



ESCUELA DE INGENIERÍA DE FUENLABRADA

GRADO EN INGENIERÍA BIOMÉDICA

TRABAJO FIN DE GRADO

**IDENTIFICACIÓN RADIÓMICA DE FALSOS NEGATIVOS
EN GANGLIOS CENTINELA METASTÁSICOS CONFIRMADOS
POR LA TÉCNICA DE GANGLIO CENTINELA**

Autor: María de Castro Lozano

Tutor: Ángel Torrado Carvajal

Co-tutor: Irene Vicente Zapata

Curso académico 2022/2023

Agradecimientos

Para empezar me gustaría agradecer a mi tutor Ángel Torrado Carvajal por darme la oportunidad de realizar este trabajo y profundizar más en el mundo del machine learning y la radiómica. Agradecer también a Borja Rodríguez Vila, por estar presente y guiarme durante todo el proyecto.

También quiero agradecer a la Dra. Irene Vicente Zapata y Dra. Teresa Presa Abós, por mostrarme su pasión por la medicina y la investigación, y por su disponibilidad estos meses.

Gracias a las amigas que me ha dado esta etapa académica, por compartir preocupaciones, horas de estudio y apoyarnos constantemente. Con personas como vosotras a mi lado, las cosas son mucho más fáciles.

A mis padres, por dárme todo para llegar hasta aquí, por confiar siempre en mí más que yo misma y por celebrar cada uno de mis logros, esto también es vuestro. A mi hermana, por estar siempre para escucharme, sacarme de mis preocupaciones y alegrarme los días.

A Jorge, por acompañarme, entenderme cuando ni yo me entendía y tener siempre las palabras adecuadas. Por darme espacio cuando lo necesitaba y también apoyo durante todos los días de estos 4 años.

Resumen

El cáncer de mama es el cáncer más común en las mujeres y la primera causa de muerte por cáncer. Por su cercanía, los ganglios linfáticos axilares son la principal vía de diseminación del cáncer de mama. Su diagnóstico supone un antes y después en el tratamiento de la paciente y su pronóstico.

La resonancia magnética es la técnica por imagen que mayor resolución tiene del interior del cuerpo y, por lo tanto, la que más información aporta al diagnóstico. En ocasiones para los radiólogos es difícil detectar ganglios metastásicos que morfológicamente son considerados sanos en estas imágenes. Por ello, la biopsia del ganglio centinela es la técnica invasiva usada para corroborar la afectación ganglionar.

El machine learning y la radiómica son ciencias con un gran potencial en la investigación y en la mejora del diagnóstico. Así pues, ayudado de ellas, el objetivo de este estudio es el desarrollo de una herramienta que permita detectar ganglios aparentemente sanos para así evitar cualquier intervención.

Inicialmente se realiza una selección de imágenes de dos modalidades, ponderadas en T1 y del estudio dinámico de 7 minutos. Se realiza una segmentación manual de los ganglios, extracción y selección de las características radiómicas. A continuación, se realizan cuatro distintas clasificaciones de los ganglios para evaluar el rendimiento utilizando diferentes bases de datos. Estas clasificaciones se llevan a cabo mediante el entrenamiento de un modelo *Random Forest* usando *Cross Validation*.

Una vez obtenidas las métricas de rendimiento para las 4 clasificaciones, se concluye que el mejor clasificador es el que categoriza ganglios patológicos vs. ganglios sanos utilizando imágenes ponderadas en T1 e imágenes del estudio dinámico. Este clasificador tiene una exactitud de 0.8, una precisión de 0.783, una sensibilidad de 0.867, una especificidad de 0.733 y el área bajo la curva ROC es de 0.82.

Además, gracias al algoritmo SHAP, se ha podido dar explicabilidad a la clasificación y entender cuáles han sido las características más relevantes en la misma. Éstas están relacionadas con la homogeneidad y asimetría de la imagen, y han sido extraídas principalmente de las imágenes del estudio dinámico. Además, las características más importantes han sido extraídas a partir de distintos filtros aplicados a la imagen original.

Este estudio evidencia el potencial de la radiómica unida al machine learning. Éstos son capaces de extraer características inapreciables al ojo humano y clasificarlas, pudiendo llegar a desarrollar herramientas de apoyo al radiólogo durante el diagnóstico de ganglios metastásicos.

Índice general

Agradecimientos

Resumen

Índice de figuras **V**

Índice de tablas **VII**

Lista de acrónimos y abreviaturas **IX**

1. Introducción y contexto clínico **1**

1.1. Definición del Cáncer de Mama y Anatomía de la Mama 1

1.2. Factores de riesgo del Cáncer de Mama 2

1.3. Tipos y prevalencia de Cáncer de Mama 3

1.4. Diagnóstico y estadificación del Cáncer de Mama 3

1.5. Tratamientos del Cáncer de Mama 5

1.6. Ganglios linfáticos axilares y biopsia del ganglio centinela 6

1.7. Resonancia Magnética en el Cáncer de Mama 8

2. Marco Teórico y Estado del Arte **13**

2.1. Machine Learning 13

2.2. Radiómica 14

2.3. Estado del Arte	15
3. Motivación y Objetivos	17
3.1. Hipótesis	17
3.2. Motivación	17
3.3. Planificación y objetivos específicos	18
4. Materiales	21
4.1. Creación de la base de datos	21
4.2. Herramientas utilizadas	22
4.2.1. 3D Slicer	22
4.2.2. Python	22
5. Métodos	25
5.1. Selección y extracción de Imágenes	25
5.2. Segmentación de Ganglios Axilares	26
5.3. Extracción de Características Radiómicas	29
5.4. Selección de Características Radiómicas	31
5.5. Creación de cuatro conjuntos de datos	32
5.6. Clasificación mediante Machine Learning	33
5.7. Evaluación de la Clasificación	36
5.8. Estudio de las Características Radiómicas	38
6. Resultados	39
6.1. Resultados de la evaluación de los Clasificadores	39
6.2. Resultados del estudio de las Características Radiómicas	41
7. Discusión	45
7.1. Rendimiento de los Clasificadores	45
7.2. Importancia de las Características Radiómicas	46

8. Conclusiones	49
8.1. Aportaciones del estudio	49
8.2. Líneas futuras	50
A. Anexo	51
Bibliografía	53

Índice de figuras

1.1. Anatomía de la mama desde la parte externa y la parte interna [1].	2
1.2. Biopsia del Ganglio Centinela [2].	8
1.3. Imagen de RM de la base de datos ponderada en T1.	10
1.4. Imagen de RM de la base de datos ponderada en T2.	10
1.5. Imagen de RM de la base de datos de un estudio dinámico.	11
1.6. Imagen de RM de la base de datos de difusión.	11
2.1. Esquema representativo de las etapas del proceso de desarrollo de un modelo ML [3].	14
3.1. Diagrama de Gantt.	19
4.1. Página principal del software 3D Slicer con la visualización de un paciente. . .	23
5.1. Bases de datos DICOM cargada en el software 3D Slicer.	26
5.2. Visualización en 3D Slicer de una imagen ponderada en T1 a la derecha y la barra de herramientas de <i>Segment Editor</i> a la izquierda.	27
5.3. Ejemplo de la segmentación de 2 ganglios en imágenes ponderadas en T1. . . .	28
5.4. Ejemplo de la segmentación de 2 ganglios en imágenes del minuto 1 del estudio dinámico.	28
5.5. Ejemplo visual de un árbol de decisión con nodo principal, nodos de decisión y nodos hoja [4].	33
5.6. Ejemplo visual del algoritmo Random Forest [5].	34
5.7. Ejemplo visual del método k-Fold con 5 subconjuntos [6].	36

5.8. Ejemplo visual de varias curvas ROC [7].	38
5.9. Ejemplo visual del algoritmo SHAP.	38
6.1. Curva ROC y valor AUC para cada clasificador.	40
6.2. Gráficos con las 5 características más importantes empleadas en cada clasificador	42

Índice de tablas

1.1. Clasificación BI-RADS	4
4.1. Datos de las imágenes de RM utilizadas para la clasificación	22
6.1. Figuras de Mérito para cada uno de los clasificadores.	40

Lista de acrónimos y abreviaturas

BGC Biopsia de Ganglio Centinela

BI-RADS Sistema de Base de Datos e Informes de Imágenes de la Mama

CD Carcinoma ductal

CL Carcinoma lobulillar

CM Cáncer de Mama

CNN Red Neuronal Convolucional

CV Cross Validation

GC Ganglio Centinela

HURyC Hospital Universitario Ramón y Cajal

IA Inteligencia Artificial

ML Machine Learning

PACS Sistema de archivo y comunicación de imágenes

RM Resonancia Magnética

ROI Region of Interest

TFG Trabajo Fin de Grado

Capítulo 1

Introducción y contexto clínico

1.1. Definición del Cáncer de Mama y Anatomía de la Mama

El cáncer de mama (CM) se debe a la proliferación incontrolada y rápida de células del epitelio glandular de la mama. La capacidad reproductiva de estas células cancerosas es excesiva, pudiendo llegar a diseminarse y alcanzar otros tejidos a través de la sangre o los vasos linfáticos cercanos. Allí, si las células se adhieren y proliferan, puede formarse metástasis.

En cuanto a la anatomía, la mama está constituida por 10 o 20 grupos llamados lóbulos. Cada lóbulo está formado por unidades más pequeñas llamadas lobulillos. Los lobulillos tienen glándulas cuya función es producir la leche durante la lactancia. Los lobulillos están unidos al pezón por unos tubos llamados ductos, por donde fluye la leche. El espacio que hay entre los diferentes lobulillos y los ductos es ocupado por grasa y tejido fibroso [1].

Además, la axila es una de las zonas del cuerpo donde se acumulan más ganglios linfáticos y, por lo tanto, puede ser una de las principales vías de diseminación del cáncer. La mama tiene muchos vasos linfáticos que están conectados con los ganglios linfáticos axilares, cuya función es la protección del organismo, atrapar bacterias, células tumorales y sustancias nocivas.

En la Figura 1.1 se puede observar una ilustración de la anatomía de la mama y los ganglios axilares.

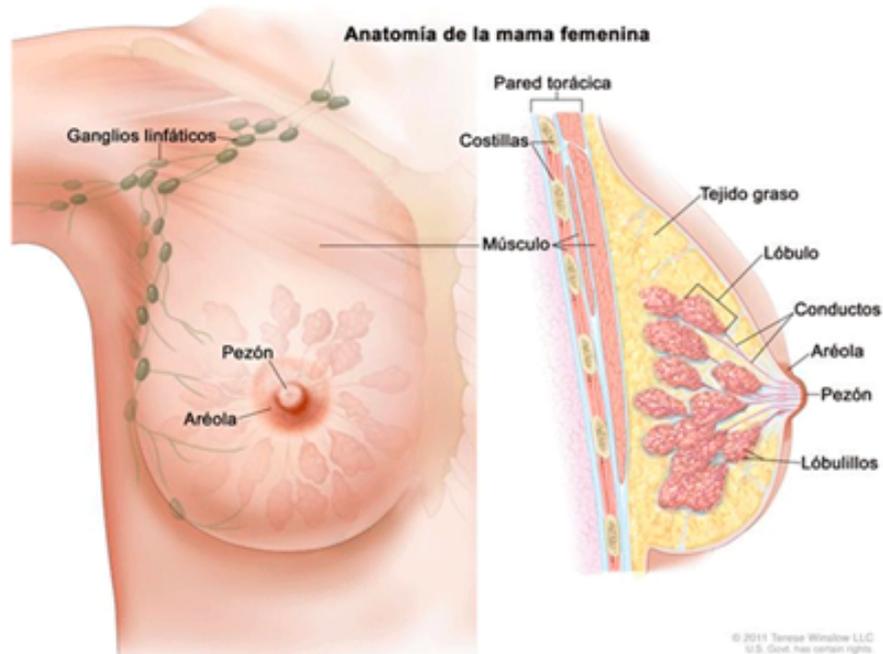


Figura 1.1: Anatomía de la mama desde la parte externa y la parte interna [1].

1.2. Factores de riesgo del Cáncer de Mama

Más del 99% de los casos de CM ocurren en mujeres. Es el cáncer más frecuente en la mujer española. Aproximadamente 1 de cada 8 mujeres españolas tendrá un CM en algún momento de su vida. Debido a los programas de cribado y la mejora en los tratamientos, la mortalidad por CM se ha visto disminuida, pero sigue siendo la primera causa de muerte por cáncer en las mujeres españolas [8].

Las causas de este cáncer no están claras, sin embargo, existen unos factores de riesgo asociados [8]:

- El principal factor de riesgo es la edad. La máxima incidencia está por encima de los 50 años, sin embargo, el 10% de los casos son en menores de 40 años.
- Densidad mamaria elevada en pruebas de imagen.
- Factores reproductivos como la aparición temprana de la primera regla, la menopausia tardía o no haber estado embarazada nunca.
- La exposición a radiaciones ionizantes y las mutaciones relacionadas con el CM.
- El estilo de vida, como puede ser el consumo de alcohol o la obesidad.

- Historia personal previa de CM y de hiperplasia atípica.
- El uso de terapia hormonal sustitutiva después de la menopausia.
- La predisposición genética. El cáncer hereditario se debe a alguna mutación genética, frecuentemente en los genes BRCA 1 y 2.

1.3. Tipos y prevalencia de Cáncer de Mama

Existe distintos tipos de CM que se diferencian por el lugar en donde empiezan a desarrollarse, su proliferación, diseminación y determinadas características sobre su comportamiento.

El CM que no se ha extendido y esta localizado en el tejido mamario donde comenzó, se denomina CM no invasivo o in situ. También puede denominarse afección precancerosa. Según la localización donde se desarrollan existen dos tipos [9]:

- El carcinoma ductal in situ, se originó en los ductos y no se extendió. No se considera mortal, pero es el precursor del CM invasivo. El 16 % de los diagnósticos son de este tipo.
- El carcinoma lobular in situ, se originó en los lobulillos y no se extendió. Es una afectación benigna y no es un CM real.

Sin embargo, el CM invasivo o infiltrante es el que se ha extendido a los tejidos mamarios cercanos. Al igual que en el no invasivo, existen varios tipos:

- El carcinoma ductal invasivo, comienza en los ductos e invade los tejidos colindantes, el 80 % de los diagnósticos son de este tipo, es el cáncer de mama más frecuente.
- El carcinoma lobular invasivo, comienza en los lobulillos e invade los tejidos colindantes, es el segundo CM más frecuente, constituye un 10 % de los casos.

1.4. Diagnóstico y estadificación del Cáncer de Mama

Para la prevención y el diagnóstico precoz existen diversos programas de cribado cuyo objetivo es la detección precoz del CM y así poder mejorar su pronóstico y evitar la mortalidad prematura o la discapacidad. Un ejemplo es el programa de detección precoz del CM de la Comunidad de Madrid (DEPRECAM) [10]. Este se dirige a mujeres con edades entre 50 y 69

años, asintomáticas. Consiste en la realización, de una mamografía cada dos años, que permite detectar el cáncer en etapas muy precoces, previniendo así el desarrollo de la enfermedad.

Estas mamografías deben ser analizadas y diagnosticadas por un experto radiólogo, para ello existe el Sistema de Base de Datos e Informes de Imágenes de la Mama (BI-RADS, por sus siglas en inglés) [11]. El Colegio Estadounidense de Radiología (ACR) estableció un sistema para realizar el diagnóstico de mamografías de manera uniforme en que se debe concluir con una puntuación del atlas BI-RADS. Consta de 7 categorías estándar que incluyen un plan para orientar a los facultativos sobre el seguimiento, se puede observar en la Tabla 1.1.

Categoría	Evaluación	Seguimiento
0	Incompleto	Antes de asignar una categoría, es necesario realizar pruebas de imagen complementarias.
1	Normal (negativo)	Seguir con las mamografías de control.
2	Lesión benigna (no cancerosa)	Seguir con las mamografías de control.
3	Lesión probablemente benigna	Cada 6 meses realizar una mamografía de seguimiento.
4	Lesión sospechosa (probablemente maligna)	Es recomendable una biopsia.
5	Altamente indicativo de cáncer	Requiere una biopsia.
6	Cáncer confirmado mediante biopsia	Se debe confirmar el cáncer mediante una biopsia antes de comenzar el tratamiento.

Tabla 1.1: Clasificación BI-RADS

Una vez se ha asignado una categoría BI-RADS a la prueba diagnóstica realizada, si es suficiente se actúa en consecuencia y si no, puede ser necesario completar el estudio con otras técnicas de imagen adicionales para poder otorgar una categoría BI-RADS definitiva [12]:

- Ecografía mamaria: obtiene imágenes del interior del cuerpo mediante el uso de ondas de sonido de alta energía (ultrasonidos).
- Mamografía de diagnóstico: Es una radiografía más detallada de la mama.
- Imagen por resonancia magnética (IRM): Esta técnica es la que crea imágenes más detalladas de áreas dentro de la mama.

- **Biopsia:** Se extrae una muestra o líquido del tejido de la mama con el objetivo de estudiarlo en el microscopio o hacer más pruebas. Existen distintos tipos de biopsias (aspiración con aguja fina, biopsia con aguja gruesa o biopsia abierta).

Por último, cuando se ha confirmado el estadiaje, los radiólogos llevan a cabo el seguimiento indicado o remiten a las mujeres a un especialista en mama o a un cirujano que procederá a indicarles el tratamiento a seguir.

1.5. Tratamientos del Cáncer de Mama

El tratamiento empleado para erradicar el CM varía dependiendo del estadio de CM, biomarcadores, el estado de salud general de la paciente y sus preferencias [13].

Existen diversos tratamientos habituales y algunos nuevos en ensayos clínicos. La cirugía es la opción a la que se someten la mayoría de las pacientes y existen distintos tipos:

- La cirugía con conservación de la mama, lumpectomía o mastectomía parcial, donde no se extirpa la mama, simplemente el tumor y tejido que lo rodea. A su vez, mediante una incisión separada, durante la operación se pueden extirpar algunos ganglios linfáticos de la axila para su posterior análisis patológico. Esta es la técnica llamada biopsia de ganglio linfático centinela.
- La mastectomía total o mastectomía simple es la operación en la que se extirpa la mama completa que tiene cáncer. Asimismo, también se puede llevar a cabo la biopsia de ganglio centinela.
- La mastectomía radical modificada, donde se extirpa toda la mama incluyendo el pezón y la piel que recubre toda la mama. En estos casos, se suele extirpar también la mayoría de los ganglios linfáticos de la axila.

Otro de los tratamientos más comunes es la quimioterapia [14]. Consiste en el uso de medicamentos destinados a impedir la formación de células cancerosas, mediante su destrucción o interrumpir su proliferación. Este tratamiento suele ser sistémico, se administra por vena o por vía oral, llega al torrente sanguíneo y puede alcanzar cualquier célula cancerosa del cuerpo. En muchas ocasiones la quimioterapia no es el único tratamiento, sino que es complementario a otros tratamientos como la cirugía. Se puede utilizar como:

- Quimioterapia neoadyuvante: Disminuir el tamaño del tumor antes de la cirugía o radioterapia.
- Quimioterapia adyuvante: Eliminar las posibles células cancerosas restantes después la cirugía o radioterapia.
- Contribuir a un mejor resultado de otros tratamientos o destruir células que han migrado a otros tejidos.

En la radioterapia se usan rayos X de alta energía u otros tipos de radiación ionizante mediante los cuales se daña el ADN de las células con el objetivo de eliminarlas o interrumpir su proliferación [15]. Su forma de administración dependerá del cáncer y su estadiaje. Existen dos tipos de radioterapia:

- Radioterapia interna o braquiterapia: la fuente de radiación se coloca dentro del cuerpo en forma de semilla o cápsulas, cerca del cáncer.
- Radioterapia externa: desde el exterior se emiten rayos que atraviesan el cuerpo del paciente hacia la zona donde se encuentra el cáncer. Este es el tipo de radioterapia que se utiliza en el tratamiento de CM.

Estos tres son los principales tratamientos utilizados en el CM. Sin embargo, existen otras técnicas también utilizadas como es la terapia con hormonas, cuyo objetivo es detener el crecimiento de células cancerosas bloqueando o eliminando hormonas que pueden unirse a tumores y hacerlos proliferar. La terapia dirigida es otro tratamiento que utiliza medicamentos, los anticuerpos monoclonales, creados en el laboratorio y administrados por infusión para combatir contra el cáncer. Y, por último, la inmunoterapia encaminada a fortalecer el propio sistema inmune del paciente para luchar contra el tumor. Aparte de todos ellos se siguen probando nuestros tratamientos en ensayos clínicos.

1.6. Ganglios linfáticos axilares y biopsia del ganglio centinella

Como se menciona en la anatomía de la mama, ésta está rodeada de ganglios linfáticos que forman parte del sistema linfático. La función de este sistema es transportar la linfa por todo el organismo a través de los vasos linfáticos. Este líquido contiene glóbulos blancos y

otros elementos que protegen al organismo y combaten contra enfermedades. Por su parte, los ganglios linfáticos son agrupaciones de glóbulos blancos (linfocitos B y linfocitos T) que filtran la linfa capturando así células extrañas o dañadas.

Hay grupos de ganglios por diferentes partes del cuerpo. Una de esas zonas es la axila y al ubicarse muy cerca de la mama puede ser una de las principales vías de diseminación de las células cancerosas de mama.

En concreto, el ganglio centinela (GC) es el primer ganglio que se encuentra más caudal y más cercano del tumor y, por lo tanto, es la primera vía por donde se pueden transportar células cancerosas y crear metástasis. Su estado es representativo de lo que ocurre en el resto de los ganglios axilares y, por ello, su diagnóstico es muy importante para la planificación del tratamiento [16].

Hace unos años, el principal tratamiento para evitar la propagación del CM era mediante una linfadenectomía, es decir, eliminar mediante una cirugía todos los ganglios de la axila de la mama afectada [17]. Sin embargo, este tratamiento, en ocasiones innecesario, tiene asociada una gran morbilidad y complicaciones como linfedema, alteraciones sensibles del brazo, dolor o malestar. Por ello, se optó por la biopsia de ganglio centinela (BGC), con la cual se reduce la morbilidad y las cirugías innecesarias.

La BGC comienza el día de antes de la cirugía mediante la inyección de un radiofármaco, usualmente Tecnecio 99, cercano a la zona del tumor [18]. Éste difundirá por los diferentes vasos hasta alcanzar el GC. Mediante una gammagrafía el experto podrá detectar los ganglios con mayor captación y podrá marcarlos en la piel del paciente para su posterior extracción al día siguiente.

El día de la operación, el cirujano busca el GC guiándose por la marca de la piel y la gammagrafía ya que el radiofármaco acumulado en el ganglio probablemente seguirá emitiendo radiación que puede ser captada. Una vez localizado, se extirpa el GC y es enviado al patólogo que deberá analizarlo [19]. La intervención realizada por el cirujano puede verse ilustrada en la Figura 1.2.

Una vez analizado, si este es negativo, el cáncer no se ha diseminado al resto de ganglios y órganos cercanos. Si la biopsia es positiva, significa que el cáncer ha alcanzado el GC y posiblemente ganglios y órganos cercanos. En este caso los médicos deberán determinar el estadio del cáncer y el tratamiento a realizar. Habitualmente, este procedimiento se realiza a la vez de la extirpación del tumor primario de mama, por ello, si el patólogo tiene los resultados se pueden extirpar otros ganglios en el momento.

Todo este proceso es llevado a cabo para poder conocer el estado del ganglio, determinar el estadio del cáncer y evitar consecuencias mayores.



Figura 1.2: Biopsia del Ganglio Centinela [2].

1.7. Resonancia Magnética en el Cáncer de Mama

La imagen por resonancia magnética (RM) es una técnica de imagen no invasiva, que no utiliza radiación y que nos permite obtener imágenes anatómicas en tres dimensiones.

La RM utiliza campos magnéticos generados por potentes imanes (1.5T - 3T) y ondas de radiofrecuencia para producir las imágenes. Dentro de los tejidos del paciente se encuentran protones que son capaces de girar y alinearse con diferentes campos magnéticos. En su estado normal, los protones se alinean al azar. Cuando el paciente entra dentro de la máquina de RM, los protones se alinean con el campo magnético generado por los imanes de la máquina. Después, se aplica un pulso de radiofrecuencia en el sentido contrario al del campo magnético de la RM por lo que los protones quedarán alineados en contra del campo en un estado de alta energía. Cuando la onda de radiofrecuencia cesa, los protones se relajan y vuelven a alinearse con el campo magnético de RM, liberando la energía acumulada [20].

Las bobinas de la máquina son las encargadas de registrar las diferentes intensidades dependiendo de la velocidad y magnitud con la que llega la energía liberada de los protones. Esta información es procesada por algoritmos computarizados que analizarán las señales recibidas y producirán las imágenes anatómicas.

En el caso de la mama, la RM, junto a la ecografía y la mamografía, son las técnicas por excelencia utilidad para el diagnóstico de CM. La RM es la técnica con mayor sensibilidad ya que podemos observar los tejidos blandos con mayor definición [21]. Es principalmente

utilizada para el estadiaje del cáncer conocido, el diagnóstico precoz en pacientes con riesgo elevado y la evaluación de la respuesta tras la administración de quimioterapia. Además, es útil para reconocer lesiones que pueden encontrarse ocultas desde los planos de otras técnicas de imagen.

Las imágenes por RM de la mama deben tener un grosor de sección no superior a 2,5 mm, ya que deben poder evaluarse los tumores de 5 mm o más grandes. La paciente debe tumbarse boca abajo sobre la camilla y las mamas deben quedar encajadas en los huecos de la camilla para que no sean presionadas.

La RM con contraste es la modalidad más utilizada en el CM. Al introducir un contraste intravenoso (gadolinio), se puede evaluar la permeabilidad de los vasos sanguíneos y la existencia de neoangiogénesis, formación de nuevos vasos, lo cual es un indicativo de cáncer. Se mide la diferencia de realce local entre una imagen tomada sin contraste y la imagen tomada después de la inyección del mismo, viendo que ésta ha tenido un realce más rápido debido al aumento de su permeabilidad sanguínea. Este fue el comienzo de la RM con contraste para el CM, sin embargo, hoy se realizan estudios dinámicos multiparamétricos que nos permiten evaluar la situación más exhaustivamente.

De manera habitual encontramos estudios radiológicos que incluyen las siguientes secuencias de RM:

- Imagen ponderada en T1: la reconstrucción de esta imagen se hace midiendo el tiempo de relajación longitudinal, el tiempo que se requiere para lograr la realineación con el campo magnético de la RM. Nos informa sobre cómo de rápido se mueve el protón y emite la energía acumulada. T1 es la constante de tiempo que determina la velocidad a la que los protones excitados vuelven al equilibrio. Es una medida del tiempo que tardan los protones en realinearse con el campo magnético externo, y se fija como el tiempo que tarda la magnetización en recuperar un 63% de su valor. El tejido adiposo aparece brillante (señal de alta intensidad) en las imágenes ponderadas en T1 y el agua y los líquidos aparecen relativamente oscuros. Las imágenes en T1 muestran más nítidamente la anatomía del tejido blando y la grasa (nos permiten confirmar una masa que tiene grasa) [22].



Figura 1.3: Imagen de RM de la base de datos ponderada en T1.

- Imagen ponderada en T2: la reconstrucción de esta imagen se hace midiendo el decaimiento de la magnetización transversal. En este caso se mide cómo la señal inicial de radiofrecuencia decae con una constante de tiempo T2. T2 es la constante de tiempo que determina la velocidad a la que los protones excitados alcanzan el equilibrio o se desfasan entre sí. Es una medida del tiempo que tardan los protones en perder la coherencia de fase entre los núcleos que giran perpendiculares al campo principal, y se fija como el tiempo que tiene que transcurrir para que la magnetización transversal pierda el 63 % de su valor. En el caso del T2, el tejido adiposo aparece relativamente oscuro (señal de baja intensidad) y el agua y los líquidos aparecen brillantes. Las imágenes en T2 muestran de manera óptima líquido y alteraciones como tumores, inflamación, traumatismo [22]. En la práctica, las imágenes ponderadas en T1 y T2 proporcionan información complementaria, por lo que ambas son importantes para la caracterización de alteraciones.

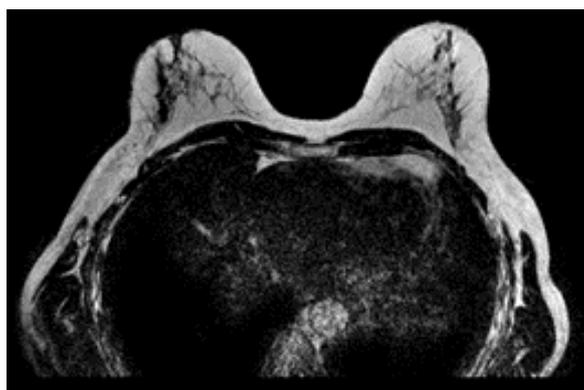


Figura 1.4: Imagen de RM de la base de datos ponderada en T2.

- Estudio dinámico: es la toma de imágenes ponderadas en T1 durante la inyección del contraste. Se toman 7 imágenes del paciente durante los 7 minutos siguientes a la admi-

nistración en las que se puede observar la irrigación sanguínea a lo largo del tiempo.

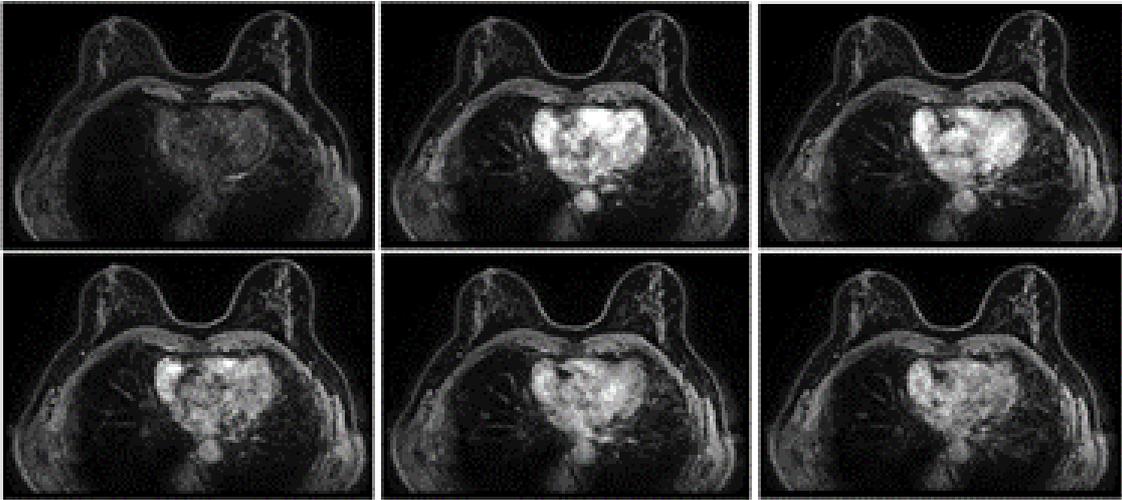


Figura 1.5: Imagen de RM de la base de datos de un estudio dinámico.

- Imagen de difusión: se basa en la detección del movimiento aleatorio de las moléculas de agua en los tejidos, detección que solo es posible a través de la RM debido a que nos permite recoger la traslación de moléculas. En el caso del cáncer, los tumores con importante vascularización y con gran masa celular muestran alta señal ya que suponen una gran restricción a la difusión y movimiento de las moléculas [23].

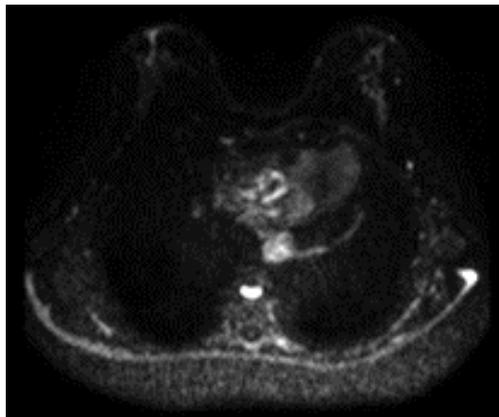


Figura 1.6: Imagen de RM de la base de datos de difusión.

Todas estas diferentes imágenes de RM hacen posible que los radiólogos puedan evaluar exhaustivamente el CM y examinar los ganglios axilares para poder determinar también la existencia de metástasis de una manera no invasiva, evitando intervenciones innecesarias y pudiendo determinar su tratamiento más rápidamente.

Capítulo 2

Marco Teórico y Estado del Arte

2.1. Machine Learning

La Inteligencia Artificial (IA) es quizás la mayor revolución tecnológica del siglo XXI. Hoy en día, no se puede entender prácticamente ningún desarrollo tecnológico sin el uso o influencia de la IA. Se puede encontrar desde la cafetera que nos prepara un café hasta un robot que realiza una cirugía compleja. Es difícil encontrar una definición exacta de IA pero una buena aproximación sería “conjunto de técnicas informáticas que permiten a una máquina (por ejemplo, un ordenador, un teléfono) realizar tareas que, por lo común, requieren inteligencia tales como el razonamiento o el aprendizaje” [24].

En concreto, una rama de la IA es el Machine Learning (ML) o aprendizaje automático. Consiste en algoritmos capaces de encontrar patrones recurrentes en conjuntos de datos. Esos datos pueden ser números, palabras, imágenes, estadísticas, etc.

Al detectar patrones en esos datos, los algoritmos de ML aprenden de forma autónoma a realizar una tarea o hacer predicciones a partir de datos y mejorar su rendimiento. Una vez entrenado, el algoritmo podrá encontrar los patrones en nuevos datos [25]. En concreto, el desarrollo de un modelo de ML debe ser supervisado por la figura del científico de datos y sigue cinco etapas que se enumeran a continuación y que pueden verse ilustradas en la Figura 2.1.

1. Seleccionar y preparar la base de datos que se utilizará para entrenar el modelo y que aprenda a resolver el problema concreto que se quiere solucionar.
2. Seleccionar un algoritmo para ejecutar sobre el conjunto de datos de entrenamiento. El tipo de algoritmo que se emplea depende del tipo y del volumen de datos de entrenamiento

y del tipo de problema que haya que resolver.

3. Entrenar el algoritmo. Es un proceso de repetición. Las variables se ejecutan a través del algoritmo y los resultados se comparan con los que debería haber producido. Los «pesos» y el sesgo se pueden ajustar para aumentar la precisión del resultado. El algoritmo entrenado es el modelo de ML.
4. Realizar pruebas del modelo entrenado para validar su rendimiento.
5. Uso y mejora del modelo. Se utiliza el modelo sobre nuevos datos, cuyo origen depende del problema que haya que resolver. Por ejemplo, en los correos electrónicos se usará un modelo de ML diseñado para detectar spam.

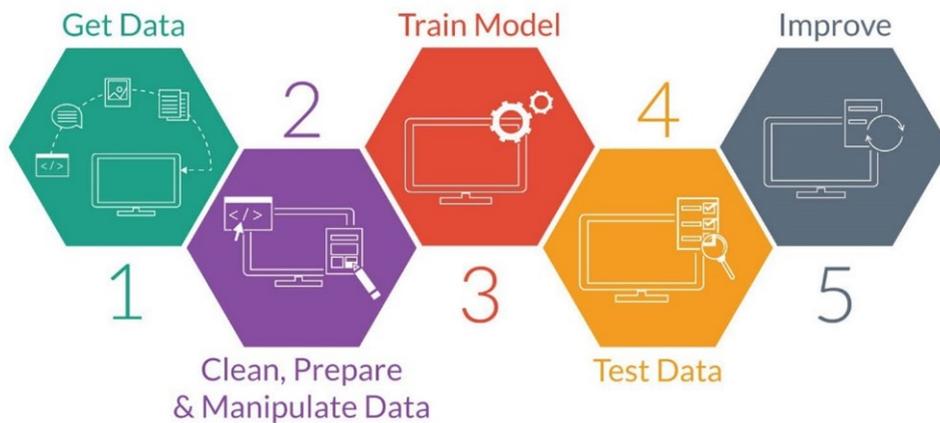


Figura 2.1: Esquema representativo de las etapas del proceso de desarrollo de un modelo ML [3].

2.2. Radiómica

Hoy en día, gracias a las diferentes herramientas de procesamiento de señales y datos, tenemos imágenes médicas con mejor resolución, mejor contraste y menos artefactos que nos permiten realizar un mejor análisis de los parámetros y un diagnóstico más preciso.

Unido a estos avances, aparece una ciencia que nos permite hacer un análisis cuantitativo de las imágenes médicas de manera no invasiva, la radiómica. La radiómica, mediante algoritmos computacionales, es capaz de analizar numerosas características inapreciables al ojo humano (las conocidas como características radiómicas) de cualquier imagen médica existente (RM, TC o PET) [26].

Las características que están relacionadas con un fenotipo o estado fisiológico característico son denominadas biomarcadores radiómicos. Éstos informan sobre la predisposición o presencia de enfermedades y, por lo tanto, pueden servir de apoyo a la toma de decisiones de los clínicos.

Esta nueva ciencia nos presenta muchas ventajas como es poder analizar imágenes del antes y después del tratamiento o poder analizar zonas de tejido menos delimitadas a diferencia de otras técnicas como la biopsia. Además, analiza una gran cantidad de datos que permiten profundizar en el conocimiento sobre las enfermedades o diferentes estados fisiológicos.

2.3. Estado del Arte

Como mencionamos anteriormente, el ML esta siendo utilizado en todas las áreas, pero especialmente en áreas de investigación como puede ser el análisis de imágenes médicas. En el caso del diagnóstico de CM, se están realizando numerosos estudios para poder realizar un diagnóstico precoz que permita aumentar hasta un 95 % la supervivencia [27]. La ecografía y mamografía son técnicas de imagen muy utilizadas, sin embargo, carecen de la resolución, sensibilidad y especificidad que sí tiene la RM. Es por ello que la mayoría de las investigaciones de análisis radiómico se hacen en RM ya que obtienen una mayor precisión.

El estudio de Vishwa S. Parekh y Michael A. Jacobs [28] trata de realizar una asociación entre las características radiómicas en RM y los diferentes tejidos mamarios, generando unos mapas de características radiómicas para la evaluación de pacientes. Este método con una sensibilidad y especificidad del 93 % y 85 %, permite a los clínicos diagnosticar basándose en tablas sin emplear ML. El estudio de Matheus de Freitas Oliveira Baffa y Aura Conci [29] tiene como objetivo ayudar a la detección del CM en su fase más inicial evitando falsos negativos. Este estudio trata de clasificar mediante el protocolo *K-Fold Cross-Validation* imágenes de infrarrojos de la mama representadas por un vector de características radiómicas. Se obtuvo una buena precisión de 98,61 %.

También hay varios estudios radiómicos del CM que evalúan las regiones tumorales para predecir la respuesta completa a la quimioterapia neoadyuvante. El estudio de Nathaniel M. Braman et al [30], utilizando un clasificador de análisis discriminante lineal diagonal (DLDA) es capaz de predecir la respuesta tumoral y determinar las características más útiles según los diferentes receptores.

En cuanto a la investigación de la afectación ganglionar existen menos estudios. Sin embar-

go, determinar el estado de los ganglios linfáticos axilares sigue siendo un requisito obligatorio del proceso de diagnóstico, muy importante para conocer el estadio del paciente y poder impedir su expansión a todo el organismo.

El estudio en 2017 de Dong et al. [31] consiste en un modelo radiómico capaz de predecir la metástasis del GC en pacientes con CM basada en las secuencias de RM en T2 y por difusión DWI. Concluyó que las características radiómicas obtenidas podían ayudar al diagnóstico de metástasis ganglionar.

El estudio de Richard Ha et al. [32] tiene como objetivo clasificar los ganglios linfáticos en imagen RM utilizando una red neuronal convolucional (CNN). Se seleccionaron una serie de ganglios linfáticos axilares metastásicos y ganglios linfáticos de control negativos en T1 con contraste y se segmentaron gracias al software 3D Slicer. Se diseñó una CNN para la predicción de ganglios linfáticos basada en cada una de estas imágenes segmentadas. Esta CNN realiza una clasificación de ganglios con un 84,3% de precisión, concluyendo así que es factible crear una CNN para predecir la probabilidad de metástasis ganglionar que mejoraría teniendo una base de datos más extensa.

El presente trabajo se podría entender como continuación al trabajo de fin de grado “Investigación en la predicción de afectación de ganglios centinela mediante aprendizaje maquina basado en características radiómicas de imágenes de resonancia magnética” presentado por Alberto Martínez Caballero en la Universidad Rey Juan Carlos. En él se realiza una recolección de imágenes de RM en T1 de pacientes con metástasis axilar. Estos pacientes han sido diagnosticados previamente mediante imagen y biopsia por los radiólogos del Hospital Universitario Ramón y Cajal (HURyC). Posteriormente se extrajeron las características radiómicas de las RM para entrenar un modelo basado en *Random Forest*. Se obtuvo una muy buena precisión (0.90), sensibilidad (0.88) y especificidad (0.93).

En este TFG se plantea una situación similar, pero con una diferencia básica: los pacientes seleccionados para este TFG tienen un diagnóstico positivo en el análisis mediante BGC, pero en el estudio radiológico no se detectaron como tal.

La información sobre todos estos estudios analizados nos lleva a afirmar que la radiómica tiene un gran potencial para predecir la afectación ganglionar, clasificar los tipos de CM, predecir la respuesta al tratamiento y su pronóstico.

Capítulo 3

Motivación y Objetivos

3.1. Hipótesis

La hipótesis de este trabajo de fin de grado de Ingeniería biomédica es que es posible **identificar falsos negativos en ganglios centinela metastásicos confirmados por la técnica de ganglio centinela usando la radiómica en las imágenes de RM.**

El objetivo principal es tratar de identificar y clasificar aquellos ganglios centinela que los radiólogos no han sido capaces de identificar como malignos debido a su morfología aparentemente normal en las imágenes de RM y que luego mediante la BGC han sido diagnosticados como malignos.

Un objetivo secundario del trabajo de fin de grado es realizar la clasificación automática de los ganglios de las pacientes con cáncer lobulillar (CL) y cáncer ductal (CD) por separado para poder comparar la clasificación de cada cáncer y estudiar la importancia de las características radiómicas para cada uno de ellos.

3.2. Motivación

El presente estudio ha sido propuesto y supervisado desde el Servicio de Radiología del Hospital Universitario Ramón y Cajal (HURyC), en concreto por la sección de Radiología mamaria. La idea del proyecto surge debido a los numerosos casos de mujeres cuyos ganglios axilares en imagen no han sido reconocidos como malignos y se pensaba que no había riesgo metastásico, pero que después de una BGC han resultado ser positivos. La BGC conlleva la realización de

una cirugía que retrasa el proceso diagnóstico produciendo una demora en el tratamiento de la paciente y empeoramiento del pronóstico. La intervención puede acarrear complicaciones postquirúrgicas como son infecciones, linfedema o falta de movilidad. Así mismo, la introducción de un fármaco radiactivo para su localización puede conllevar problemas de rechazo.

Por todo ello, la principal motivación de este estudio es poder aumentar la esperanza de vida de los pacientes mediante un diagnóstico cada vez más precoz, realizando el menor número de intervenciones posible. También está motivado por la tendencia actual a utilizar métodos más conservadores y menos invasivos para el paciente, tanto en el diagnóstico como durante el tratamiento. Además, esta investigación sobre ganglios linfáticos se puede extrapolar al diagnóstico de otras enfermedades, no solo de CM sino también linfomas o leucemia.

3.3. Planificación y objetivos específicos

A continuación se muestran las diversas tareas u objetivos secundarios a cumplimentar para llegar a nuestro principal objetivo que es la identificación radiómica de falsos negativos en ganglios centinela metastásicos.

- Revisión del Estado del Arte.
- Selección de mujeres con CM con ganglios axilares negativos en pruebas de imagen y BGC positiva.
- Selección, anonimización y descarga de las imágenes de RM en las modalidades T1 y dinámico con contraste.
- Segmentación de ganglios patológicos confirmados mediante BGC y ganglios control sanos.
- Extracción de las características radiómicas y selección de las características más influyentes.
- Entrenamiento de un modelo de ML para la clasificación de los ganglios.
- Evaluación del rendimiento del modelo y de la importancia de las características radiómicas.
- Escritura de la memoria.

Se puede observar un diagrama de Gantt en la Figura 3.1. En éste se representan cronológicamente cada uno de los pasos realizados y mencionados anteriormente.

DIAGRAMA DE GANTT



Figura 3.1: Diagrama de Gantt.

Capítulo 4

Materiales

4.1. Creación de la base de datos

Para este proyecto el criterio de inclusión fue el de pacientes con carcinoma lobulillar (CL) y carcinoma ductal (CD) que habían sido registrados con biopsia de adenopatía axilar positiva y que no tenían ganglios metastásicos identificables mediante pruebas de imagen.

El servicio de anatomía patológica del HURyC facilitó una base de datos con 30 pacientes con CL y 144 casos con CD que cumplían con el criterio de inclusión desde enero de 2020 a marzo de 2023. Se ha optado por crear una base de datos balanceada con el mismo número de pacientes con CL y CD para poder luego comparar las características radiómicas más relevantes de cada uno de los dos carcinomas. Por ello, los pacientes con CL limitan la N de pacientes de cada tipo y se comienza el estudio con 30 pacientes con CL y 30 pacientes con CD, en total 60 pacientes. De estos pacientes se seleccionan los GC positivos y los ganglios axilares sanos del lado contralateral que serán los ganglios control, lo que hacen un total de 120 ganglios.

Estos ganglios serán evaluados en imágenes de RM realizadas en HURyC, mediante sus equipos de RM de 1.5 T de Philips Medical Systems. Las imágenes deben pertenecer a un estudio dinámico con contraste, ya que se van a evaluar dos tipos de imágenes, las imágenes ponderadas en T1 (ver en la Figura 1.3) e imágenes dinámicas de 7 minutos en las que se estudiará el minuto 1 (ver en la Figura 1.5). Al tener que segmentar los ganglios manualmente en cada modalidad, las imágenes ponderadas en T2 y las imágenes de difusión quedan descartadas, y se opta por incluir únicamente los dos tipos de imagen ya mencionados ya que son las más utilizadas en el diagnóstico.

A continuación se muestra la Tabla 4.1 con los datos de cada secuencia:

Datos	Imágenes ponderadas en T1	Imágenes de estudios dinámicos
TR (tiempo de repetición)	604.647 ms	5.697 ms
TE (tiempo de eco)	5.318 ms	2.806 ms
Dimensiones de la imagen	448 x 448 x 70	448 x 448 x 178
Tamaño del vóxel	0.8116mm x 0.8116mm	0.8116mm x 0.8116mm
Grosor de corte	0.75 mm	0.8013 mm

Tabla 4.1: Datos de las imágenes de RM utilizadas para la clasificación

Ésta es la base de datos con la que se inicializa el proyecto. Sin embargo, debido a las características que requiere el proyecto, como es la inclusión de la zona axilar en las imágenes o la existencia de estudios con contraste, la base de datos se ve reducida durante el proceso de selección de imágenes y segmentación, descritos en el Capítulo 5.

4.2. Herramientas utilizadas

4.2.1. 3D Slicer

3D Slicer es un software gratuito de código abierto que permite visualizar, procesar, segmentar y analizar las imágenes médicas y cualquier tipo de imagen 3D [33] (ver en la Figura 4.1).

Se utiliza este software multifuncional, concretamente su versión Slicer 5.2.2, principalmente para visualizar los ganglios, segmentarlos manualmente mediante el módulo *Segmentations* y posteriormente exportar imagen y máscara en el formato adecuado para que pueda ser leído y procesado por el siguiente software.

4.2.2. Python

Python es un lenguaje de programación orientado a objetos, potente y fácil de aprender. Se utiliza a través de Anaconda, una de las mejores distribuciones de Python para el desarrollo científico [34]. Python tiene múltiples librerías. Éstas son conjuntos de funcionalidades que permiten al usuario realizar diferentes tareas. En este estudio las librerías más utilizadas son:

- Pandas: útil para la lectura y el manejo de Dataframes.

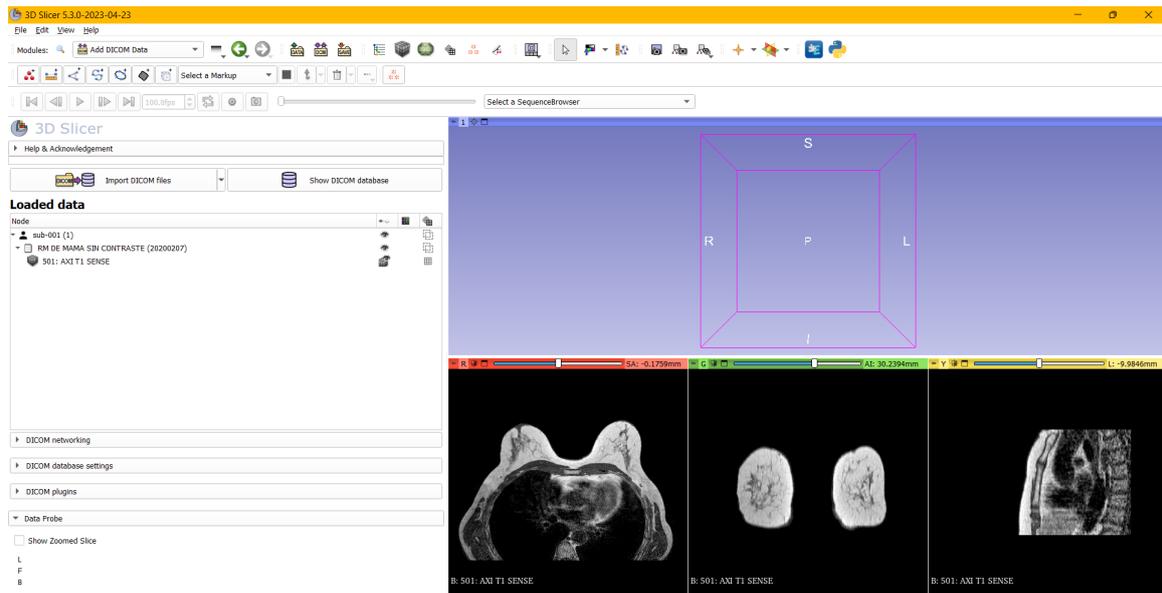


Figura 4.1: Página principal del software 3D Slicer con la visualización de un paciente. En la barra superior se pueden observar todas las herramientas disponibles.

- Matplotlib: puede generar una gran variedad de gráficos de calidad, como histogramas o diagramas de barras.
- Scipy y Numpy: permiten tener una estructura de datos numéricos universal y su optimización y tratamiento.
- Pyradiomics: aceptado recientemente en la investigación del cáncer, es un paquete de Python de libre acceso que permite la extracción de características radiómicas de imágenes médicas.
- Scikit-learn: permite la construcción de modelos de aprendizaje automático, así como el análisis de datos.
- SHAP: se utiliza para el cálculo de la importancia de las características en los modelos ML.

Capítulo 5

Métodos

5.1. Selección y extracción de Imágenes

Como bien se mencionó en la sección 4.1, se han seleccionado 30 pacientes con CL y 30 pacientes con CD cuyos ganglios axilares en imagen no han sido reconocidos como malignos pero que después de una BGC han resultado ser positivos. Por lo tanto, partimos con 60 pacientes.

Desde los equipos del HURyC, se procede a la selección de esos pacientes a través del IntelliSpace Portal de Philips que, mediante el número de historia clínica del paciente, permite buscar todos los estudios radiológicos del paciente registrados por el Sistema de archivo y comunicación de imágenes (PACS) y descargarlos en formato DICOM para su posterior análisis.

Antes de la descarga, se deben anonimizar los pacientes para poder tratar con la información durante la investigación. Para ello, se elimina cualquier dato personal y se sustituye por "sub_XXX", asignándole un número a cada paciente.

Durante la búsqueda y anonimización de cada paciente, hay que fijarse que el paciente tenga estudios dinámicos de RM de mama antes de una BGC, es decir, que sus imágenes hayan sido tomadas con contraste, y que estos estudios contengan imágenes ponderadas en T1 e imágenes de estudio dinámico de 7 minutos. Además, existen estudios de pacientes realizados en otros centros, cuyas imágenes han sido digitalizadas y traspasadas al PACS del HURyC, esas imágenes están en un formato distinto a las imágenes tomadas por los equipos del HURyC y serían descartadas. Asimismo, durante la visualización de las imágenes puede haber imágenes movidas o que no incluyan la zona axilar donde hay que segmentar los ganglios.

Durante este proceso de los primeros 30 pacientes con CL, la base de datos se vió reducida a 13 pacientes con CL:

- 4 pacientes no tenían estudios dinámicos de RM de mama.
- 2 pacientes tenían estudios dinámicos de RM de mama postquirúrgicos.
- 4 pacientes tenían estudios digitalizados en un formato diferente al del hospital.
- 4 pacientes no incluían la zona axilar.
- 3 pacientes tenían imágenes en movimiento.

Aunque no es necesario tener un conjunto de datos balanceado, se quería tener un número similar de pacientes con CL y CD para poder comparar los resultados. Por tanto, la base de datos final de imágenes ponderadas en T1 es: 13 pacientes con CL y 13 pacientes con CD, en total 26 pacientes, de los cuales se analiza 2 ganglios de cada uno, un total de 52 ganglios.

A continuación, se selecciona el minuto 1 de las imágenes de estudio dinámico de 7 minutos de cada paciente. Algunos de estos pacientes no tienen el registro de los 7 minutos. Por lo tanto, la base de datos final de imágenes del minuto 1 del estudio dinámico es: 8 pacientes con CL y 7 pacientes con CD, en total 15 pacientes, de los cuales se analiza 2 ganglios de cada uno, un total de 30 ganglios.

5.2. Segmentación de Ganglios Axilares

Para el proceso de segmentación se hará uso de 3D Slicer (mencionado en la sección 4.2.1). Éste nos permite realizar una segmentación manual. Primero se deben cargar todas las imágenes DICOM anonimizadas en el software (ver Figura 5.1).

DICOM database

Patient name	Patient ID	Birth date	Sex	Studies	Last study date	Date added
sub-012	12	1974-01-01	F	1	2021-12-10	2023-05-...0:05.426
sub-011	11	1954-01-01	F	1	2021-10-08	2023-05-...9:53.806
sub-010	10	1960-01-01	F	1	2021-02-19	2023-05-...9:41.237
sub-009	9	1965-01-01	F	1	2021-08-06	2023-05-...6:22.831
sub-007	7	1943-01-01	F	1	2021-05-20	2023-05-...5:57.100
sub-006	6	1969-01-01	F	1	2021-05-07	2023-05-...5:45.306
sub-005	5	1968-01-01	F	1	2020-11-06	2023-05-...2:26.189
sub-004	4	1965-01-01	F	1	2022-06-27	2023-05-...2:16.031
sub-003	3	1942-01-01	F	1	2020-10-23	2023-05-...2:03.154
sub-001	1	1948-01-01	F	1	2020-02-07	2023-05-...1:32.255

Figura 5.1: Bases de datos DICOM cargada en el software 3D Slicer.

A continuación, se selecciona un paciente y una imagen concreta para visualizar y posteriormente segmentar. Para la visualización, se pueden ver los diferentes cortes del estudio gracias

a la barra de desplazamiento, moviéndose hacia la derecha para avanzar en sentido craneal y hacia la izquierda en sentido caudal. Gracias a la herramienta *Draw* del módulo *Segment Editor*, se puede dibujar manualmente el contorno de los ganglios y crear su segmentación. En la Figura 5.2 podemos ver a la derecha un ejemplo de imagen ponderada en T1 del paciente sub_001 y a la izquierda las diferentes herramientas ofrecidas por *Segment Editor* para la segmentación.

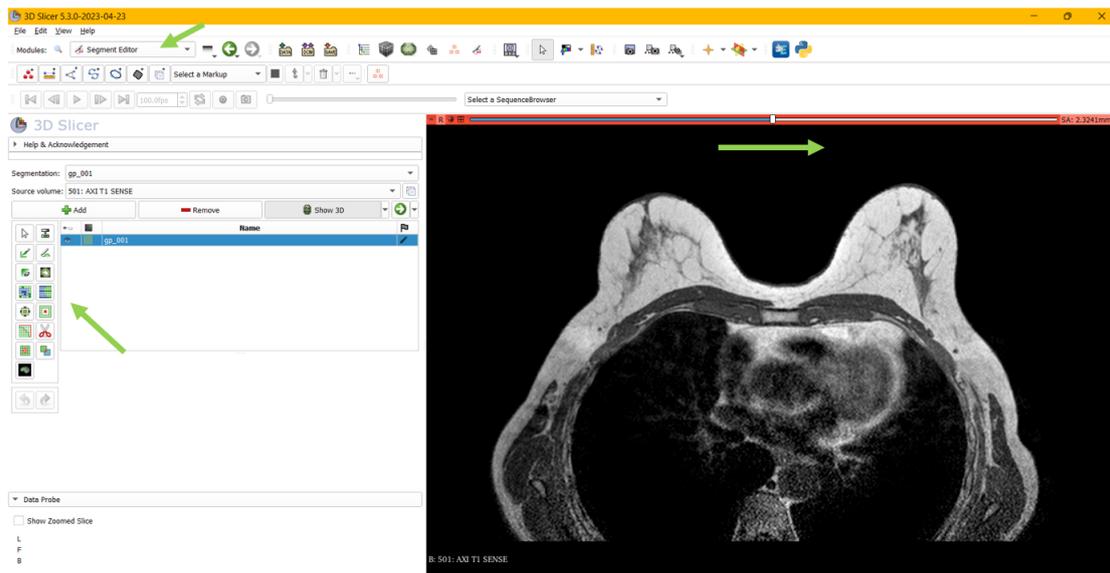


Figura 5.2: Visualización en 3D Slicer de una imagen ponderada en T1 a la derecha y la barra de herramientas de *Segment Editor* a la izquierda.

Antes de comenzar a segmentar, hay que localizar los ganglios. En la base de datos proporcionada por el HURyC, viene la información de qué lado, derecho o izquierdo, es el patológico. En el caso del lado patológico, es muy importante encontrar el GC que ha sido biopsiado y, por lo tanto, es maligno. El GC es el ganglio más caudal y cercano a la mama, de modo que para localizarlo se debe comenzar a analizar el estudio desde la imagen más caudal hasta la imagen más craneal (de izquierda a derecha, como se indica en la Figura 5.2). El ganglio que primero se localice será el ganglio patológico que se segmentará. En el caso de haber varios a la misma altura, se segmenta el que tenga una forma más patológica, bien por ser más grande, más redondo o con forma más irregular. En el caso del lado control, donde se necesita segmentar un ganglio sano, se puede elegir cualquier ganglio ya que todos ellos son benignos.

Una vez localizados, se puede proceder a la segmentación contorneando los ganglios en todos los cortes en los que aparecen. La suma de todos los cortes segmentados conforma una figura tridimensional del ganglio. Se pueden observar varios ejemplos del proceso de segmentación en las Figuras 5.3 y 5.4.

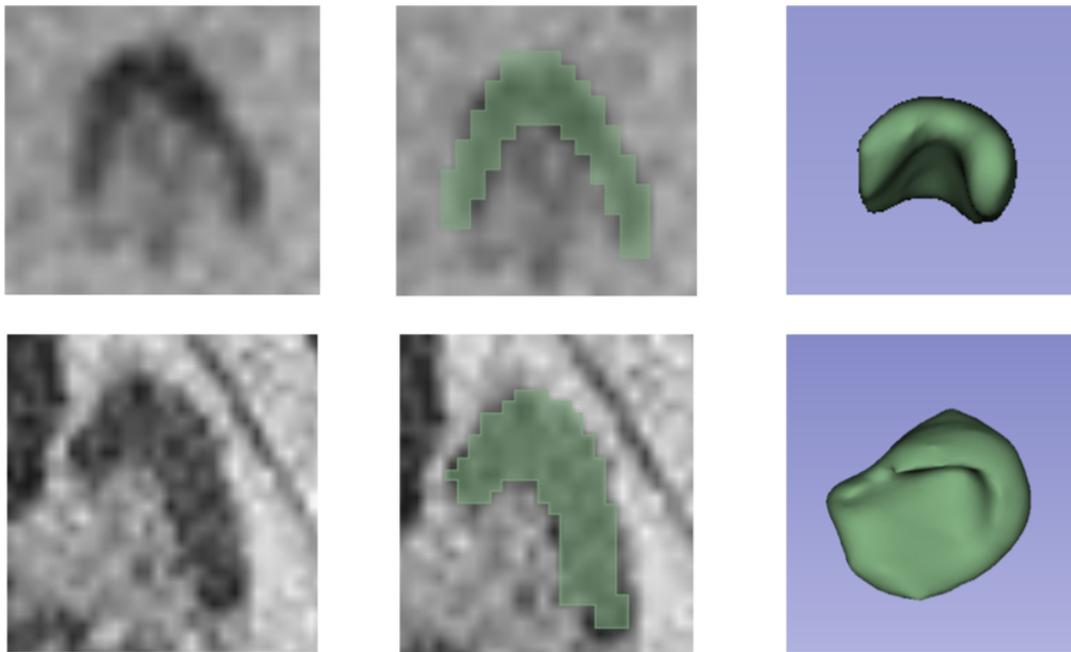


Figura 5.3: Ejemplo de la segmentación de 2 ganglios en imágenes ponderadas en T1. El superior es un ganglio control y el inferior un ganglio patológico.

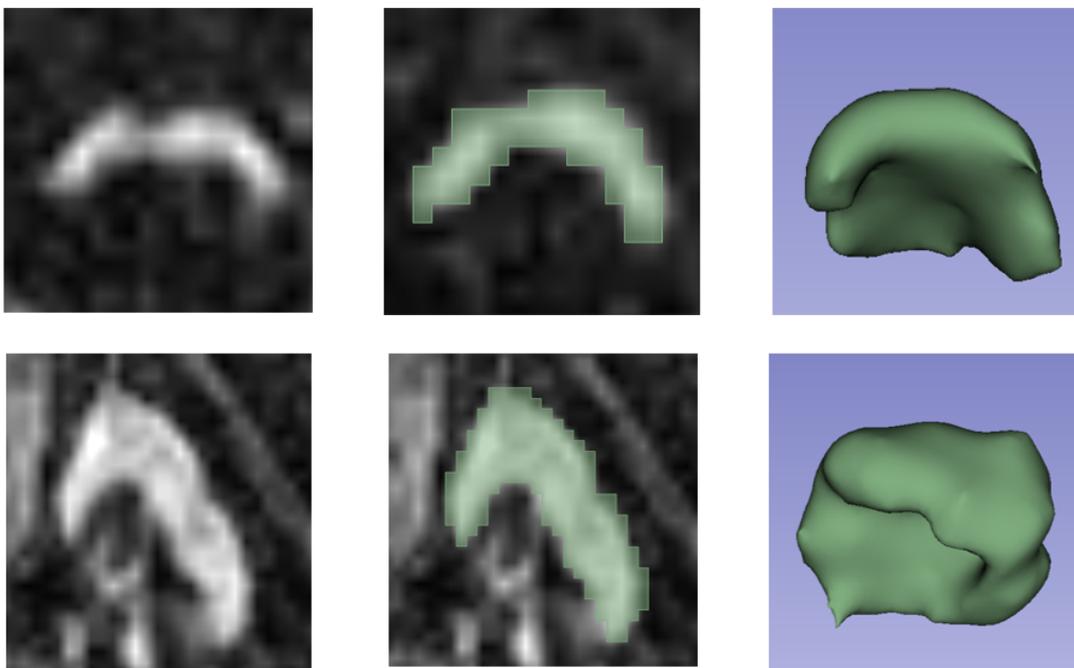


Figura 5.4: Ejemplo de la segmentación de 2 ganglios en imágenes del minuto 1 del estudio dinámico. El superior es un ganglio control y el inferior un ganglio patológico.

Para el proceso de localización de ganglios y segmentación de los mismos ha sido imprescindible la ayuda de la Dra. Irene Vicente Zapata y la Dra. Teresa Presa Abós que enseñaron a la autora a localizar los ganglios y cómo se debían segmentar, así como su supervisión después de la segmentación.

En total, se segmentan 52 ganglios en imágenes ponderadas en T1 y 30 ganglios en imágenes del minuto 1 de estudio dinámico.

De esta manera, una vez finalizado el proceso de segmentación, los archivos se guardan en formato NIfTI para posteriormente ser analizados. Finalmente, se obtienen los siguientes archivos que serán los utilizados en los siguientes métodos:

- Imágenes ponderadas en T1:
 - gp_001_mask.nii.gz (máscara que contiene el volumen del ganglio patológico)
 - gc_001_mask.nii.gz (máscara que contiene el volumen del ganglio control)
 - sub_001.nii.gz (imagen original ponderada en T1)
- Imágenes del minuto 1 del estudio dinámico:
 - gp_001_dyn_mask.nii.gz (máscara que contiene el volumen del ganglio patológico)
 - gc_001_dyn_mask.nii.gz (máscara que contiene el volumen del ganglio control)
 - sub_001_dyn.nii.gz (imagen original del minuto 1 del estudio dinámico)

5.3. Extracción de Características Radiómicas

Para la extracción de características, primero se carga la base de datos que contiene todas las imágenes segmentadas en formato NIfTI en un notebook de Python. Para ello, se utiliza la librería Pandas que nos permite manejar dataframes fácilmente.

A continuación, se utiliza *PyRadiomics*, que permite extraer de manera automática diferentes características de la imagen original o de la imagen obtenida a partir de la aplicación de diferentes filtros. Para determinar qué características y qué tipo de imagen se quiere analizar, *PyRadiomics* permite incluir un archivo, '*params.yaml*', donde se indican este tipo de factores. En el caso de este trabajo, se extraen todas las características de todos los tipos de imágenes posibles (ver en el Anexo A).

Las características que se van a extraer forman parte de estos 8 grupos [35]:

- **First Order Statistics:** Este grupo de características evalúan la distribución de las intensidades de los vóxeles dentro de la región de la imagen delimitada por la máscara a través de métricas básicas como puede ser el cálculo de la energía o la entropía.
- **Shape-based (3D):** Estos estadísticos incluyen descriptores del tamaño y forma tridimensional de la región de interés (ROI por sus siglas en inglés) como puede ser la esfericidad. No tienen en cuenta la intensidad del nivel de gris en el ROI y, por lo tanto, solo se calculan en la imagen y la máscara sin filtros.
- **Shape-based (2D):** En este grupo de características se incluyen descriptores del tamaño y forma bidimensional del ROI como puede ser el área. Al igual que el grupo anterior, solo se calculan en la imagen y la máscara sin filtros.
- **Gray Level Co-occurrence Matrix:** Una Matriz de Co-ocurrencia de Nivel de Gris describe la función de probabilidad conjunta de segundo orden del ROI.
- **Gray Level Run Length Matrix:** Una matriz de longitud de ejecución de nivel de gris cuantifica la longitud en número de píxeles, de píxeles seguidos que tienen el mismo nivel de gris.
- **Gray Level Size Zone Matrix:** Zona de tamaño de nivel de gris cuantifica las zonas de nivel de gris en una imagen. Es decir, el número de vóxeles unidos que tienen la misma intensidad de nivel de gris.
- **Neighbouring Gray Tone Difference Matrix:** Una matriz de diferencia de tonos de gris vecina evalúa la diferencia entre un valor de gris y el valor de gris promedio de sus vecinos dentro de la distancia δ .
- **Gray Level Dependence Matrix:** Una matriz de dependencia del nivel de gris mide las dependencias del nivel de gris en una imagen. En otras palabras, el número de vóxeles conectados dentro de la distancia δ dependientes del vóxel central.

Los tipos de imágenes posibles de los que se van a extraer características son [36]:

- **Original:** La imagen cargada en Python sin filtros.
- **Wavelet:** : Filtrado de wavelet, produce 8 descomposiciones por nivel, todas las combinaciones posibles de aplicar un filtro de paso alto o bajo.
- **LoG:** Filtro laplaciano del gaussiano, realza los bordes y áreas de cambio de nivel de gris.

- Square: Toma el cuadrado de las intensidades de la imagen, manteniendo los valores negativos, y las escala linealmente al rango original.
- SquareRoot: Realiza la raíz cuadrada de las intensidades de imagen absolutas manteniendo los valores negativos, y las reduce al rango original.
- Logarithm: Efectúa el logaritmo de la intensidad absoluta y escala los valores al rango original.
- Exponential: Toma la exponencial, donde la intensidad filtrada es exponencial de la intensidad absoluta.
- Gradient: Devuelve la magnitud del gradiente local.
- LBP2D: Calcula el patrón binario local en una operación por segmento (2D).
- LBP3D: Calcula el patrón binario local en 3D utilizando armónicos esféricos.

Después de la extracción de un total de 1688 características para 52 ganglios en imágenes ponderadas en T1 y 1688 características para 30 ganglios en imágenes del minuto 1 de estudio dinámico, éstas son almacenadas en un documento tipo Dataframe. En éste, los ganglios serán cada una de las filas y las columnas serán cada una de las características.

5.4. Selección de Características Radiómicas

Antes de introducir las características radiómicas a un modelo que clasifique los ganglios, se debe de hacer una selección con el objetivo de reducir la dimensionalidad del problema. La selección de características tiene muchas ventajas, entre ellas [37]:

- Evita el sobreajuste u *overfitting*, este es el principal objetivo de la selección de características. Cuando un modelo tiene demasiados datos de entrenamiento no categoriza bien los datos debido a que aprende del ruido y de demasiadas características inexactas.
- Interpretabilidad, es más fácil explicar la importancia de las características siendo menos. Además, elimina variables irrelevantes que no aporta información a la clasificación haciéndolo más sencillo.
- Reducción de costos computacionales, al ser menor información que procesar.

Para ello, en nuestro modelo se utiliza un método basado en información mutua. La información mutua permite conocer la dependencia estadística entre dos variables. Mide la información que comparten dos características y , por lo tanto, su redundancia. Para ello, se utilizó `SelectPercentile` del módulo de `feature_selection` de `Scikit-Learn` [38]. Indicando un porcentaje, permite seleccionar las características más útiles para la clasificación según su información mutua. En este caso, se seleccionará el 6% de las características ya que es el porcentaje que mejor resultados proporciona al modelo.

5.5. Creación de cuatro conjuntos de datos

En este proyecto, hay diferentes objetivos a desarrollar. Por ello, se crean 4 conjuntos de datos para entrenar un modelo de ML y obtener 4 clasificadores diferentes de los que extraer conclusiones. Estas distintas 4 bases de datos se crean seleccionando los pacientes y las imágenes del conjunto de 52 ganglios en imágenes ponderadas en T1 y 30 ganglios en imágenes dinámicas.

El primer objetivo es la creación de un clasificador de ganglios patológicos (1) o sanos (0). Como se tiene dos tipos de imágenes, se hace un primer conjunto de datos solamente con imágenes ponderadas en T1 (son las más utilizadas en el diagnóstico) y un segundo conjunto de datos con imágenes ponderadas en T1 e imágenes dinámicas, para así comprobar la información que aportan cada una de las imágenes a la clasificación.

Por otro lado, el segundo objetivo es realizar dos clasificaciones por separado de pacientes con CL y CD, para así ver si existen diferencias al realizar por separado la clasificación según el tipo de cáncer. Estos clasificadores categorizarán los ganglios patológicos (1) o sanos (0) para cada tipo de cáncer, considerando imágenes ponderadas en T1 e imágenes dinámicas. Así pues, se crean dos conjuntos de datos con un número de pacientes similar para así poder comparar su rendimiento, una con los ganglios de 8 pacientes con CL y otra con 7 pacientes con CD.

Tras la selección del 6% de las características, éstas son las 4 bases de datos resultantes:

- **Base de datos 1:** 119 características radiómicas de imágenes ponderadas en T1 de 52 ganglios. (52, 119)
- **Base de datos 2:** 203 características radiómicas de imágenes ponderadas en T1 e imágenes del minuto 1 del estudio dinámico de 30 ganglios. (30, 203)

- **Base de datos 3:** 203 características radiómicas de imágenes ponderadas en T1 e imágenes del minuto 1 del estudio dinámico de 16 ganglios de pacientes con CL. (16, 203)
- **Base de datos 4:** 203 características radiómicas de imágenes ponderadas en T1 e imágenes del minuto 1 del estudio dinámico de 14 ganglios de pacientes con CD. (14, 203)

5.6. Clasificación mediante Machine Learning

Para la realización de las clasificaciones, se utiliza un modelo de aprendizaje supervisado, concretamente de tipo *embesed*, denominado *Random Forest* (RF). RF es un algoritmo que combina diferentes árboles de decisión para llegar a un resultado, y puede utilizarse para problemas de clasificación o de regresión. En este caso, se quieren clasificar ganglios sanos (con etiqueta 0) y patológicos (con etiqueta 1), una clasificación binaria.

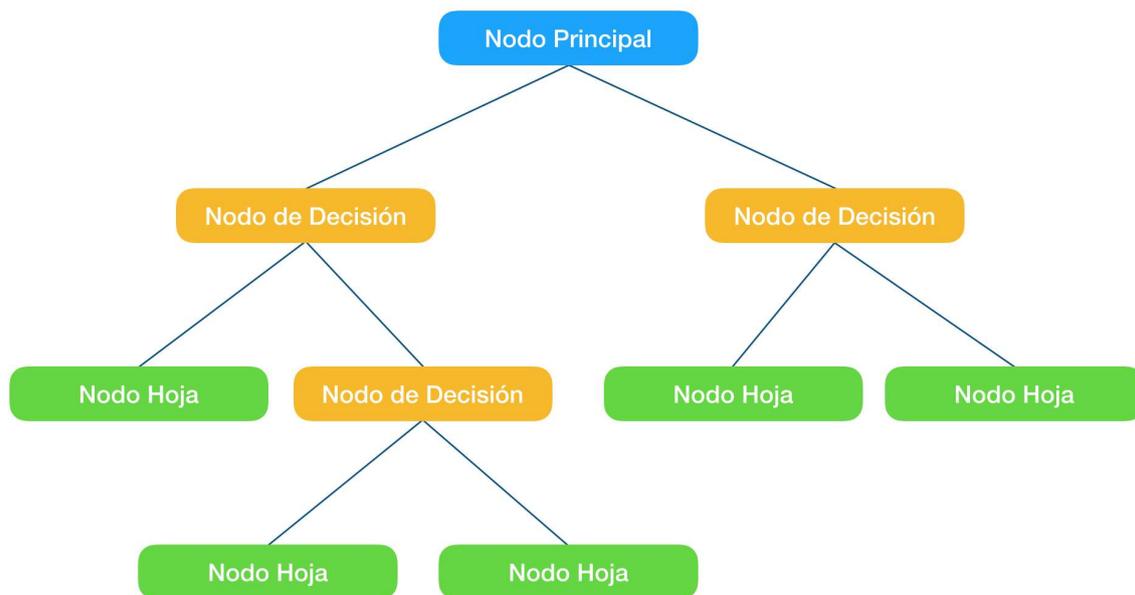


Figura 5.5: Ejemplo visual de un árbol de decisión con nodo principal, nodos de decisión y nodos hoja [4].

Para comprenderlo, se debe entender cómo funciona un árbol de decisión. Un árbol de decisión, de arriba a abajo, está formado por un nodo raíz o principal (donde se introduce la base de datos con las muestras o características), ramas, nodos internos o de decisión y un nodo hoja (donde se da la predicción alcanzada) (ver Figura 5.5). En cada uno de los nodos de decisión,

la base de datos se debe evaluar en función de las características disponibles y formar subconjuntos homogéneos, es decir, dividir entre ganglios sanos (0) y ganglios patológicos (1). Estos subconjuntos se dividirán en dos ramas y pasaran a otros nodos internos que realizaran la clasificación hasta llegar a los nodos hoja o finales donde se presentan las predicciones realizadas.

En el algoritmo de RF o bosque aleatorio de árboles de decisión, cada árbol del conjunto se compone de una muestra de datos aleatoriamente extraídos de un conjunto de entrenamiento. Posteriormente, cada uno de ellos se entrena de manera independiente y la mayoría de las predicciones que coincidan, será la estimación final del clasificador RF. Con esta aleatoriedad, se pretende evitar el sobreajuste y llegar a una clasificación más precisa evitando los errores que se pueden producir en algún árbol de decisión [39].

En la Figura 5.6 se puede observar un ejemplo del algoritmo RF. Está compuesto por 3 árboles de decisión que toman 3 muestras aleatorias y después de procesarlo por los diferentes nodos de decisión, se obtienen 3 predicciones, cuya mayoría es la decisión final.

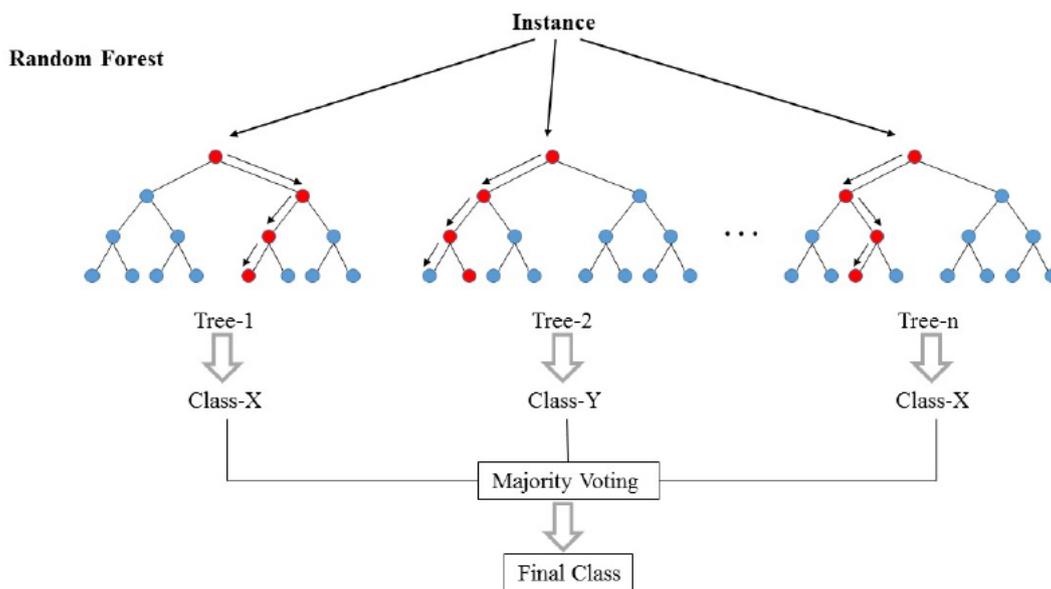


Figura 5.6: Ejemplo visual del algoritmo Random Forest [5].

Se puede implementar y ejecutar un modelo RF a través de la librería *Scikit-Learn* de Python [40]. Para ello, primero deben configurarse unos hiperparámetros. Estos pueden ser el tamaño del nodo, la cantidad de árboles o la cantidad de características muestreadas. Para la selección de hiperparámetros se utiliza la función *GridSearchCV* de *Scikit-Learn* [41], que realiza una búsqueda exhaustiva sobre valores de parámetros específicos para un modelo. Se introduce el modelo y una lista '*param_grid*' con los hiperparámetros y los posibles valores de cada uno de ellos, se prueban todas las combinaciones posibles hasta quedarse con el estimador con mayor

precisión. En este caso, los parámetros son:

- **N_estimators:** El número de árboles en el bosque. Establecido en 500 ya que es el valor indicado para el mejor estimador.
- **Criterion:** La función de pérdida que se utiliza para medir la calidad de una división. En este caso «*entropy*» (entropía) que es una medida de información que indica el desorden de las características con el objetivo. La división óptima es elegida por la característica con menos entropía [42].
- **Random_state:** Semilla aleatoria para controlar la aleatoriedad del modelo. Debe estar fijado en un valor concreto para que siempre coja los mismos datos y así se obtenga siempre el mismo modelo con las mismas métricas.
- **Max_depth:** La profundidad máxima de cada árbol en el bosque. Se establece en 10 de esta manera no se expande de forma indefinida.
- **Min_samples_split:** El número mínimo de muestras necesarias para dividir un nodo interno. En este caso 2.
- **N_jobs:** Si se establece en -1, se utilizan todos los procesadores disponibles.

Aunque RF gracias a su aleatoriedad evita en gran parte el sobreajuste, existen los métodos de *Cross Validation* (CV) que ayudan a que el modelo sea más preciso. CV permite estimar la capacidad predictiva de los modelos cuando se aplican a nuevos datos, haciendo uso de los datos que ya tenemos. CV ajusta el modelo empleando un subconjunto de observaciones del conjunto de datos y se evalúa (calcula la exactitud, que mide cómo de bueno es el modelo) con las observaciones restantes. Esto se repite varias veces y se obtiene la media de los resultados. Gracias a las múltiples evaluaciones, se compensan las posibles desviaciones que puedan surgir en cada predicción. La diferencia entre los distintos métodos de CV es la forma en la que se generan los subconjuntos de entrenamiento y validación [43].

En este caso se usa un *Repeated Stratified K-Fold* de la librería *Scikit-Learn* [44]. En los métodos *K-Fold*, se realiza una división de la base de datos en k muestras distintas. De éstas, k-1 subconjuntos son utilizados para el entrenamiento y uno es utilizado para la validación. Este proceso se repite k veces hasta que todos los subconjuntos han sido utilizados para la validación. *Stratified* hace referencia a que se mantiene la misma proporción de clases (0 y 1) en el entrenamiento y en la validación. *Repeated* hace referencia al número de veces que se repetirá este proceso de división de la base de datos y entrenamiento. En este estudio, se realizan

5 *folds* o subconjuntos que servirán para hacer 5 entrenamientos y validaciones al igual que en la Figura 5.7. El número de repeticiones es 1 ya que es el valor con el que se obtienen mejores resultados y aumentando su número se realizan demasiados entrenamientos con los mismos datos, reduciendo su precisión.

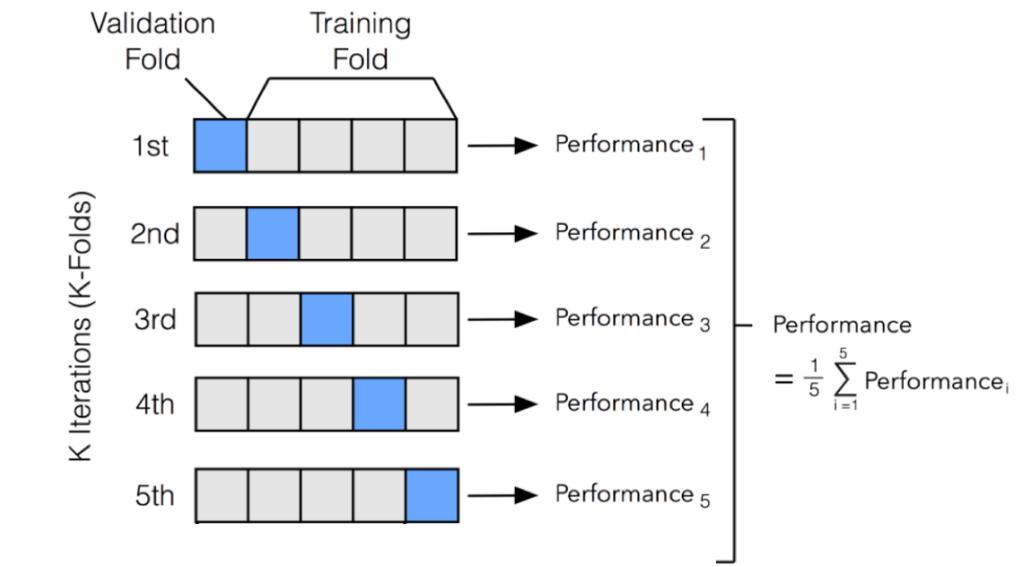


Figura 5.7: Ejemplo visual del método k-Fold con 5 subconjuntos [6].

5.7. Evaluación de la Clasificación

Se utilizan diferentes figuras de mérito para poder evaluar los clasificadores obtenidos:

- **Accuracy o Exactitud:** mide la fracción de predicciones que el modelo realizó de manera acertada. Es un valor entre 0 y 1 [45]. Útil para conjuntos de datos balanceados.

$$\frac{TP + TN}{TP + FP + TN + FN} \quad (5.1)$$

- **Sensibilidad:** mide la proporción de positivos (patológicos) que están identificados correctamente por el modelo entre todos los positivos reales [45]. Informa de la cantidad de positivos que el modelo de ML es capaz de identificar.

$$\frac{TP}{TP + FN} \quad (5.2)$$

- Especificidad: mide la proporción de negativos (sanos) que están identificados correctamente por el modelo entre todos los negativos reales [45]. Informa de la cantidad de negativos que el modelo de ML es capaz de identificar.

$$\frac{TN}{TN + FP} \quad (5.3)$$

- Precisión: mide la proporción de positivos (patológicos) que están identificados correctamente por el modelo entre todos los positivos clasificados [45]. Permite medir la calidad del modelo, cuantos positivos son identificados correctamente.

$$\frac{TP}{TP + FP} \quad (5.4)$$

- F1 Score: combina precisión y sensibilidad realizando un balance entre la cantidad y la calidad de la clasificación de positivos [45].

$$2 * \frac{\textit{precision} * \textit{sensibilidad}}{\textit{precision} + \textit{sensibilidad}} \quad (5.5)$$

- ROC-AUC (área debajo de la curva ROC): se utiliza para evaluar el rendimiento de los modelos de clasificación binaria. Se puede ver un ejemplo en la Figura 5.8 [46]. En el eje X se representa la tasa de falsos positivos (1-especificidad) y en el eje Y la tasa de verdaderos positivos (sensibilidad). La línea discontinua diagonal representa un clasificador aleatorio. Si la curva se presenta por encima, el modelo entrenado proporciona más poder de predicción que el clasificador aleatorio. El AUC es una métrica numérica que mide el área bajo la curva ROC. Cuánto mayor es la puntuación AUC, mejor es el rendimiento de un modelo. En un clasificador perfecto el AUC=1 y en un clasificador aleatorio AUC=0.5.

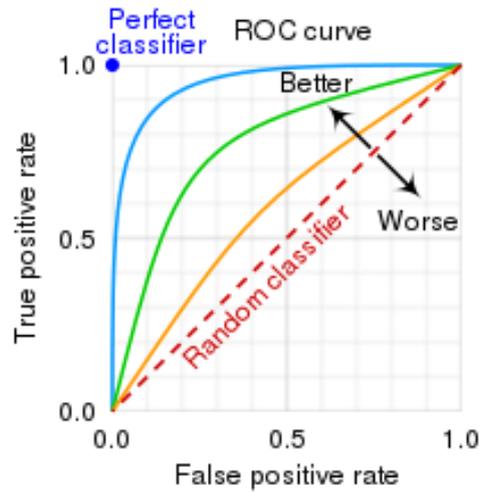


Figura 5.8: Ejemplo visual de varias curvas ROC [7].

5.8. Estudio de las Características Radiómicas

Una vez evaluado el modelo, se requiere ver cuáles han sido las características más relevantes para su clasificación y cómo estas describen la patología que se quería clasificar. Para ello se utiliza SHAP, una de las librerías de explicabilidad de modelos más conocida y utilizada [47]. Este algoritmo utiliza la teoría de juegos para descubrir qué variables son las más influyentes en las predicciones realizadas por un modelo de ML. Se utiliza *shap.TreeExplainer(model)* ya que se realiza una clasificación mediante árboles de decisión. Ver ejemplo en la Figura 5.9.

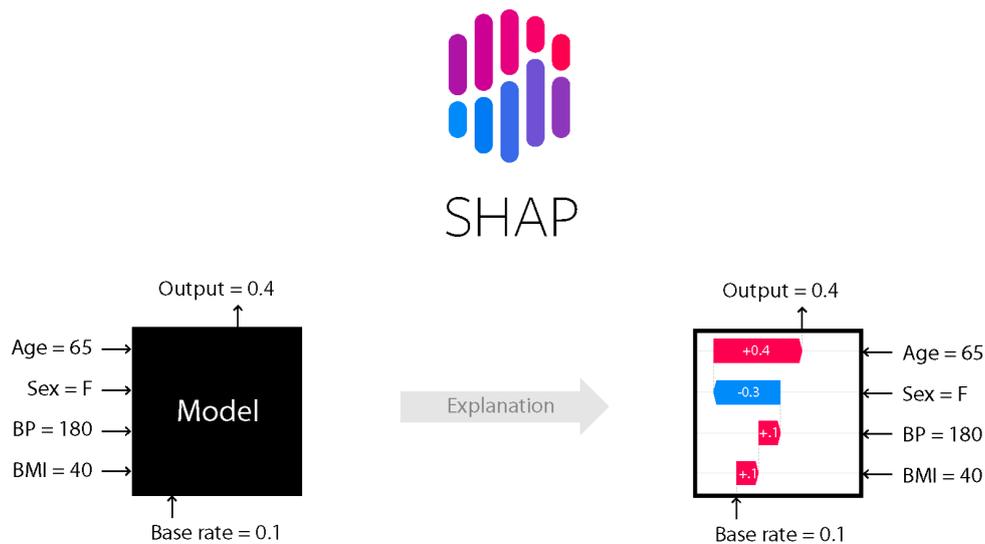


Figura 5.9: Ejemplo visual del algoritmo SHAP.

Capítulo 6

Resultados

6.1. Resultados de la evaluación de los Clasificadores

En la sección 5.5 se presentan las 4 bases de datos a utilizar. A continuación, se muestran los 4 clasificadores obtenidos después de entrenar el modelo de ML con cada una de las bases de datos.

- **Clasificador 1:** clasifica cualquier ganglio en patológico (1) o sano (0) considerando solamente imágenes ponderadas en T1. Entrenado con una base de datos de 52 ganglios y 119 características radiómicas.
- **Clasificador 2:** clasifica cualquier ganglio en patológico (1) o sano (0) considerando 2 tipos de imágenes distintas, imágenes ponderadas en T1 e imágenes dinámicas. Entrenado con una base de datos de 30 ganglios y 203 características radiómicas.
- **Clasificador 3:** clasifica ganglios de pacientes con CL en patológico (1) o sano (0) considerando 2 tipos de imágenes distintas, imágenes ponderadas en T1 e imágenes dinámicas. Entrenado con una base de datos de 16 ganglios y 203 características radiómicas.
- **Clasificador 4:** clasifica ganglios de pacientes con CD en patológico (1) o sano (0) considerando 2 tipos de imágenes distintas, imágenes ponderadas en T1 e imágenes dinámicas. Entrenado con una base de datos de 14 ganglios y 203 características radiómicas.

Para evaluar el rendimiento de los clasificadores se utilizan las diferentes figuras de mérito mencionadas en la sección 5.7.

En la Tabla 6.1, podemos ver a la izquierda las figuras de mérito y a la derecha los valores obtenidos y su desviación estándar para cada clasificador. Los valores han sido obtenidos a partir de la media de las métricas calculadas para cada uno de los subconjuntos del *Repeated Stratified K-Fold*. Además, la desviación estándar nos indica cuál es la dispersión de los datos validados.

Métrica	Clasificador 1	Clasificador 2	Clasificador 3	Clasificador 4
Exactitud	0.658 ± 0.145	0.800 ± 0.067	0.671 ± 0.258	0.600 ± 0.327
Sensibilidad	0.587 ± 0.148	0.867 ± 0.163	0.600 ± 0.406	0.700 ± 0.400
Especificidad	0.747 ± 0.254	0.733 ± 0.133	0.750 ± 0.354	0.600 ± 0.374
Precisión	0.756 ± 0.224	0.783 ± 0.113	0.613 ± 0.405	0.600 ± 0.374
F1 Score	0.631 ± 0.144	0.808 ± 0.074	0.581 ± 0.373	0.600 ± 0.327

Tabla 6.1: Figuras de Mérito para cada uno de los clasificadores.

En la Figura 6.1, se puede observar la curva ROC junto al valor AUC asociado para cada uno de los clasificadores. En este caso, la curva ROC también es la media de todas las curvas calculadas para cada uno de los subconjuntos del *Repeated Stratified K-Fold*.

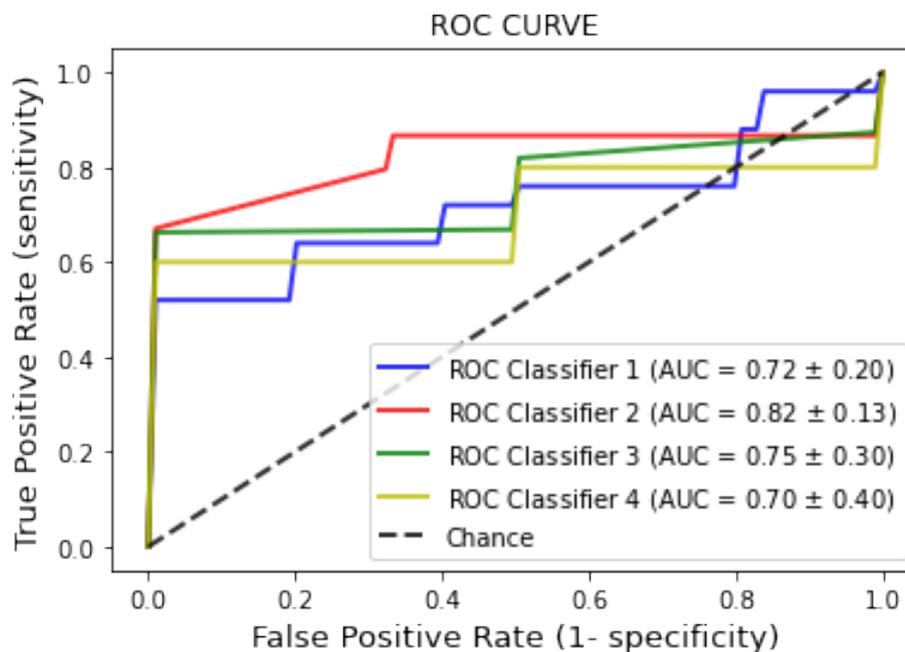


Figura 6.1: Curva ROC y valor AUC para cada clasificador.

6.2. Resultados del estudio de las Características Radiómicas

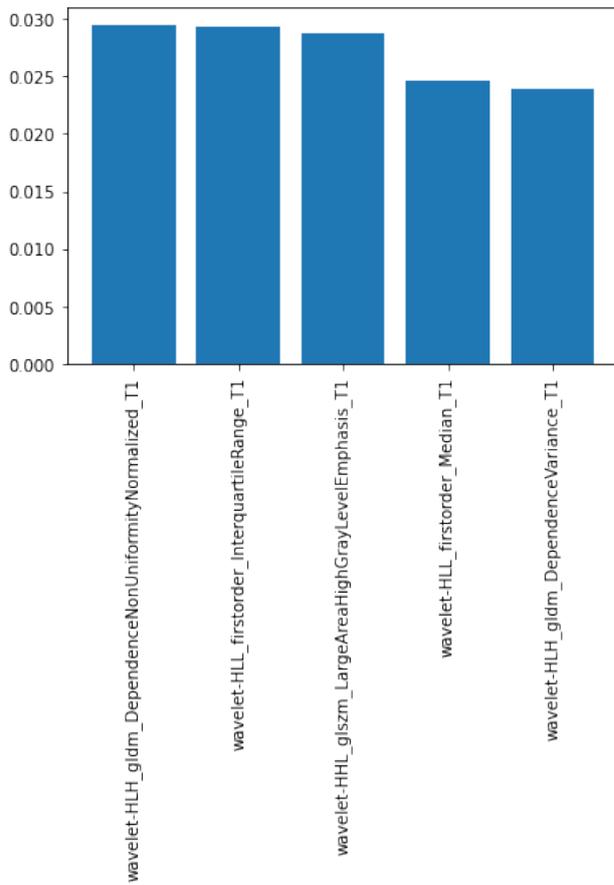
Aparte de la evaluación del rendimiento de los clasificadores, se realiza un estudio de características gracias al algoritmo SHAP que permite conocer cuáles han sido las características más influyentes. En la Figura 6.2 se presentan 4 gráficos con las 5 características más importantes de cada clasificador y a continuación, se explicarán cada una de ellas.

En el primer clasificador, se evalúa la clasificación de ganglios patológicos (1) y sanos (0) en imágenes ponderadas en T1. Se puede observar que las 5 características más importantes han sido extraídas con el filtro wavelet que es el filtro que produce 8 descomposiciones. Las características son:

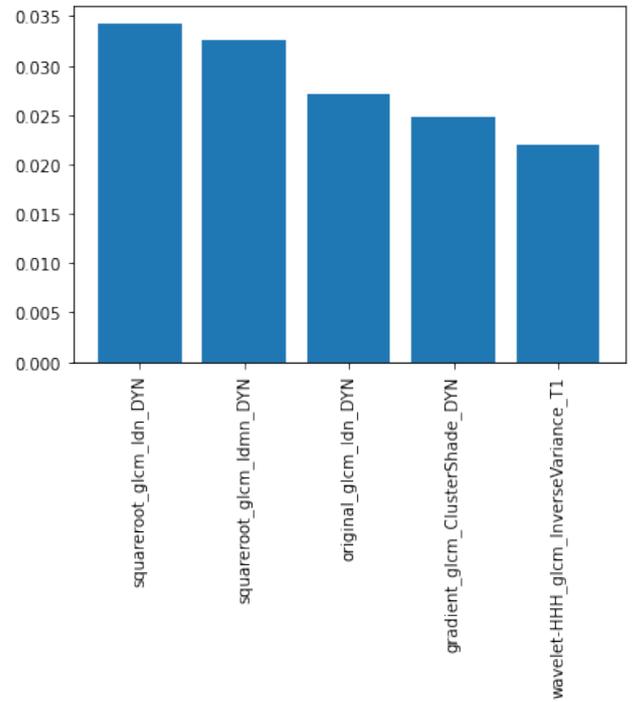
- **DependenceNonUniformityNormalized**: Mide la similitud de la dependencia en toda la imagen, donde un valor más bajo indica más homogeneidad entre las dependencias de la imagen.
- **InterquartileRange**: son los percentiles 25 y 75, ordena de menor a mayor los elementos de la matriz de imágenes.
- **LargeAreaHighGrayLevelEmphasis**: mide la proporción en la imagen de zonas grandes con valores de nivel de gris más altos.
- **Median**: calcula la mediana de la intensidad del nivel de gris.
- **DependenceVariance**: Mide la variación de dependencia en la imagen.

En el segundo, se clasifican ganglios patológicos (1) y sanos (0) en imágenes ponderadas en T1 e imágenes dinámicas. Al introducir las imágenes dinámicas, éstas obtienen mayor importancia ya que 4 de las 5 características más importantes han salido de imágenes dinámicas. Las características se han extraído de diferentes filtros pero todas ellas del grupo de características Gray Level Co-occurrence Matrix. Las características en cuestión son:

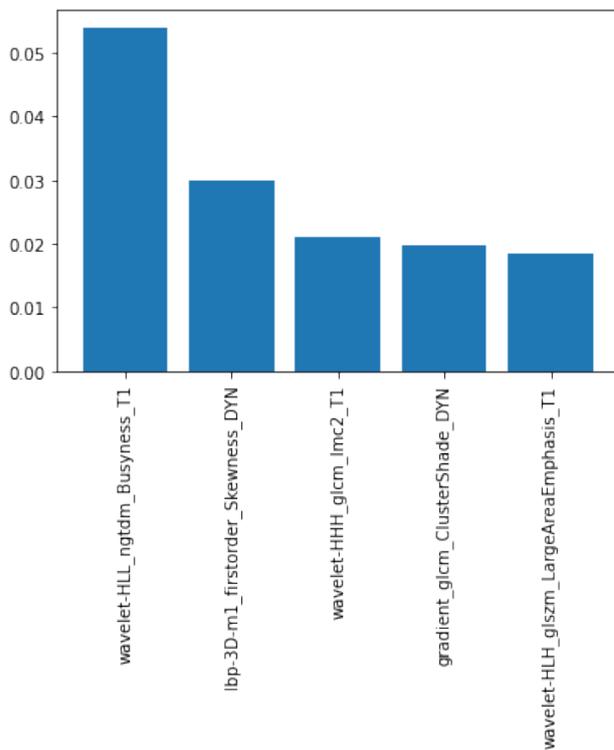
- **Idn**: es una medida de la homogeneidad local de la imagen original.
- **Idn**: es una medida de la homogeneidad local de la imagen filtrada por *squareRoot* que realiza la raíz cuadrada de las intensidades.
- **Idmn**: momento de diferencia inversa normalizado, es otra medida de la homogeneidad local de una imagen.



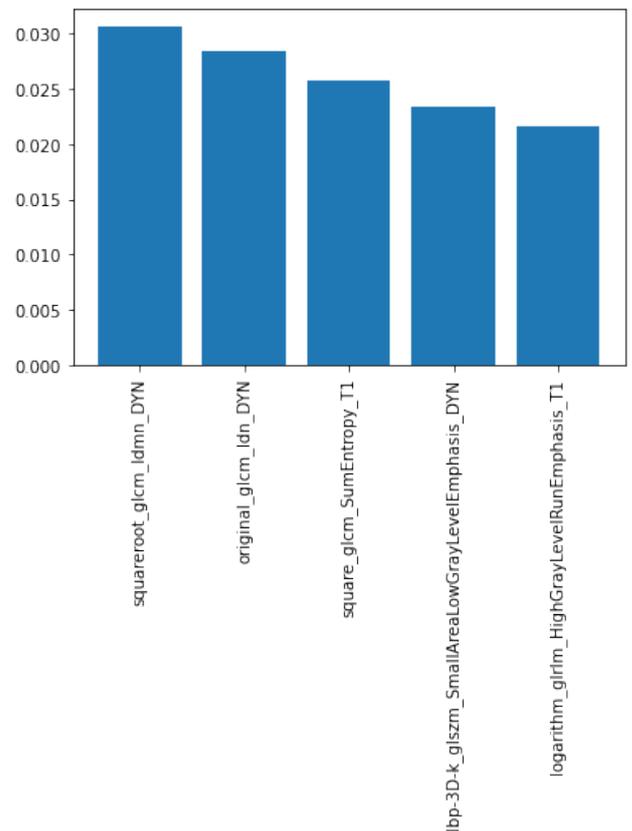
(a) Clasificador 1



(b) Clasificador 2



(c) Clasificador 3



(d) Clasificador 4

Figura 6.2: Gráficos con las 5 características más importantes empleadas en cada clasificador

- ClusterShade: es una medida de la asimetría y uniformidad.
- InverseVariance: es la inversa de la media de las distancias al cuadrado de cada valor de intensidad.

En el tercero, se clasifican ganglios de pacientes con CL en patológicos (1) y sanos (0) en imágenes ponderadas en T1 e imágenes dinámicas. En este caso las características pertenecen tanto a imágenes T1 como imágenes dinámicas. Éstas son obtenidas a partir de diferentes filtros, sin embargo, el filtro wavelet es el más utilizado. Las características son:

- Busyness: Una medida del cambio de un píxel a su vecino.
- Skewness: mide la asimetría de la distribución de valores sobre el valor medio.
- Imc2: evalúa la correlación entre 2 distribuciones (cuantificando la complejidad de la textura).
- ClusterShade: al igual que en el clasificador 2.
- LargeAreaEmphasis: es una medida de la distribución de grandes zonas.

En el cuarto, se clasifican ganglios de pacientes con CD en patológicos (1) y sanos (0) en imágenes ponderadas en T1 e imágenes dinámicas. En este caso, las características pertenecen tanto a imágenes T1 como imágenes dinámicas. Las características son:

- Idmn y Idn: al igual que en el clasificador 2.
- Sumentropy: es la suma de las diferencias de valores de intensidad entre vecinos.
- SmallAreaLowGrayLeveEmphasis: mide la proporción en la imagen de zonas pequeñas con valores de nivel de gris más bajos.
- HighGrayLevelRunEmphasis: mide la concentración de valores de nivel de gris altos en la imagen.

Capítulo 7

Discusión

7.1. Rendimiento de los Clasificadores

No se registran estudios relacionados con la clasificación de ganglios metastásicos que son falsos negativos. Sin embargo existen diversos, mencionados en el Estado del Arte, sobre la clasificación de ganglios patológicos en los que se obtenían muy buenos resultados como una precisión de 0.9. En esta investigación se pueden comparar los resultados con los de otras investigaciones ya que, al fin y al cabo, es una clasificación de GC utilizando radiómica y ML. En cambio, los resultados son notablemente inferiores debido a la dificultad añadida que supone tener ganglios patológicos no detectables en imagen. Por ello, la mejor comparación que se puede hacer es entre los diferentes clasificadores que se han creado para conocer los beneficios y cómo se podrían mejorar para encontrar un buen clasificador de estos falsos negativos.

En el primer clasificador, en el que solo se tienen en cuenta imágenes ponderadas en T1 de 52 ganglios, obtenemos una exactitud de 0.658. Este parámetro mide las predicciones que realiza el modelo adecuadamente, significa que ha acertado más de la mitad de las predicciones, pero no es un valor que se pueda considerar bueno. La sensibilidad es 0.587 y la especificidad es 0.747, lo que indica que al tener una mayor especificidad, el clasificador tiene mayor capacidad de detectar ganglios sanos que enfermos, por lo que sigue identificando algunos ganglios patológicos como sanos. Por consiguiente, no cumple con el objetivo de detectar falsos negativos. En el caso de la Curva ROC, el valor AUC es 0.72, el modelo realiza mejores predicciones que un clasificador aleatorio, pero sigue estando lejos de un clasificador ideal, situándose en algunos puntos por debajo de la predicción aleatoria.

En el segundo clasificador el número de ganglios se ve reducido a 30. En cambio, tenemos

el doble de características radiómicas debido a que se utilizan imágenes en dos modalidades, ponderadas en T1 y dinámicas. En este caso obtenemos mejores resultados que en el anterior clasificador: la precisión es de 0.783, lo cual indica que el 78,3% de los positivos han sido identificados correctamente. Este valor es bastante bueno, además se encuentra bastante cerca del 84,3% de precisión obtenido en el estudio de Richard Ha et al. [32]. La sensibilidad es de 0.867, es decir, tiene una alta capacidad para la detección de ganglios patológicos. En el caso de la curva ROC y su valor AUC de 0.82, indica que el rendimiento de este modelo es mayor al del clasificador 1. Una de las principales conclusiones que podemos extraer de estos resultados y sus comparaciones es que la introducción de imágenes dinámicas mejora notablemente la clasificación.

En el caso del tercer y cuarto clasificador, se divide en dos el conjunto de datos utilizado para el segundo clasificador. Se evalúan ganglios de pacientes con CL y CD por separado. En estos clasificadores, los valores de las figuras de mérito disminuyen. Por ejemplo, la precisión es de 0.613 y 0.600 respectivamente. Estos valores probablemente sean debido a la poca cantidad de datos que se tiene. Se demuestra que no existe diferencia entre ganglios de pacientes con CL y CD, y por lo tanto, se puede utilizar un clasificador general sin hacer distinción del cáncer que padece el paciente.

Como conclusión de la evaluación de los modelos, el clasificador 2 con resultados alrededor del 0.8 podría considerarse un buen clasificador aunque mejorable. Éste tiene en cuenta las imágenes de las dos modalidades de RM más utilizadas para el diagnóstico. Lo que demuestra, que es fundamental considerar más de un tipo de imagen. De esta manera, se podrían añadir otras modalidades como es la imagen ponderada en T2 o la imagen de difusión que pueden aportar otras características y así mejorar la clasificación y hacerla aplicable a la práctica real.

7.2. Importancia de las Características Radiómicas

En el estudio de Dong et al. [31] ya se evidenciaba la importancia de las características radiómicas. Por ello, en esta investigación se obtienen también las características más relevantes, para poder entender cómo funciona la clasificación y poder ayudar a los radiólogos a un mejor diagnóstico.

De las 5 características del segundo y mejor clasificador se pueden extraer diversas conclusiones. Las imágenes dinámicas contribuyen más a la clasificación que las imágenes en T1. Las características han sido extraídas de diversos filtros, y esto es importante, ya que normalmente los radiólogos examinan las imágenes originales. Además, las características que más ayudan a

la clasificación son de la categoría Gray Level Co-occurrence Matrix, es decir, la homogeneidad, simetría y diferencias de intensidades dentro de la imagen, detalles que a simple vista no son fáciles de medir.

Un objetivo secundario de este estudio era valorar la diferencia de características en la clasificación de ganglios de pacientes con CL y la clasificación de ganglios de pacientes con CD. En ambos clasificadores se valoran imágenes tanto ponderadas en T1 como dinámicas, así como imágenes tratadas con diferentes tipos de filtros, lo que evidencia, una vez más, la aportación de información por parte de diferentes imágenes. En cuanto a las características, la clasificación de ganglios de pacientes con CL se centra una vez más en la asimetría, los cambios entre los píxeles vecinos y su correlación, es decir, analiza la distribución de intensidades en la imagen. En cuanto a la clasificación de ganglios de pacientes con CD, también se repiten características relacionadas con la homogeneidad y distribución de píxeles, pero además se tienen en cuenta la concentración de los grises y el tamaño de las zonas con el mismo valor de gris. Como ya se mencionaba en la discusión del rendimiento del modelo, estos dos últimos clasificadores no son los más precisos, por ello, la evaluación de sus características debe ser acorde a los resultados. Sin embargo, se puede observar que ambos clasificadores, a través de diferentes cálculos de características, tienen en cuenta los mismos parámetros (homogeneidad y distribución de la información en la imagen), y no existen diferencias a la hora de su clasificación.

Las características extraídas son fruto de cálculos computacionales, resultados que no son posibles medir a simple vista, además de la extracción de características de diferentes imágenes a través de diferentes tratamientos. Esto evidencia la capacidad que la radiómica junto al ML tienen para detectar y clasificar detalles inapreciables al ojo humano.

Capítulo 8

Conclusiones

8.1. Aportaciones del estudio

Como se viene demostrando desde el comienzo del proyecto, la investigación en el cáncer de mama es una de las principales líneas de investigación hoy en día debido al impacto que tiene en la sociedad. Además, el diagnóstico precoz mediante técnicas no invasivas también es uno de los principales objetivos actuales en la medicina para conseguir de esta manera una mayor supervivencia de los pacientes y menor invasión.

El ML, como se menciona en otros Capítulos, ya ha demostrado su gran impacto en la medicina y su posible aplicación a cualquier ámbito de la salud. La radiómica es una ciencia recientemente aceptada para la investigación del cáncer, sin embargo, prometedora, pudiendo formar parte de potentes clasificadores o aumentando el conocimiento sobre las imágenes médicas y la anatomía.

En este estudio, se ha realizado una clasificación de GC que los radiólogos del HURyC no han sido capaces de identificar como malignos debido a su morfología aparentemente normal en las imágenes de RM y que luego mediante la BGC han sido diagnosticados como malignos.

Para ello, se han seleccionado los pacientes que cumplían los requisitos del estudio y sus imágenes de RM del PACS. Se han segmentado manualmente los ganglios a través de 3D Slicer, se han extraído las 1688 características radiómicas y se ha seleccionado las más relevantes. A continuación, se ha entrenado un modelo RF con CV para evaluar las aportaciones de diferentes bases de datos y se han generado las diferentes figuras de mérito que permiten evaluar la validez y precisión de los diferentes clasificadores.

Con todo ello, se ha obtenido un clasificador de ganglios con buena precisión a través de

imágenes de RM mamaria ponderadas en T1 y del minuto 1 del estudio dinámico.

La creación de una buena base de datos es importante y determinante en el rendimiento del modelo. Durante la etapa de creación de la base de datos, se descartan pacientes con CD que podrían servir para el entrenamiento del modelo, ya que este clasificador no hace distinción entre el tipo de cáncer y no es necesario tener una base de datos equilibrada entre CL y CD. El aumento del número de pacientes haría más robusto el modelo (como se menciona a continuación en las líneas futuras).

Además, se ha evidenciado la importancia de valorar diferentes tipos de imagen. La imagen dinámica aporta mayor información que la imagen ponderada en T1 y también las imágenes a las que se han aplicado diversos filtros muestran más información que la original. Asimismo, se han demostrado que los rasgos más relevantes son los de homogeneidad y simetría. Éstas son características difíciles de medir a simple vista que prueban la gran capacidad de la radiómica y pueden servir de guía a los radiólogos para un mejor diagnóstico.

8.2. Líneas futuras

El clasificador construido, a pesar de tener una buena precisión, podría tener mejores resultados y de esta manera ser aplicable a la práctica real. Por ello, existen diferentes puntos de esta investigación que podrían extenderse:

- Añadir otros tipos de imagen que puedan aportar mayor información como puede ser la imagen de RM de difusión o la imagen de RM ponderada en T2.
- Añadir más imágenes de otros pacientes ya que la base de datos del estudio es bastante reducida y así hacer más robusto el modelo.
- Entrenar otros modelos de ML que pueden ofrecer otros resultados.
- Realizar una segmentación automática mediante herramientas de ML que redujera la subjetividad y el sesgo propio de la segmentación manual.

Además de las mejoras que puede tener este estudio, se ha demostrado la capacidad que tiene la radiómica y el ML en la investigación del cáncer, en concreto en la investigación de la predicción de afectación ganglionar. La extensión de este estudio podría implicar un cambio importante en el tratamiento del CM, desarrollando una aplicación que apoyará al facultativo durante el proceso diagnóstico evitando intervenir al paciente, mejorando la eficacia de los tratamientos y la calidad de vida de los pacientes.

Apéndice A

Anexo

Anexo A.1: Documento *params.yaml* que indica los filtros y las características a extraer por PyRadiomics.

```
# Image types to use: "Original" for unfiltered image ,
# for possible filters , see documentation .
imageType:
  Original: {} # for dictionaries / mappings ,
  # None values are not allowed .
  Wavelet: {}
  LoG: {}
  Square: {}
  SquareRoot: {}
  Logarithm: {}
  Exponential: {}
  Gradient: {}
  LBP2D: {}
  LBP3D: {}

# Featureclasses , from which features must be calculated .
featureClass:
  shape:
  firstorder: []
```

glcm :

glrlm :

glszm :

gldm :

ngtdm :

Bibliografía

- [1] Dra. Ana Santaballa Bertrán. SEOM: Sociedad Española de Oncología Médica. Cáncer de mama, 2023. Consultado el 23 de mayo de 2023. Disponible en <https://seom.org/info-sobre-el-cancer/cancer-de-mama>.
- [2] Osakidetza. Biopsia ganglio linfático centinela. Consultado el 31 de mayo de 2023. Disponible en <https://osieec.osakidetza.eus/esano/biopsia-ganglio-linfatico-centinela/>.
- [3] Rodolfo de Juana. Machine learning: cuando aprenden las máquinas, 2019. Consultado el 5 de junio de 2023. Disponible en <https://www.muycomputerpro.com/2019/03/07/machine-learning-cuando-aprenden-las-maquinas>.
- [4] Vicente Rodríguez. Decision trees/Árboles de decisión para clasificar en python, 2018. Consultado el 21 de junio de 2023. Disponible en <https://vincentblog.xyz/posts/decision-trees-arboles-de-decision-para-clasificar-en-python>.
- [5] Gerard Sánchez. Random forest classifier (sp500), 2021. Consultado el 21 de junio de 2023. Disponible en <https://gsnchez.com/blog/article/Random-forest-classifier-sp500>.
- [6] model_selection. Consultado el 21 de junio de 2023. Disponible en http://ethen8181.github.io/machine-learning/model_selection/model_selection.html.
- [7] Receiver operating characteristic. Consultado el 24 de junio de 2023. Disponible en https://en.wikipedia.org/wiki/Receiver_operating_characteristic#/media/File:Roc_curve.svg.
- [8] Dra. Ana Santaballa Bertrán. SEOM: Sociedad Española de Oncología Médica. Epidemiología y factores de riesgo, 2023. Consultado el 23 de mayo de 2023. Disponible en <https://seom.org/info-sobre-el-cancer/cancer-de-mama?start=2>.

- [9] Jamie DePolo. Tipos de cáncer de mama, 2022. Consultado el 23 de mayo de 2023. Disponible en <https://www.breastcancer.org/es/tipos>.
- [10] Comunidad de Madrid. Consultado el 25 de mayo de 2023. Detección precoz del cáncer de mama, 2017. Disponible en <https://www.breastcancer.org/es/tipos>.
- [11] Instituto Nacional del Cáncer. Mamografías, 2017. Consultado el 25 de mayo de 2023. Disponible en <https://www.cancer.gov/espanol/tipos/seno/hoja-informativa-mamografias#qu-es-el-sistema-de-base-de-datos-e-informes-de-imagenes-de-la-mama-bi-rads>.
- [12] Centers for Disease Control and Prevention. ¿Cómo se diagnostica el cáncer de mama?, 2022. Consultado el 25 de mayo de 2023. Disponible en https://www.cdc.gov/spanish/cancer/breast/basic_info/diagnosis.htm.
- [13] Instituto Nacional del Cáncer. Tratamiento del cáncer de seno (mama) en adultas, 2023. Consultado el 31 de mayo de 2023. Disponible en <https://www.cancer.gov/espanol/tipos/seno/paciente/tratamiento-seno-pdq>.
- [14] Instituto Nacional del Cáncer. Quimioterapia para tratar el cáncer, 2015. Consultado el 31 de mayo de 2023. Disponible en <https://www.cancer.gov/espanol/cancer/tratamiento/tipos/quimioterapia>.
- [15] Instituto Nacional del Cáncer. Radioterapia para el cáncer, 2017. Consultado el 31 de mayo de 2023. Disponible en <https://www.cancer.gov/espanol/cancer/tratamiento/tipos/radioterapia>.
- [16] Instituto Nacional del Cáncer. Biopsia de ganglio linfático centinela, 2022. Consultado el 31 de mayo de 2023. Disponible en <https://www.cancer.gov/espanol/cancer/diagnostico-estadificacion/estadificacion/hoja-informativa-ganglio-centinela>.
- [17] José Luis Fougo, Mário Dinis-Ribeiro, Cláudia Araújo, Teresa Dias, Paulo Reis, Laurinda Giesteira, Fernando Castro, and Hernâni Silva. Impacto de la linfadenectomía sobre la recurrencia axilar y la morbilidad del miembro superior en pacientes con cáncer de mama y ganglio centinela negativo. estudio prospectivo aleatorizado. *Cirugía Española*, 89(5):307–316, 2011.

- [18] Dra. Macarena Rodríguez Fraile. Ganglio centinela, 2022. Consultado el 31 de mayo de 2023. Disponible en <https://www.cun.es/enfermedades-tratamientos/pruebas-diagnosticas/ganglio-centinela>.
- [19] Pere Culell, Lluís Solernou, Manel Fraile, Jordi Tarazona, Ana Miguel, and Victoria Villa. Técnica del ganglio centinela en la cirugía del cáncer de mama en un hospital sin servicio de medicina nuclear. *Cirugía Española*, 81(3):126–129, 2007.
- [20] Kocak M. Resonancia magnética, 2021. Consultado el 3 de junio de 2023. Disponible en <https://www.msmanuals.com/es-es/professional/temas-especiales/principios-de-estudios-por-la-imagen-radiol%C3%B3gicas/resonancia-magn%C3%A9tica>.
- [21] Ritse M Mann, Nariya Cho, and Linda Moy. Breast mri: state of the art. *Radiology*, 292(3):520–536, 2019.
- [22] Marcelo Gálvez, Mauricio Farías, Takeshi Asahi, and Eduardo Bravo. Cálculo de tiempos t1 y t2 in vitro. *Revista chilena de radiología*, 11(3):109–115, 2005.
- [23] Pablo Soffía. Difusión por resonancia magnética: bases y aplicaciones oncológicas en órganos extracraneanos. *Revista chilena de radiología*, 15:17–24, 2009.
- [24] Rocío B Ruiz and Juan D Velásquez. Inteligencia artificial al servicio de la salud del futuro. *Revista Médica Clínica Las Condes*, 34(1):84–91, 2023.
- [25] DataScientest. Machine learning: definición, funcionamiento, usos, 2022. Consultado el 5 de junio de 2023. Disponible en <https://datascientest.com/es/machine-learning-definicion-funcionamiento-usos>.
- [26] Javier Blázquez Sánchez Gema Bruixola Luis Martí-Bonmatí, Ángel Alberich-Bayarri. Radiómica, 2022. Consultado el 5 de junio de 2023. Disponible en <https://www.institutoroche.es/observatorio/radiomica>.
- [27] Allegra Conti, Andrea Duggento, Iole Indovina, Maria Guerrisi, and Nicola Toschi. Radiomics in breast cancer classification and prediction. In *Seminars in cancer biology*, volume 72, pages 238–250. Elsevier, 2021.
- [28] Vishwa S Parekh and Michael A Jacobs. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric mri. *NPJ breast cancer*, 3(1):43, 2017.

- [29] Matheus de Freitas Oliveira Baffa and Aura Conci. Radiomics for breast ir-imaging classification. In *MICCAI Workshop on Medical Image Assisted Biomarkers' Discovery*, pages 10–19. Springer, 2022.
- [30] Nathaniel M Braman, Maryam Etesami, Prateek Prasanna, Christina Dubchuk, Hannah Gilmore, Pallavi Tiwari, Donna Plecha, and Anant Madabhushi. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast dce-mri. *Breast Cancer Research*, 19:1–14, 2017.
- [31] Yuhao Dong, Qianjin Feng, Wei Yang, Zixiao Lu, Chunyan Deng, Lu Zhang, Zhouyang Lian, Jing Liu, Xiaoning Luo, Shufang Pei, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of t2-weighted fat-suppression and diffusion-weighted mri. *European radiology*, 28(2):582–591, 2018.
- [32] Richard Ha, Peter Chang, Jenika Karcich, Simukayi Mutasa, Reza Fardanesh, Ralph T Wynn, Michael Z Liu, and Sachin Jambawalikar. Axillary lymph node evaluation utilizing convolutional neural networks using mri dataset. *Journal of digital imaging*, 31:851–856, 2018.
- [33] 3d slicer image computing platform. Consultado el 12 de junio de 2023. Disponible en <https://www.slicer.org/>.
- [34] Anaconda. Consultado el 12 de junio de 2023. Disponible en <https://www.anaconda.com/>.
- [35] Radiomic features. Consultado el 18 de junio de 2023. Disponible en <https://pyradiomics.readthedocs.io/en/latest/features.html>.
- [36] Customizing the extraction. Consultado el 18 de junio de 2023. Disponible en <https://pyradiomics.readthedocs.io/en/latest/customization.html>.
- [37] Redacción KeepCoding. Qué es la selección de características en machine learning, 2022. Consultado el 20 de junio de 2023. Disponible en <https://keepcoding.io/blog/seleccion-de-caracteristicas-en-machine-learning/>.
- [38] Sklearn.feature_selection.selectpercentile. Consultado el 20 de junio de 2023. Disponible en https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectPercentile.html.

- [39] What is random forest? Consultado el 21 de junio de 2023. Disponible en <https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems.>
- [40] Sklearn.ensemble.randomforestclassifier. Consultado el 21 de junio de 2023. Disponible en <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>.
- [41] Sklearn.model_selection.gridsearchcv. Consultado el 21 de junio de 2023. Disponible en https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html.
- [42] Aznar P. Decision trees: Gini vs entropy, 2020. Consultado el 21 de junio de 2023. Disponible en <https://quantdare.com/decision-trees-gini-vs-entropy/>.
- [43] J Amat Rodrigo. Validación de modelos predictivos: Cross-validation, oneleaveout, bootstrapping. 2020. Consultado el 21 de junio de 2023. Disponible en https://www.cienciadedatos.net/documentos/30_cross-validation_oneleaveout_bootstrap.
- [44] Sklearn.model_selection.repeatedstratifiedkfold. Consultado el 21 de junio de 2023. Disponible en https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RepeatedStratifiedKFold.html.
- [45] Heras JM. Precision, recall, f1, accuracy en clasificación, 2019. Consultado el 24 de junio de 2023. Disponible en <https://www.iartificial.net/precision-recall-f1-accuracy-en-clasificacion/>.
- [46] Multiclass receiver operating characteristic (roc). Consultado el 24 de junio de 2023. Disponible en https://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html.
- [47] Welcome to the shap documentation. Consultado el 24 de junio de 2023. Disponible en <https://shap.readthedocs.io/en/latest/>.