tomosíntesis mamaria con inteligencia artificial”, tesis de grado, Departamento de Tecnología médica, UCh, Santiago, 2024.
- [27] Buda, M., Saha, A., Walsh, R., Ghate, S., Li, N., Swiecicki, A., Lo, J. Y., Yang, J., y Mazurowski, M., “Breast cancer screening – digital breast tomosynthesis (bcs-dbt) (version 5)”, 2020, [doi:10.7937/E4WT-CD02](https://doi.org/10.7937/E4WT-CD02).
- [28] “Improving Segmentation of Breast Arterial Calcifications from Digital Mammography: Good Annotation is All You Need”, en *Lecture Notes in Computer Science*, pp. 134–150, Cham: Springer Nature Switzerland, 2023, [doi:10.1007/978-3-031-27066-6_10](https://doi.org/10.1007/978-3-031-27066-6_10). ISSN: 0302-9743, 1611-3349.
- [29] Mobini, N., Codari, M., Riva, F., Ienco, M. G., Capra, D., Cozzi, A., Carriero, S., Spinelli, D., Trimboli, R. M., Baselli, G., y Sardanelli, F., “Detection and quantification of breast arterial calcifications on mammograms: a deep learning approach”, *European Radiology*, vol. 33, no. 10, pp. 6746–6755, 2023, [doi:10.1007/s00330-023-09668-z](https://doi.org/10.1007/s00330-023-09668-z).
- [30] Sakaida, M., Yoshimura, T., Tang, M., Ichikawa, S., y Sugimori, H., “Development of a Mammography Calcification Detection Algorithm Using Deep Learning with Resolution-Preserved Image Patch Division”, *Algorithms*, vol. 16, no. 10, p. 483, 2023, [doi:10.3390/a16100483](https://doi.org/10.3390/a16100483).
- [31] Maguire, D., Thompson, J. D., Vadera, S., y Szczepura, K., “Approaches to Automatic Classification, Detection and Segmentation of Breast Arterial Calcification Using Deep Learning”, *Expert Systems*, vol. 42, no. 6, 2025, [doi:10.1111/exsy.70069](https://doi.org/10.1111/exsy.70069). Publisher: Wiley.
- [32] Cardoso, M. J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., Nath, V., He, Y., Xu, Z., Hatamizadeh, A., Myronenko, A., Zhu, W., Liu, Y., Zheng, M., Tang, Y., Yang, I., Zephyr, M., Hashemian, B., Alle, S., Darestani, M. Z., Budd, C., Modat, M., Vercauteren, T., Wang, G., Li, Y., Hu, Y., Fu, Y., Gorman, B., Johnson, H., Genereaux, B., Erdal, B. S., Gupta, V., Diaz-Pinto, A., Dourson, A., Maier-Hein, L., Jaeger, P. F., Baumgartner, M., Kalpathy-Cramer, J., Flores, M., Kirby, J., Cooper, L. A. D., Roth, H. R., Xu, D., Bericat, D., Floca, R., Zhou, S. K., Shuaib, H., Farahani, K., Maier-Hein, K. H., Aylward, S., Dogra, P., Ourselin, S., y Feng, A., “MONAI: An open-source framework for deep learning in healthcare”, 2022, [doi:10.48550/arXiv.2211.02701](https://doi.org/10.48550/arXiv.2211.02701). arXiv:2211.02701 [cs].
- [33] Prokop, M., Shin, H.-O., Schanz, A., y Schaefer-Prokop, C. M., “Use of maximum intensity projections in ct angiography: a basic review”, *Radiographics*, vol. 17, no. 2, pp. 433–451, 1997, [doi:10.1148/radiographics.17.2.9084083](https://doi.org/10.1148/radiographics.17.2.9084083).
- [34] Paulina Neira, V., “Densidad mamaria y riesgo de cáncer mamario”, *Revista Médica Clínica Las Condes*, vol. 24, no. 1, pp. 122–130, 2013, [doi:10.1016/s0716-8640\(13\)70137-8](https://doi.org/10.1016/s0716-8640(13)70137-8). Publisher: Elsevier BV.

Anexos

Anexo A. Resultados detallados por volumen

En esta sección se presentan los resultados detallados de los coeficientes de Dice obtenidos en cada volumen, organizados por función de pérdida, arquitectura y configuración de filtros. Estos resultados complementan los promedios reportados en el Capítulo 4.

A.1. Coeficientes de Dice para U-Net bajo distintas funciones de pérdida

Se presentan los coeficientes de Dice por volumen obtenidos usando la misma arquitectura U-Net y diferentes funciones de pérdida.

Tabla A.1: Coeficientes de Dice por volumen para U-Net bajo distintas funciones de pérdida.

Volumen	Dice Loss	Dice Focal Loss ($\alpha = 0,8,$ $\gamma = 2,0$)	Dice Focal Loss ($\alpha = 0,25,$ $\gamma = 2,0$)	Dice Focal Loss ($\alpha = 0,25,$ $\gamma = 1,0$)
DBT-P00001_MLOI	0,423	0,474	0,451	0,499
DBT-P00285_MLOI	0,390	0,166	0,096	0,097
DBT-P00303_MLOD	0,199	0,089	0,037	0,038
DBT-P00491_MLOD	0,356	0,274	0,119	0,157
DBT-P00523_MLOD	0,000	0,000	0,000	0,000
DBT-P00688_MLOD	0,044	0,130	0,008	0,178
DBT-P00887_MLOD	0,187	0,274	0,316	0,246
DBT-P01488_MLOD	0,033	0,069	0,110	0,047
DBT-P01762_MLOI	0,260	0,333	0,292	0,445
DBT-P01778_MLOI	0,012	0,027	0,144	0,075
DBT-P01957_MLOD	0,032	0,084	0,222	0,026
DBT-P02244_MLOD	0,086	0,163	0,498	0,289
DBT-P02510_MLOD	0,387	0,566	0,666	0,687
DBT-P04368_MLOD	0,366	0,068	0,078	0,131
DBT-P04396_MLOD	0,122	0,191	0,178	0,230
Promedio \pm Desviación Estándar	0,194 \pm 0,158	0,195 \pm 0,164	0,214 \pm 0,196	0,213 \pm 0,198

A.2. Coeficientes de Dice para distintas arquitecturas

Se presentan los coeficientes de Dice por volumen obtenidos usando la mejor función de pérdida encontrada y diferentes arquitecturas de redes neuronales convolucionales.

Tabla A.2: Coeficientes de Dice por volumen para distintas arquitecturas.

Volumen	U-Net	Attention U-Net	Swin UNetR
DBT-P00001_MLOI	0,451	0,563	0,146
DBT-P00285_MLOI	0,096	0,282	0,056
DBT-P00303_MLOD	0,037	0,130	0,024
DBT-P00491_MLOD	0,119	0,207	0,012
DBT-P00523_MLOD	0,000	0,000	0,000
DBT-P00688_MLOD	0,008	0,389	0,045
DBT-P00887_MLOD	0,316	0,448	0,070
DBT-P01488_MLOD	0,110	0,242	0,018
DBT-P01762_MLOI	0,292	0,464	0,083
DBT-P01778_MLOI	0,144	0,261	0,013
DBT-P01957_MLOD	0,222	0,255	0,068
DBT-P02244_MLOD	0,498	0,443	0,058
DBT-P02510_MLOD	0,666	0,689	0,123
DBT-P04368_MLOD	0,078	0,356	0,026
DBT-P04396_MLOD	0,178	0,453	0,118
Promedio \pm Desviación Estándar	0,214 \pm 0,196	0,346 \pm 0,175	0,057 \pm 0,045

A.3. Coeficientes de Dice para diferentes filtros

Se presentan los coeficientes de Dice por volumen obtenidos usando la mejor arquitectura y función de pérdida encontrada luego de ser filtrados por máscaras de arterias obtenidas con distintos segmentadores.

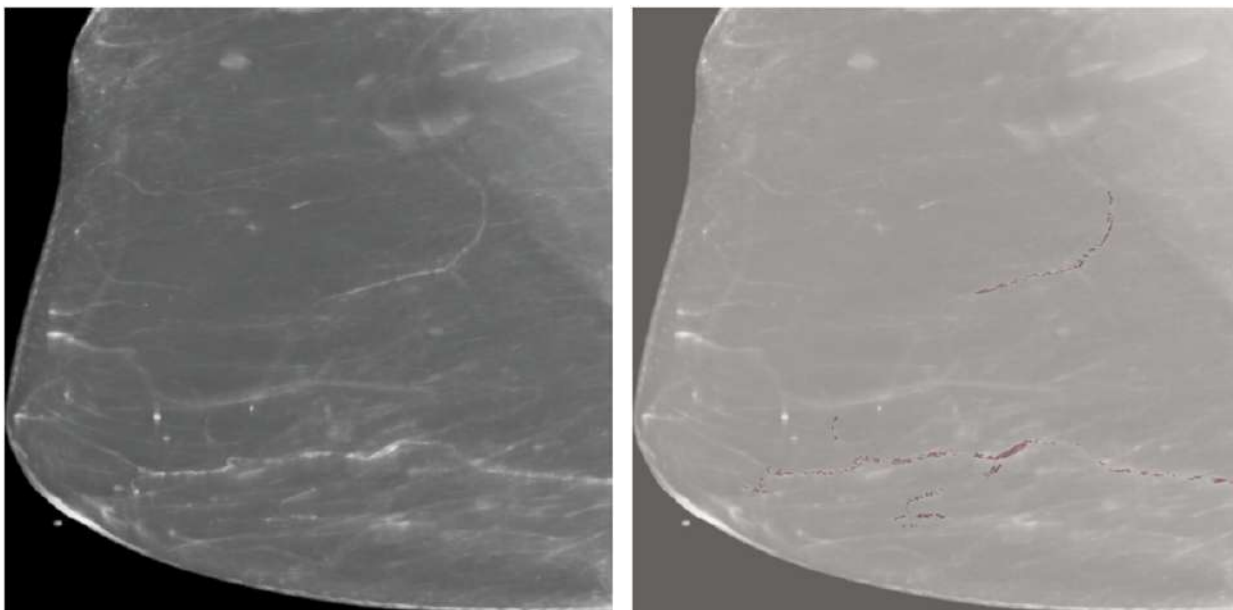
Tabla A.3: Coeficientes de Dice para diferentes filtros

Volumen	Segmentador de arterias calcificadas usando sólo su etiqueta	Segmentador de arterias general usando etiquetas de arterias calcificadas y no calcificadas	Segmentador de arterias calcificadas usando ambas etiquetas de arterias
DBT-P00001_MLOI	0,575	0,561	0,583
DBT-P00285_MLOI	0,261	0,066	0,237
DBT-P00303_MLOD	0,126	0,144	0,093
DBT-P00491_MLOD	0,200	0,176	0,168
DBT-P00523_MLOD	0,000	0,000	0,000
DBT-P00688_MLOD	0,391	0,456	0,296
DBT-P00887_MLOD	0,525	0,497	0,502

DBT-P01488_MLOD	0,262	0,313	0,197
DBT-P01762_MLOI	0,481	0,355	0,512
DBT-P01778_MLOI	0,269	0,289	0,288
DBT-P01957_MLOD	0,336	0,222	0,318
DBT-P02244_MLOD	0,451	0,460	0,460
DBT-P02510_MLOD	0,722	0,700	0,724
DBT-P04368_MLOD	0,221	0,140	0,223
DBT-P04396_MLOD	0,436	0,352	0,453
Promedio ± Desviación Estándar	0,351 ± 0,187	0,315 ± 0,196	0,337 ± 0,197

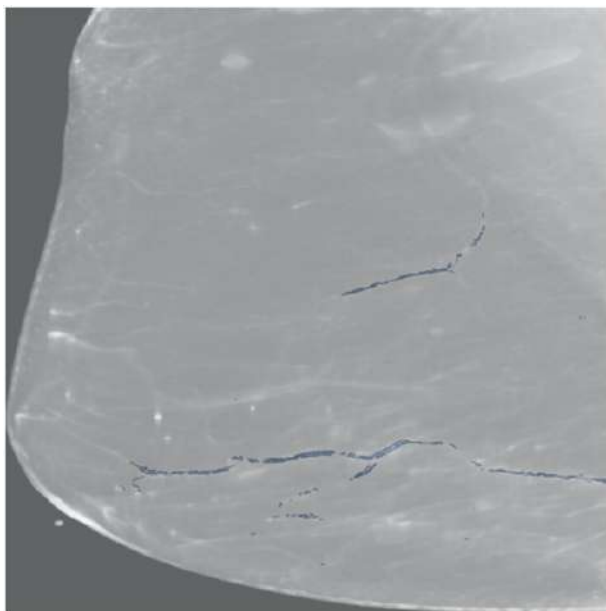
Anexo B. Predicciones del mejor modelo desarrollado

En este anexo se presentan las predicciones de todos los volúmenes del conjunto de validación generadas por el modelo con mejor arquitectura y función de pérdida, filtradas con la máscara de arterias calcificadas. Estas imágenes complementan los ejemplos representativos mostrados en el Capítulo 4.



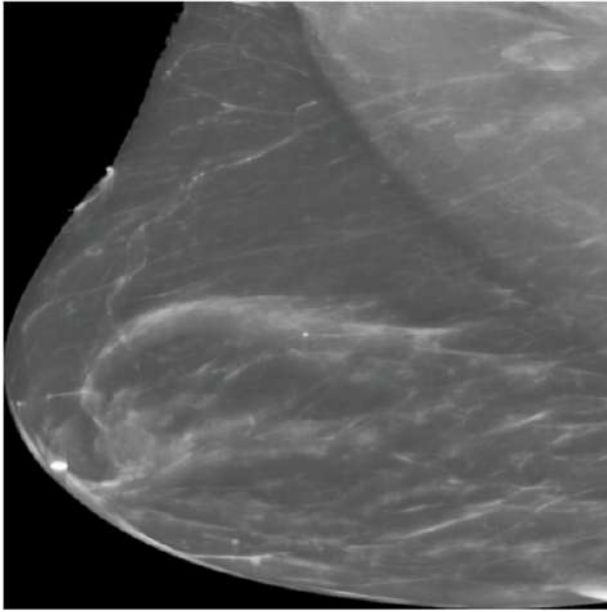
(a) MIP de la mama.

(b) Etiqueta de BACs.

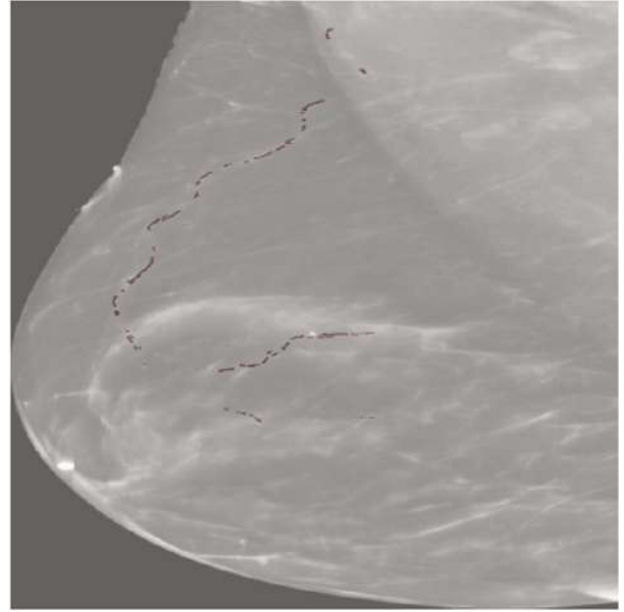


(c) Predicción filtrada.

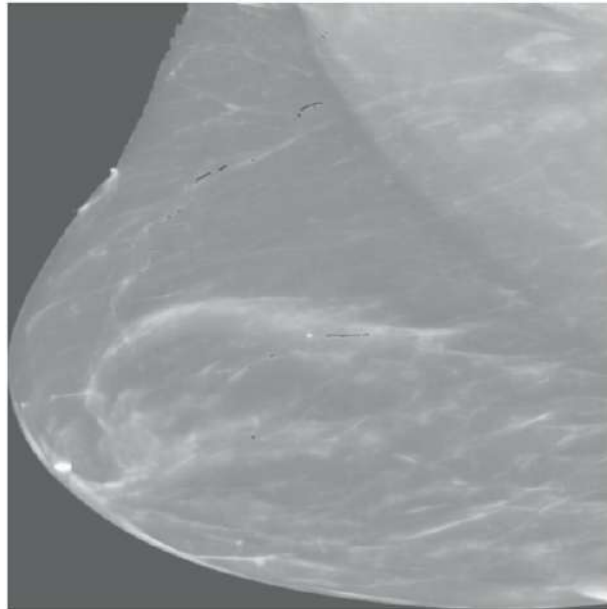
Figura B.1: Predicción del volumen DBT-P00001_MLOI con coeficiente de Dice 0,575.



(a) MIP de la mama.

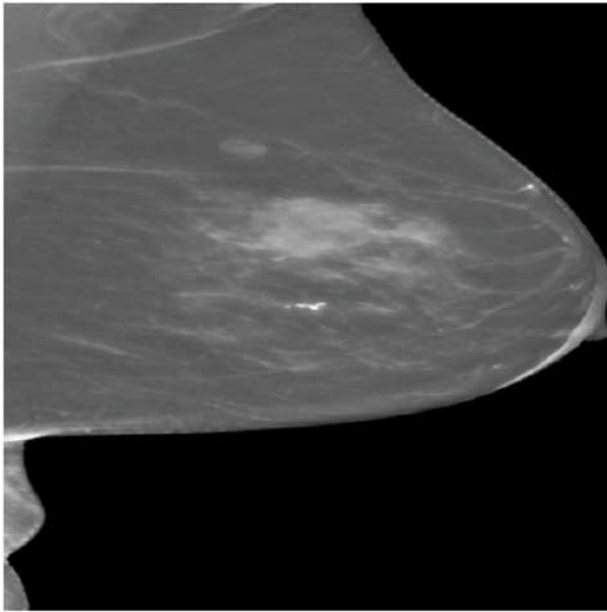


(b) Etiqueta de BACs.

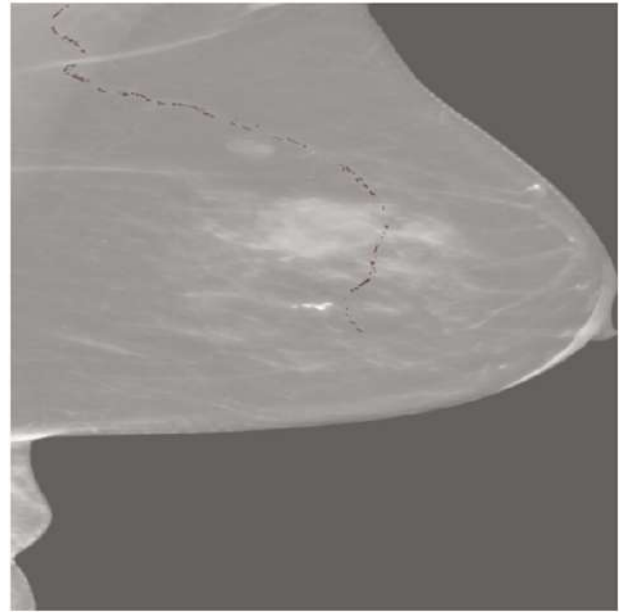


(c) Predicción filtrada.

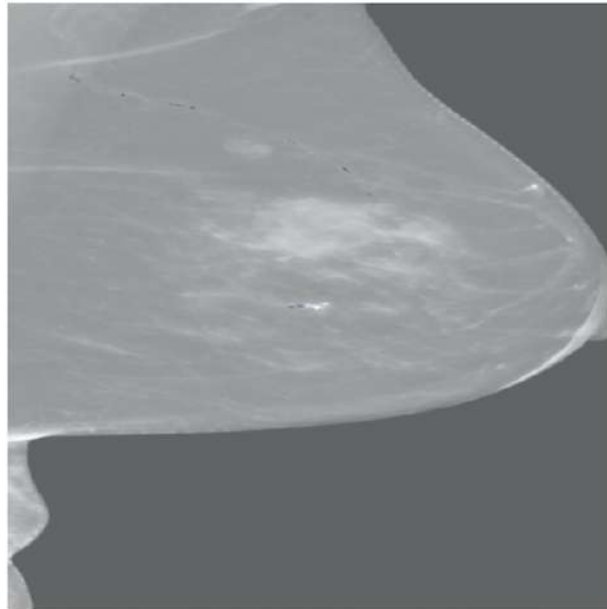
Figura B.2: Predicción del volumen DBT-P00285_MLOI con coeficiente de Dice 0,261.



(a) MIP de la mama.

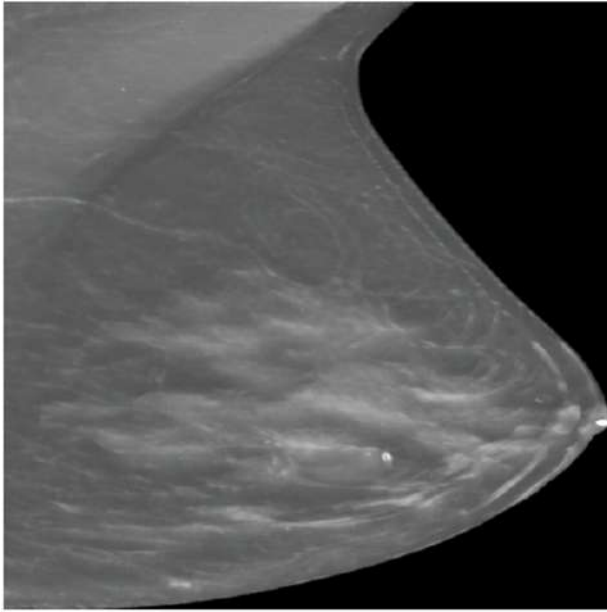


(b) Etiqueta de BACs.

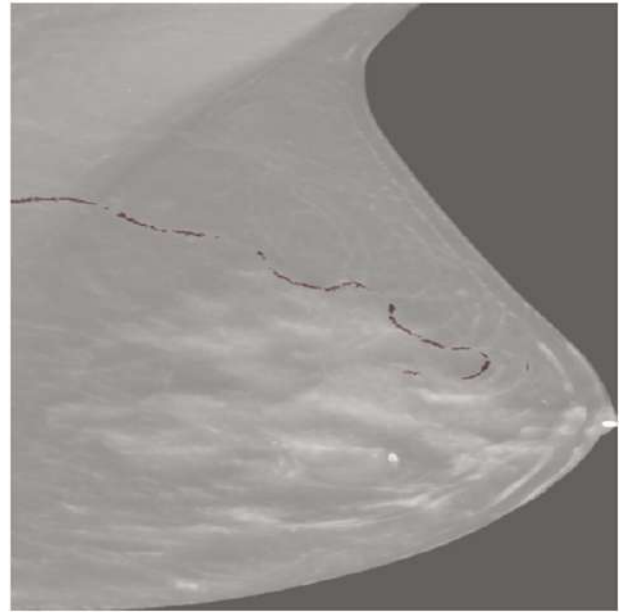


(c) Predicción filtrada.

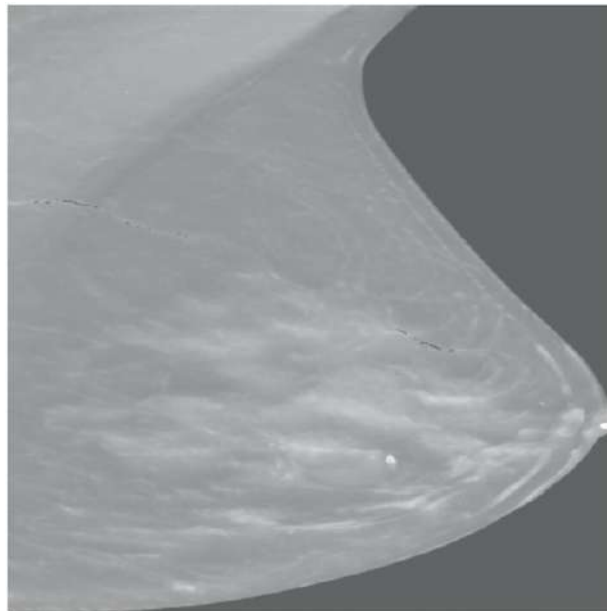
Figura B.3: Predicción del volumen DBT-P00303_MLOD con coeficiente de Dice 0,126.



(a) MIP de la mama.

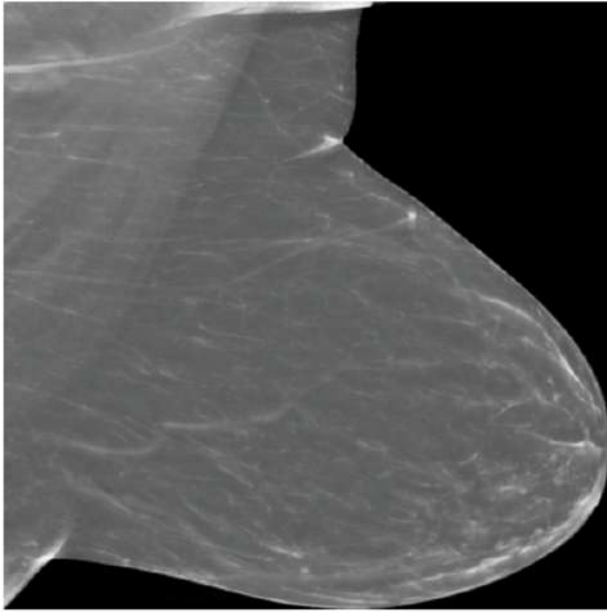


(b) Etiqueta de BACs.

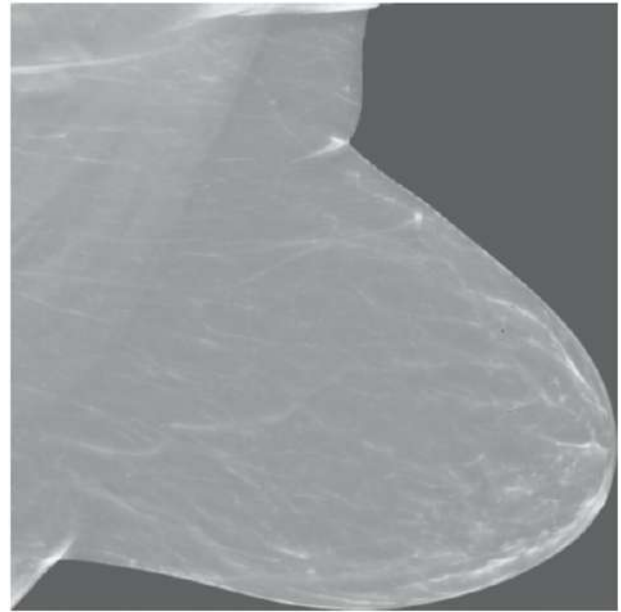


(c) Predicción filtrada.

Figura B.4: Predicción del volumen DBT-P00491_MLOD con coeficiente de Dice 0,200.

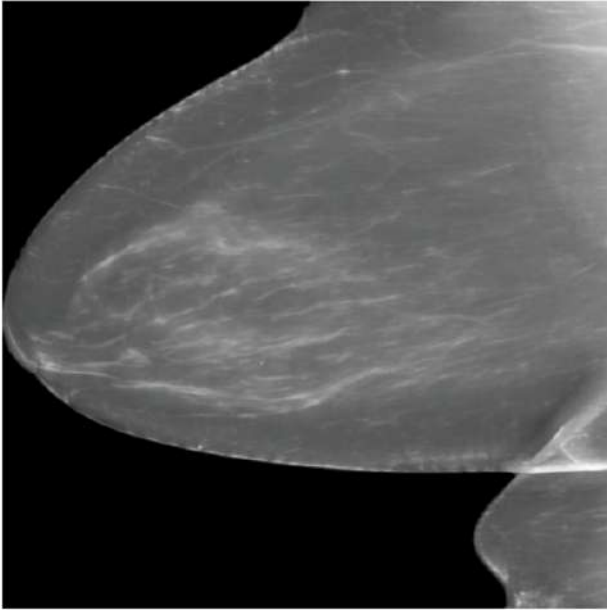


(a) MIP de la mama.

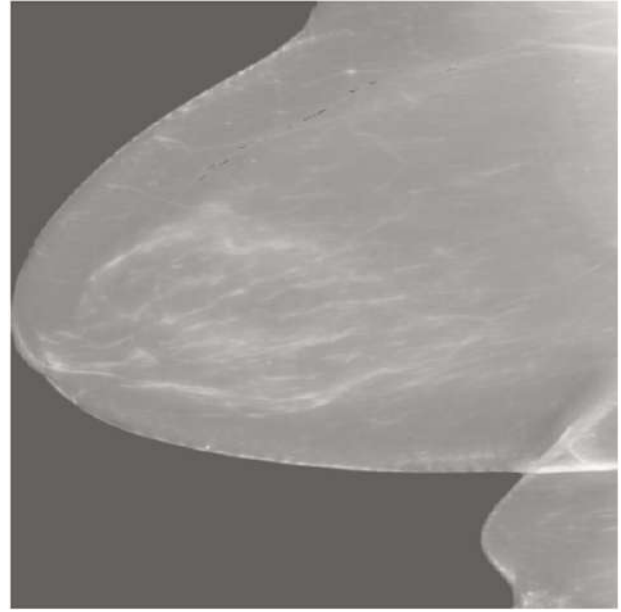


(b) Predicción filtrada.

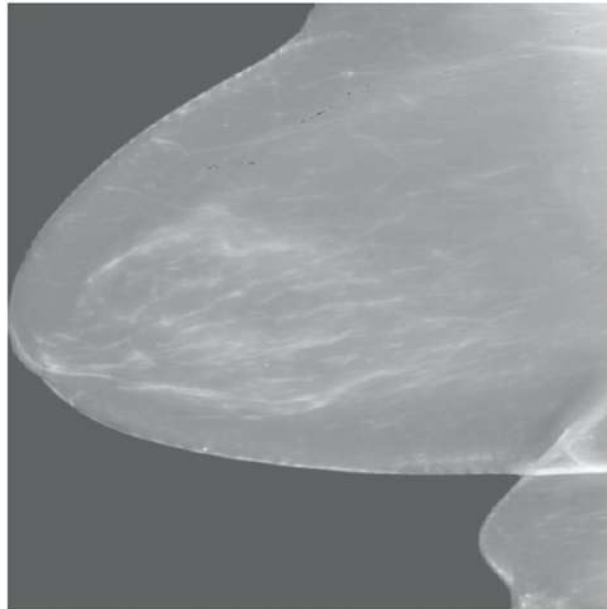
Figura B.5: Predicción del volumen DBT-P00523_MLOD, de etiqueta nula, con coeficiente de Dice 0,000.



(a) MIP de la mama.

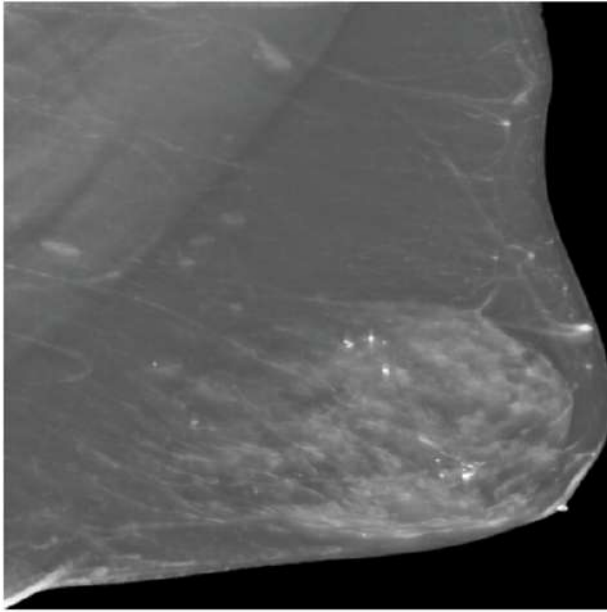


(b) Etiqueta de BACs.

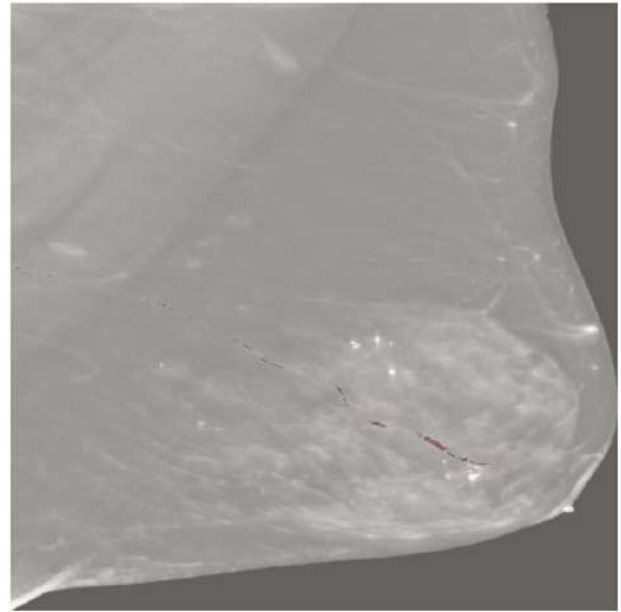


(c) Predicción filtrada.

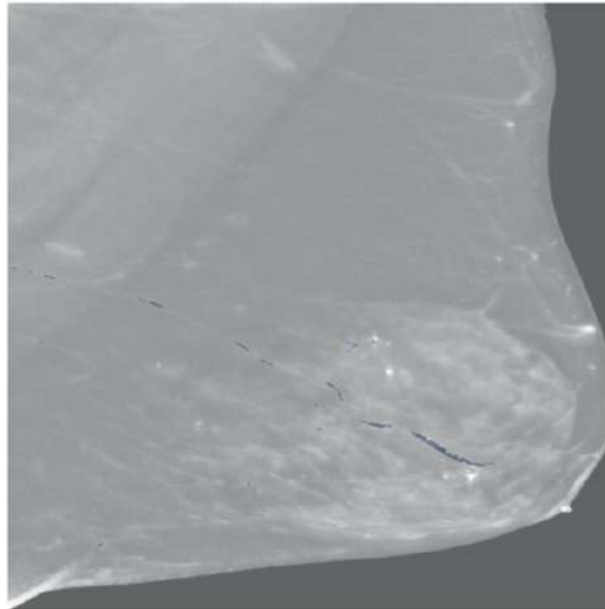
Figura B.6: Predicción del volumen DBT-P00688_MLOD con coeficiente de Dice 0,391.



(a) MIP de la mama.

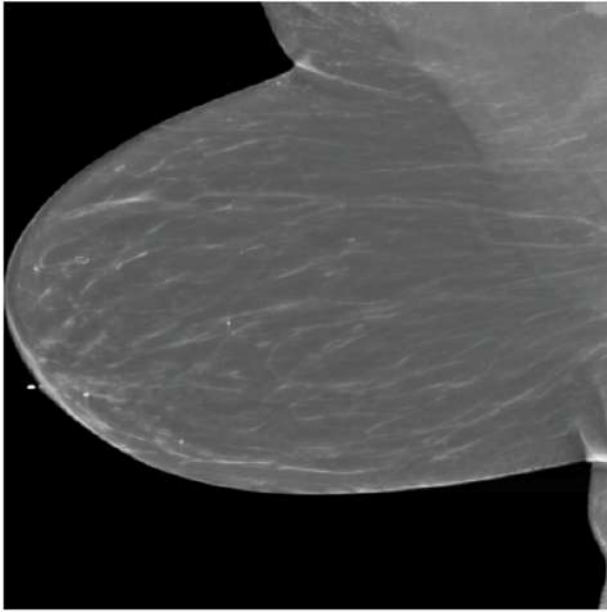


(b) Etiqueta de BACs.

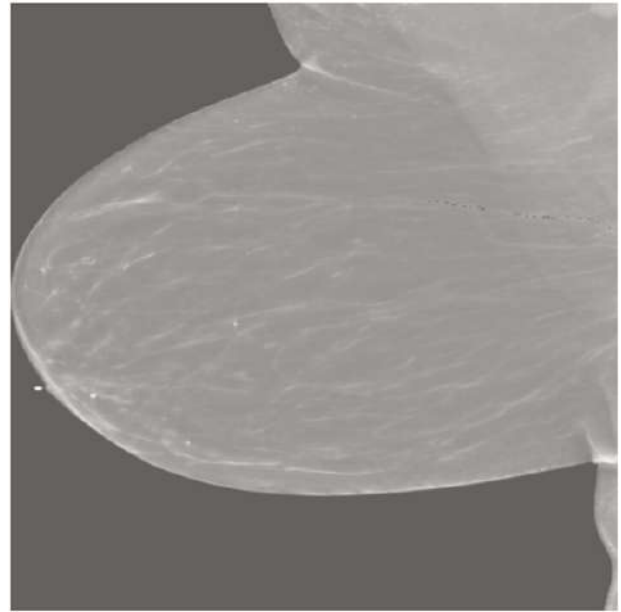


(c) Predicción filtrada.

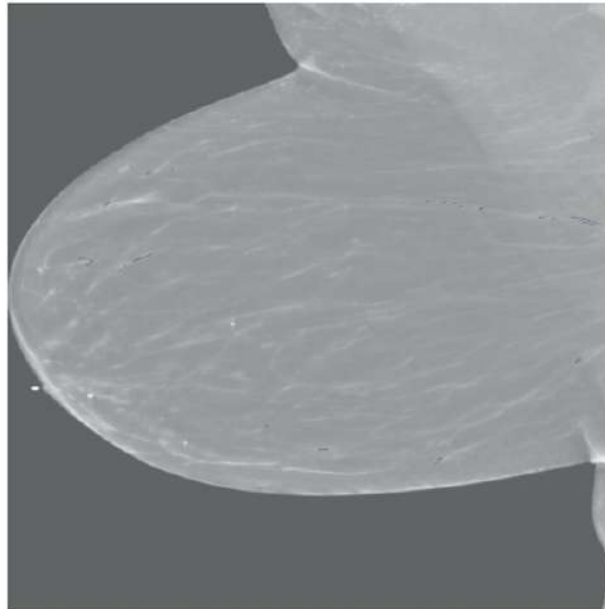
Figura B.7: Predicción del volumen DBT-P00887_MLOD con coeficiente de Dice 0,525.



(a) MIP de la mama.

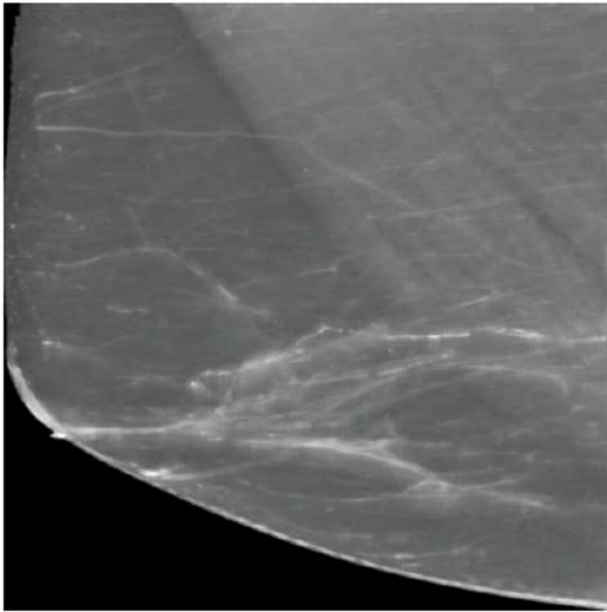


(b) Etiqueta de BACs.

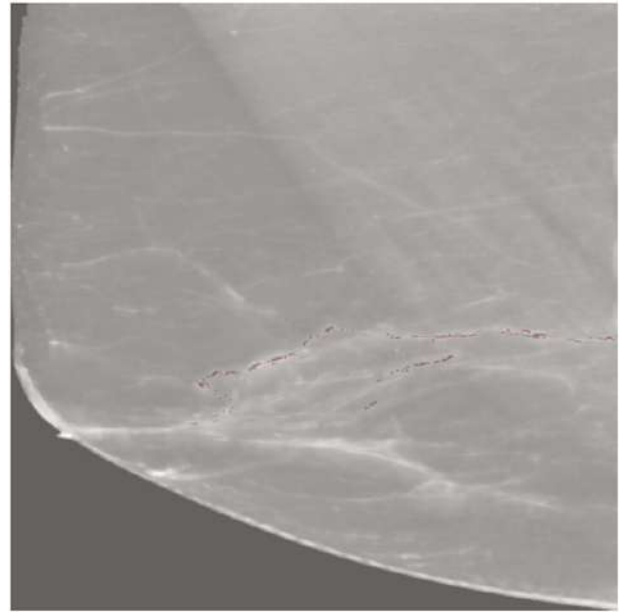


(c) Predicción filtrada.

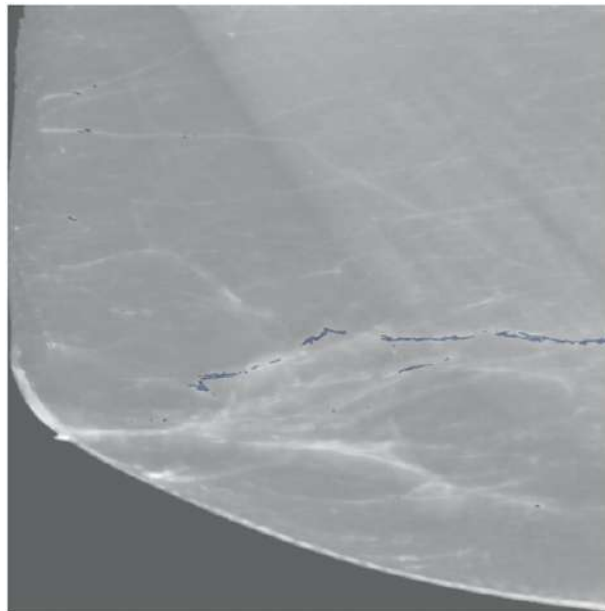
Figura B.8: Predicción del volumen DBT-P01488_MLOD con coeficiente de Dice 0,262.



(a) MIP de la mama.

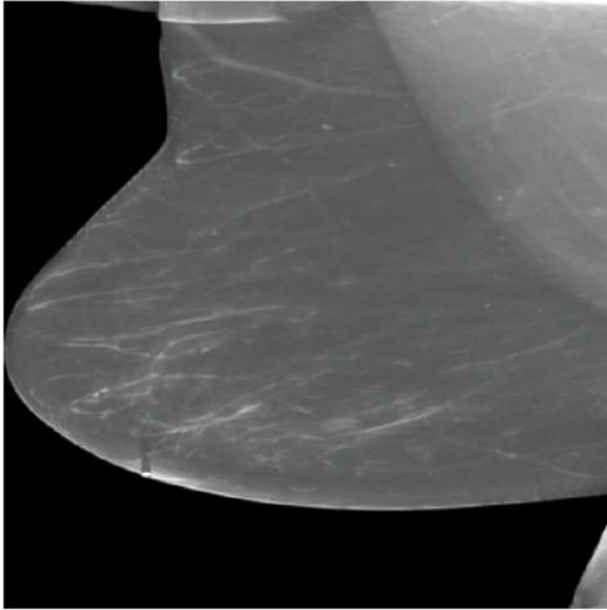


(b) Etiqueta de BACs.

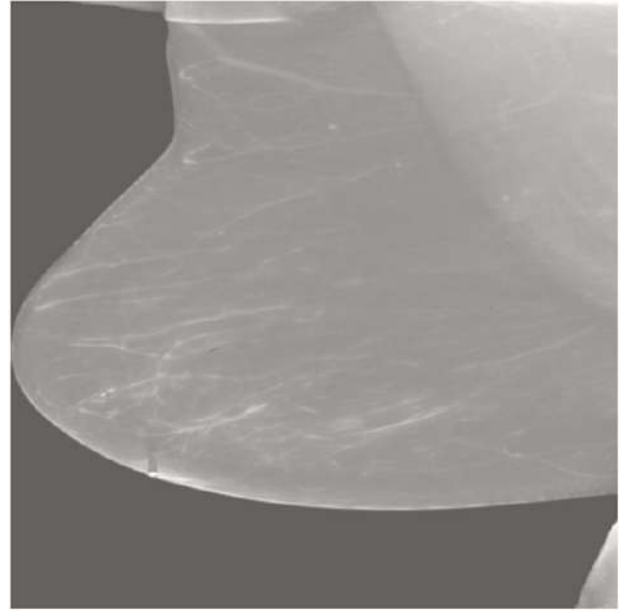


(c) Predicción filtrada.

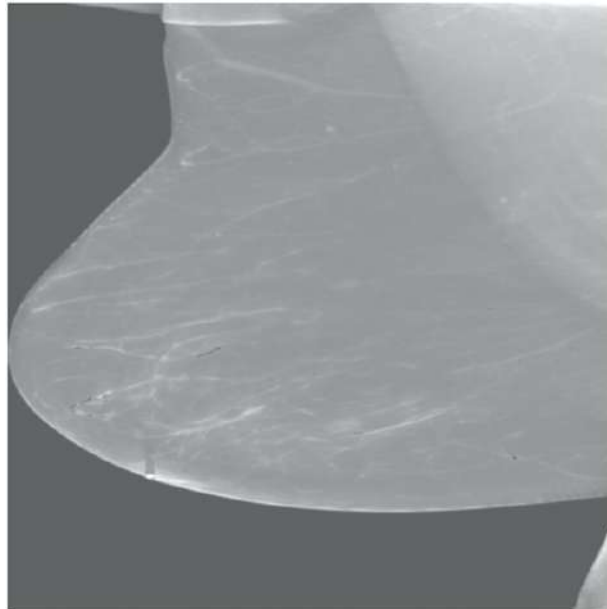
Figura B.9: Predicción del volumen DBT-P01762_MLOI con coeficiente de Dice 0,481.



(a) MIP de la mama.

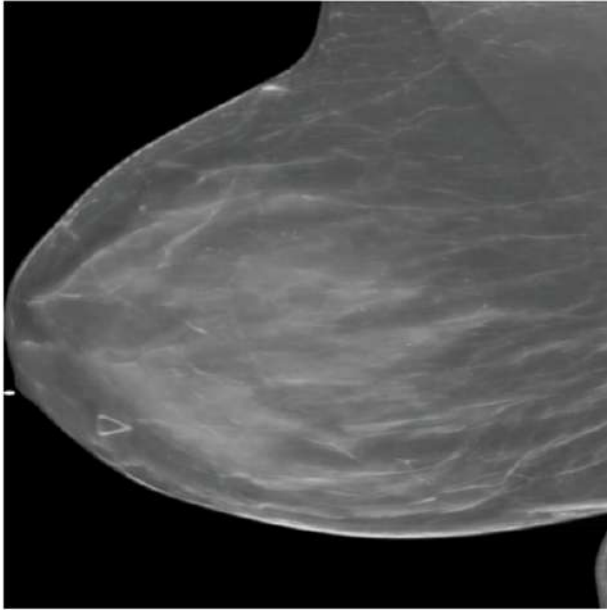


(b) Etiqueta de BACs.

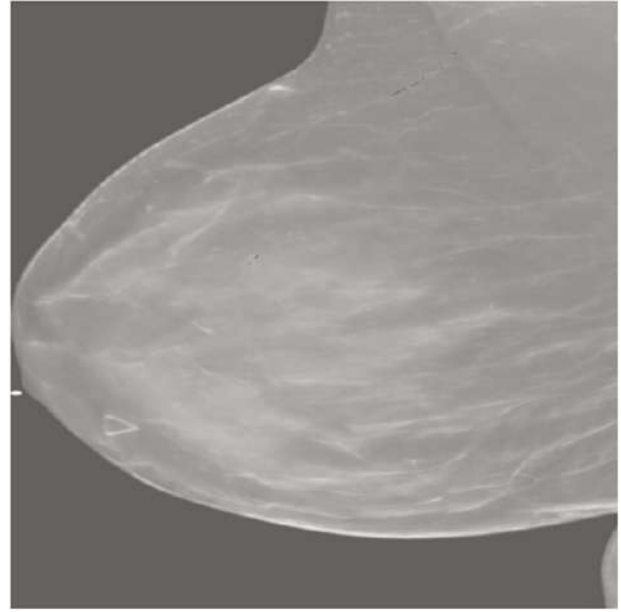


(c) Predicción filtrada.

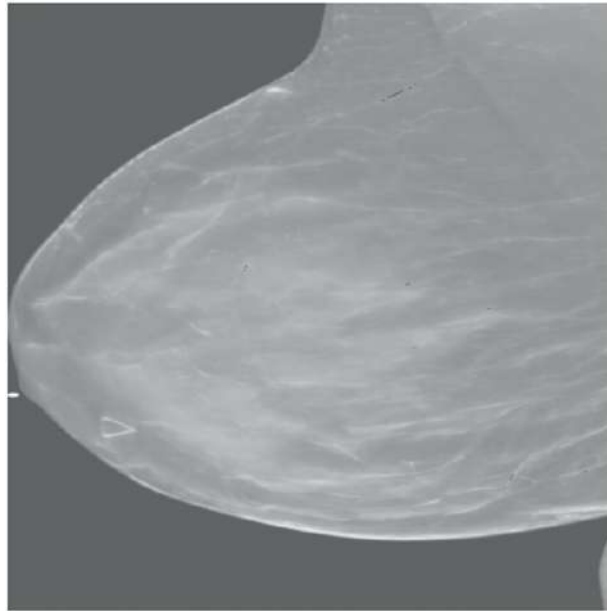
Figura B.10: Predicción del volumen DBT-P01778_MLOI con coeficiente de Dice 0,269.



(a) MIP de la mama.

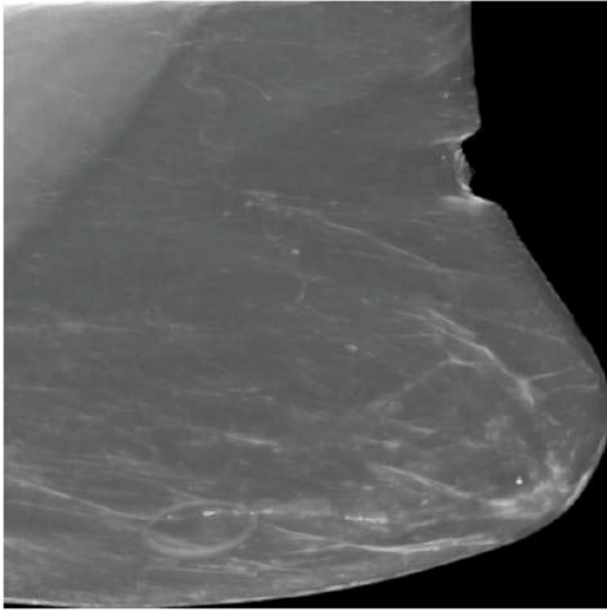


(b) Etiqueta de BACs.

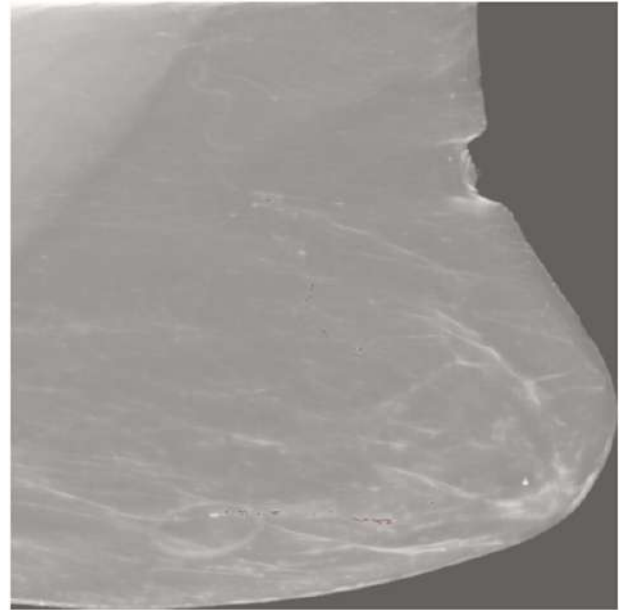


(c) Predicción filtrada.

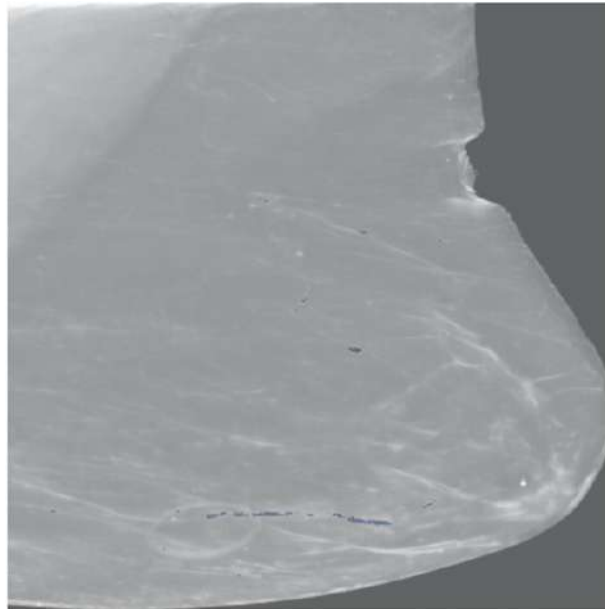
Figura B.11: Predicción del volumen DBT-P01957_MLOD con coeficiente de Dice 0,336.



(a) MIP de la mama.

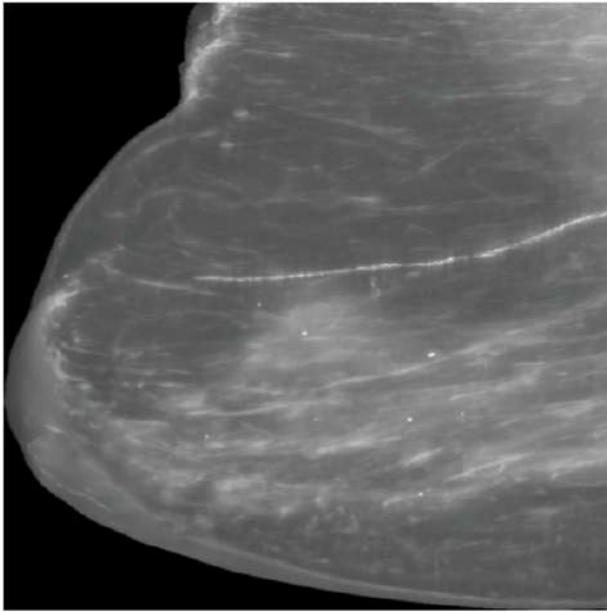


(b) Etiqueta de BACs.

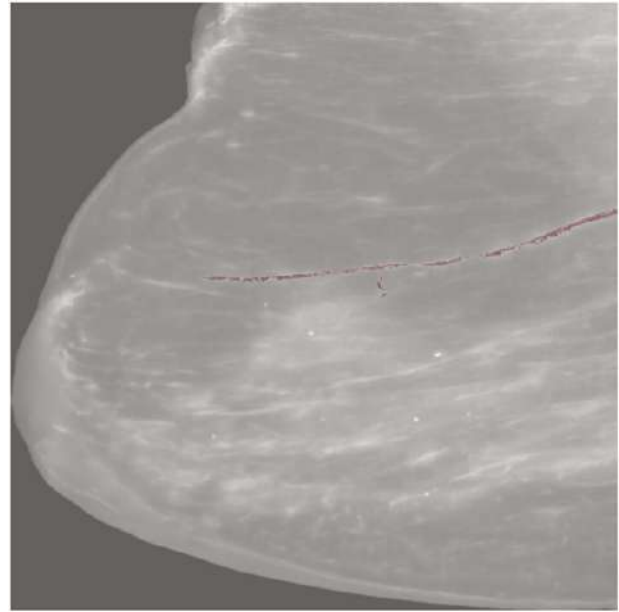


(c) Predicción filtrada.

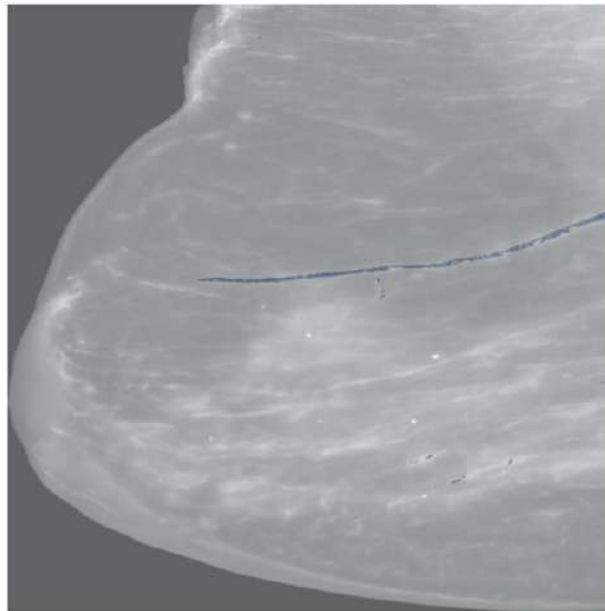
Figura B.12: Predicción del volumen DBT-P02244_MLOD con coeficiente de Dice 0,451.



(a) MIP de la mama.

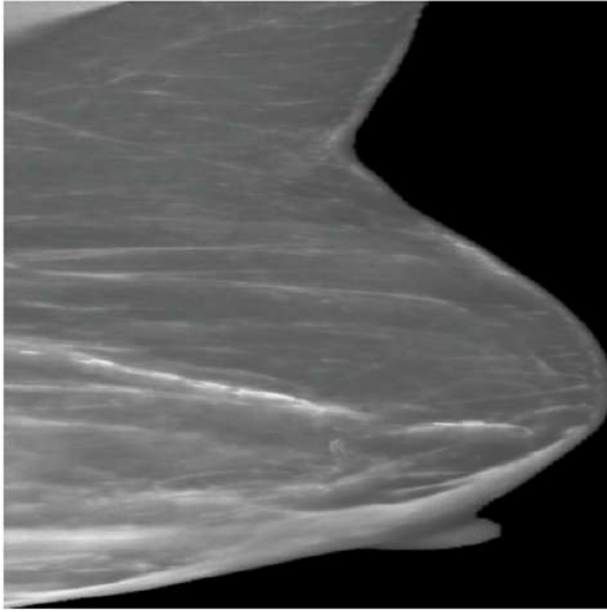


(b) Etiqueta de BACs.

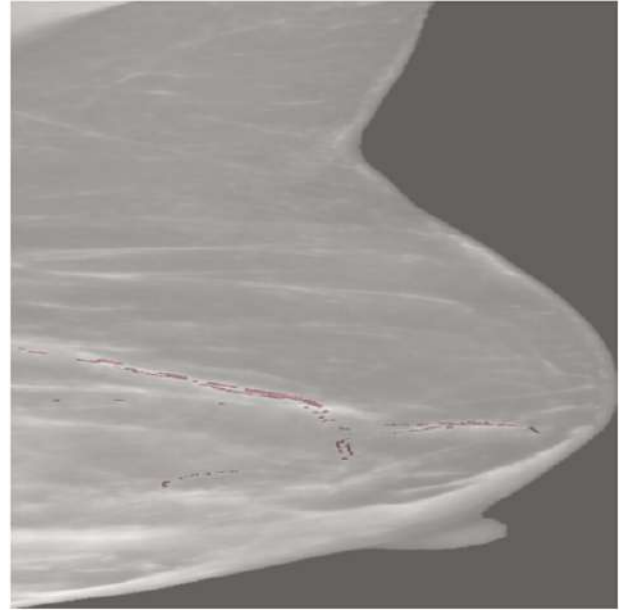


(c) Predicción filtrada.

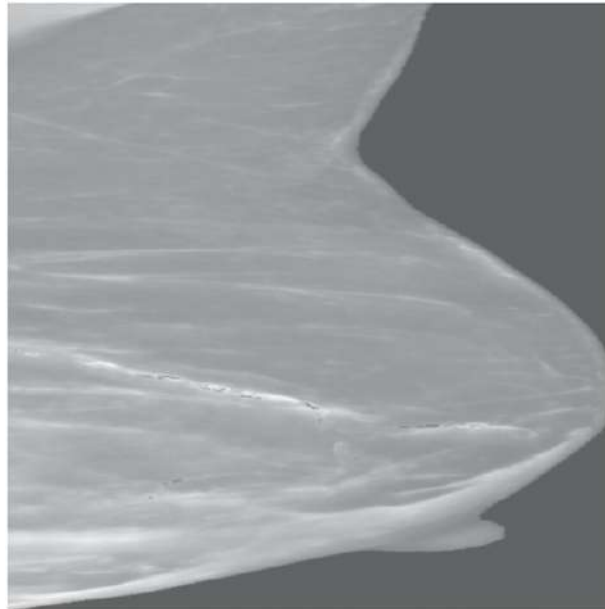
Figura B.13: Predicción del volumen DBT-P02510_MLOD con coeficiente de Dice 0,722.



(a) MIP de la mama.

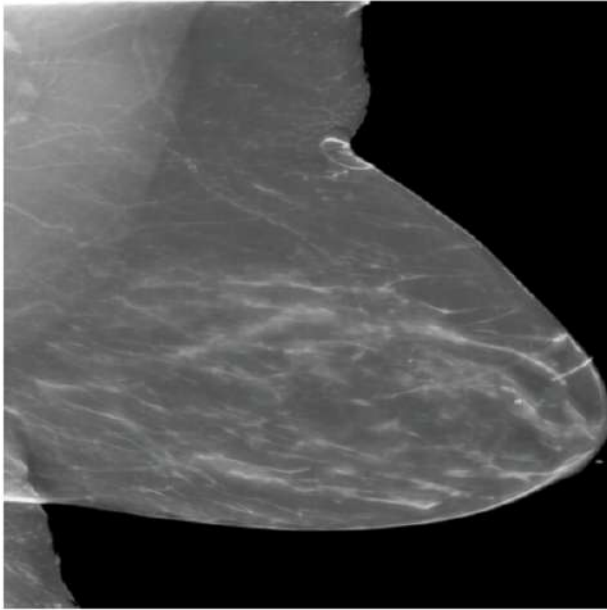


(b) Etiqueta de BACs.

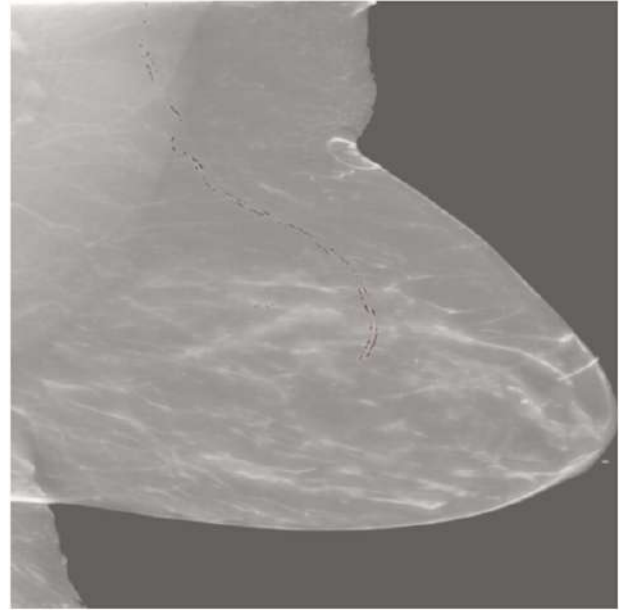


(c) Predicción filtrada.

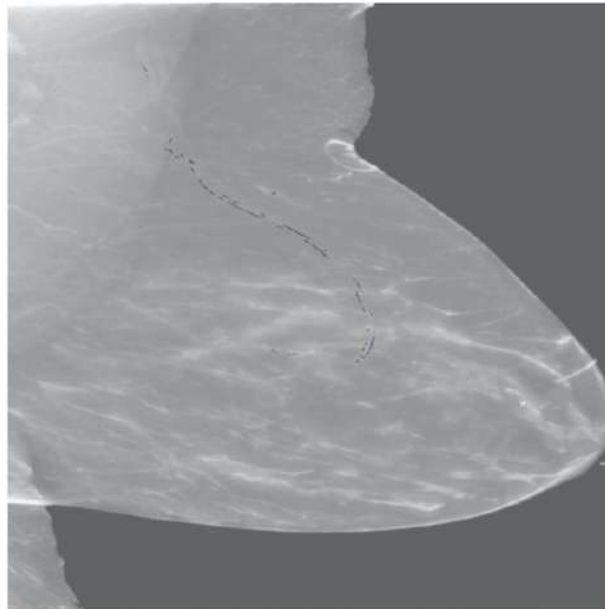
Figura B.14: Predicción del volumen DBT-P04368_MLOD con coeficiente de Dice 0,221.



(a) MIP de la mama.



(b) Etiqueta de BACs.



(c) Predicción filtrada.

Figura B.15: Predicción del volumen DBT-P04396_MLOD con coeficiente de Dice 0,436.