

Trabajo Fin de Grado
Grado en Ingeniería de las Tecnologías de
Telecomunicación

Radiómica para el análisis automático de
mamografías

Autor: José Manuel Román Moguer

Tutoras: María del Carmen Serrano Gotarredona y Begoña Acha Piñero

Dpto. Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2024



Trabajo Fin de Grado
Grado en Ingeniería de las Tecnologías de Telecomunicación

Radiómica para el análisis automático de mamografías

Autor:

José Manuel Román Moguer

Tutoras:

María del Carmen Serrano Gotarredona y Begoña Acha Piñero

Catedráticas de Universidad

Dpto. de Teoría de la Señal y Comunicaciones

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2024

Trabajo Fin de Grado: Radiómica para el análisis automático de mamografías

Autor: José Manuel Román Moguer

Tutoras: María del Carmen Serrano Gotarredona y
Begoña Acha Piñero

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2024

El Secretario del Tribunal

A mi familia

A mis maestros

A mis amigos

AGRADECIMIENTOS

Para concluir mi etapa como estudiante del grado de Ingeniería de las Tecnologías de Telecomunicación, me gustaría agradecer a todas y cada una de las personas que han hecho posible, junto a mí, el finalizar este largo y complejo camino.

En primer lugar, me gustaría agradecer a esas dos personas que nunca me han soltado de la mano, a mis padres: José Manuel y Conchi. Gracias por desviviros tanto por mí como por mis hermanas y por cuidarnos siempre tan bien como lo han hecho. Gracias a ellos por aportarme cada uno su apoyo personal e incondicional; y por permitirme ser el hombre en el que, sin darme cuenta, me he convertido.

También me gustaría agradecer a mis hermanas, Conchi y Elisabeth, por guiarme siempre y sostenerme en los momentos más complicados, por ver en mí una capacidad para lograr metas que ni yo mismo veía.

Me gustaría agradecer a una persona que ha sido bastante influyente para mí, mi cuñado Quique, el cual pese a llegar más tarde a mi vida, siempre ha estado en el momento clave, confiando y sacando lo mejor de mí.

A mis amigos del pueblo, mil gracias a cada uno de vosotros. Sin vuestras largas charlas y vuestro apoyo incondicional, tanto en los buenos como en los malos momentos, nada de esto habría sido posible.

A mis amigos de la universidad, los cuales en los momentos de mayor debilidad en este proceso, habéis sacado un empujón de mí, que os lo voy a agradecer de por vida.

También me gustaría agradecer a mis amigos de Bolonia. Llegasteis a mi vida de manera inesperada, pero os quedasteis de la forma más sana. Gracias por ser un apoyo fundamental y por estar conmigo en cada momento importante.

Por último, quiero expresar mi profundo agradecimiento a mis profesores a lo largo de la carrera, y en especial a mis tutoras de TFG, Carmen y Begoña. Su paciencia infinita conmigo durante este largo proceso y su apoyo constante han sido fundamentales para la realización de este trabajo. Su dedicación ha sido una fuente de inspiración para mí como futuro ingeniero.

Sé que es probable que me deje a algunas personas sin mencionar, pero este agradecimiento es solo una pequeña parte de lo que todos vosotros merecéis. Sin vosotros, nada de esto hubiera sido posible.

Gracias a todos, de corazón.

José Manuel Román Moguer

Sevilla, 2024

RESUMEN

El cáncer de mama es una de las principales causas de mortalidad entre las mujeres de la sociedad española, su detección temprana es crucial para mejorar las tasas de supervivencia. El uso de la inteligencia artificial en el campo de la investigación biomédica puede aportar bastante al campo de la detección de esta enfermedad.

En este proyecto se desarrollará un algoritmo, basado en aprendizaje máquina, en el cual pretendemos discriminar entre masas mamarias benignas y malignas. Para ello, se extraerán características estadísticas, de forma y de textura utilizando la librería PyRadiomics, en lenguaje Python, y *Measure* de Skimage. A continuación, con la ayuda de métodos estadísticos, como el uso del P-Value o la eliminación a partir de la correlación, se reducirá el número de las características extraídas, quedándonos únicamente con aquellas características más significativas. Finalmente, se entrenará un modelo SVM (máquina de soporte vectorial) para la clasificación, a partir de las características extraídas y seleccionadas significativamente tomando en consideración estas imágenes, que permita realizar la clasificación en masas benignas o malignas.

Con este proyecto se pretende ayudar a los profesionales sanitarios, especialistas en el campo de la detección oncológica, mejorando la detección y el tratamiento de la enfermedad.

ABSTRACT

Breast cancer is one of the main causes of mortality among women in Spanish society, and its early detection is crucial to improve survival rates. The use of artificial intelligence in the field of biomedical research can make a significant contribution to the detection of this disease.

In this project we will develop an algorithm, based on machine learning, in which we aim to discriminate between benign and malignant breast masses. To do this, statistical, shape and texture features will be extracted using the PyRadiomics library, in Python language, and Measure from Skimage. Then, with the help of statistical methods, such as the use of P-Value or elimination from correlation, the volume of extracted features will be reduced, keeping only the most significant features. Finally, an SVM (support vector machine) model will be trained for classification, based on the features extracted and significantly selected from these images, which will allow classification into benign or malignant masses.

The aim of this project is to help health professionals, specialists in the field of oncological detection, to improve the detection and treatment of the disease.

Agradecimientos	ix
Resumen	xi
Abstract	xiii
Índice	xiv
Índice de Tablas	xv
Índice de Ilustraciones	xvii
1 Contexto clínico	1
2 Estado del arte	2
2.1. <i>Inteligencia artificial</i>	2
2.1.1 Machine Learning	2
2.1.2 Deep Learning	3
2.2. <i>Radiómica</i>	5
3 Metodología	8
3.1. <i>Extracción de características</i>	8
3.2. <i>Selección de características</i>	11
3.3. <i>Clasificación de características</i>	12
4 Evaluación del algoritmo	14
5 Resultados	16
5.1. <i>Resultados de las fases de extracción y selección.</i>	16
5.2. <i>Resultados de la fase de clasificación.</i>	23
6 Conclusiones	24
Referencias	25
Códigos	28
<i>Código extracción con Radiomics.</i>	28
<i>Código extracción con Measure.</i>	29
<i>Código selección con P-Value.</i>	31
<i>Código clasificación con SVM.</i>	32

ÍNDICE DE TABLAS

Tabla 1. Parámetros de clasificación

23

ÍNDICE DE ILUSTRACIONES

Ilustración 1. Flujo de trabajo de la radiómica [25]	7
Ilustración 2. Importación de librerías	8
Ilustración 3. Extractor de radiomics (para el caso del grupo ‘gldm’)	9
Ilustración 4. Manipulación de lista y almacenado en archivo “.mat”	9
Ilustración 5. Extracción de regiones	9
Ilustración 6. Propiedades extraídas de las regiones	10
Ilustración 7. Comando de Matlab “ttest2”	12
Ilustración 8. Caso linealmente separable [47]	13
Ilustración 9. Caso linealmente no separable [47]	13
Ilustración 10. Matriz de confusión	14
Ilustración 11. Histograma “Percentil 90”	17
Ilustración 12. Histograma “Mediana”	18
Ilustración 13. Histograma “RMS”	18
Ilustración 14. Histograma “Ratio Perímetro Superficie”	19
Ilustración 15. Histograma “Esfericidad”	20
Ilustración 16. Histograma “Excentricidad”	20
Ilustración 17. Histograma “Compacidad”	21
Ilustración 18. Histograma “GLDM No dependencia en uniformidad”	21
Ilustración 19. Histograma “NGTDM Contraste”	22

1 CONTEXTO CLÍNICO

El cáncer representa uno de los principales problemas de salud para la población mundial. En España, al igual que en muchos otros países, el cáncer de mama es una de las enfermedades más frecuentes entre la población femenina [1].

Según datos extraídos de la OMS (Organización Mundial de la Salud), en el año 2012 el diagnóstico mundial fue de unos 1,67 millones de mujeres enfermas por este tipo de cáncer; mientras que en el año 2020, este número llegó a aumentar hasta los 2,3 millones de mujeres enfermas [2], lo que muestra un amplio aumento con el paso de los años, debido al incremento y al avance en las técnicas de detección.

El éxito ante cualquier enfermedad radica en el desarrollo en su detección y tratamiento. En nuestro caso, centrándonos en la detección de esta enfermedad, tras tantos años de investigación, los profesionales en el sector han llegado a la conclusión de que tanto las características de formas como las de textura de la masa tumoral definen, en su gran mayoría, la malignidad de ésta. Los parámetros característicos que alarman de malignidad son los siguientes [3]: aumento de la densidad en el tejido mamario, distorsión en la morfología mamaria, heterogeneidad en la densidad del tejido de la masa, contornos irregulares, etc.

Sin embargo, el desarrollo de la tecnología informática, en concreto el desarrollo de la conocida IA (Inteligencia Artificial) está revolucionando el mundo de la medicina en muchos campos, ya que está demostrando ofrecer una gran mejora en cuanto a resultados junto a los que ya obteníamos con técnicas de detección más tradicionales [4].

Hoy en día, se recogen enormes cantidades de imágenes en bases de datos mundialmente conocidas para ofrecerse como ayuda en el desarrollo de la investigación biomédica. En este contexto, hemos hecho uso de la conocida base de datos CBIS-DDSM (Curated Breast Imaging Subset of Digital Database for Screening Mammography) [5], la cual es una versión mejorada y actualizada de la base de datos DDSM. Dicha base de datos está especializada en imágenes mamográficas, conteniendo imágenes sobre casos malignos y benignos de masas tumorales en mamas. Esta base de datos ofrece, además, segmentaciones de ROIs (*Region Of Interest*) detalladas por profesionales, datos seleccionados y supervisados por radiólogos altamente cualificados, entre muchas otras herramientas. Además, esta base de datos ofrece las máscaras de las masas tumorales presentes en las imágenes, precisamente de aquí se extraen las características radiómicas.

2 ESTADO DEL ARTE

2.1. Inteligencia artificial

Lo que conocemos como IA está basada en una tecnología informática, la cual pretende dotar a las máquinas de la capacidad de aprender, haciendo uso de enormes cantidades de datos, con el fin de simular el funcionamiento de la inteligencia humana.

Esta tecnología posee una característica muy importante, la adaptabilidad frente a los distintos campos de la ciencia. En nuestro caso, vamos a centrarnos en su aplicación en el campo de la medicina, ya que, a pesar de ser una tecnología desarrollada en los años 50 del pasado siglo, hasta hace pocos años no ha adquirido verdaderamente su peso, como lo fue en los años 2000, con la aparición del “Deep Learning” (aprendizaje profundo, en español) [6].

Su desarrollo se remonta al año 1943, cuando la unión del matemático estadounidense Walter Pitts y el neurólogo estadounidense, Warren McCulloch, dio lugar al primer modelo de red neuronal en la historia [7].

Años después, en 1956, fue cuando se dio a conocer el nombre de “Inteligencia Artificial”, cuando un grupo de científicos estadounidenses formado por John McCarthy, Marvin Minsky, Nat Rochester y Claude Shannon durante una conferencia lo mencionaron por primera vez en la historia [8].

Su potencial está basado, principalmente, en la capacidad de predicción ante diversos puntos claves en nuestras vidas, como puede ser desde la ayuda en la educación de las próximas generaciones hasta el desarrollo de técnicas para el desarrollo de enfermedades. Esta capacidad de predicción está fundamentada en modelos analíticos relacionales y lógicos [9].

Un punto clave dentro de la IA es lo que conocemos como “Machine Learning” (aprendizaje automático, en español), lo cual dota a la máquina de la capacidad de “aprender” sin estar programada para un uso en concreto, aportando una infinidad de beneficios tecnológicos. Entre la infinidad de avances desarrollados gracias al *machine learning*, uno de los principales fue la evolución hacia las conocidas RNA (redes neuronales artificiales), en las que cada neurona interviene a fin de simular el comportamiento del cerebro humano. Sin embargo, estas redes están evolucionando hacia las que conocemos ahora como redes neuronales cada vez más profundas, las cuales se están basando en el conocido “Deep Learning”.

2.1.1 Machine Learning

El campo que abarca el *machine learning* va desde la clasificación de imágenes, la extracción de características, la predicción de datos, hasta muchos otros más, ya que es imposible mencionar todas sus aplicaciones.

Su funcionalidad está basada en el uso de matrices numéricas, vectores, tensores y cálculos estadísticos y probabilísticos [10].

Al no ser un modelo programado y definido para una finalidad en concreto, y basarse en modelos de aprendizaje de datos, se distinguen distintos tipos de modelos según su tipo de aprendizaje [11]:

- Aprendizaje supervisado: los datos que usan son previamente etiquetados, para lo que se requiere de intervención humana. El conjunto de datos que se le introduce al sistema se divide en dos: grupo de entrenamiento y grupo de prueba. Entre los distintos algoritmos que implementan este modelo de aprendizaje podemos destacar:
 - Decision tree: el cual se basa en un grafo donde se representan las distintas opciones y los resultados, se llama así por su forma de árbol.
 - Naive Bayes: basado en el teorema de Bayes, se suele usar para *clustering*.

- SVM (Support Vector Machine): se basa en una clasificación utilizando lo que se conoce como “truco del kernel” (espaciar las clases, maximizando las distancias entre clases y el margen, consiguiendo minimizar el error a la hora de la clasificación).
- Aprendizaje no supervisado: se basa, al igual que el aprendizaje supervisado, en un modelo entrenado, pero, a diferencia del anterior, éste no consta de datos previamente etiquetados, por lo que deduce patrones y salidas a partir de lo aprendido. El principal algoritmo basado en este tipo de aprendizaje es el conocido como:
 - K-Means: basado en la creación de “K” número de centroides, uno para cada *clúster*, que han de ubicarse estratégicamente para que estén lo más lejos posible unos de otros, produciendo así el mejor resultado de clasificación viable.
- Aprendizaje semisupervisado: se trata de una combinación de los dos métodos anteriores. Dentro de los algoritmos que lo implementan, podemos destacar:
 - TSVM (Transductive SVM): se usa para el etiquetado de datos, de manera que maximice la distancia entre el grupo de datos etiquetados y los no etiquetados.
- Aprendizaje de refuerzo: Se basa en un modelo de aprendizaje en el que, en función del resultado que genere, se aplica un refuerzo para mejorar su funcionamiento.

Esta tecnología del *Machine Learning* se desarrolla en distintos lenguajes de programación como son R (para el análisis estadístico y la representación de datos), Python, Matlab o Julia [12]. Entre todos estos lenguajes destacan ciertas librerías usadas según el propósito del sistema, enfatizando Scikit-Learn para la clasificación de imágenes o extracción de características, o Statsmodels para el desarrollo de modelos estadísticos.

2.1.2 Deep Learning

Abarca un subconjunto del aprendizaje automático, diseñado para el abordaje de problemas de mayor magnitud de datos, para ello, se hace uso de las conocidas redes neuronales, las cuales existen distintos tipos [13]. Estas redes han demostrado resultados bastante mejorados respecto a aquellos algoritmos pertenecientes al aprendizaje máquina.

Este citado campo de la IA ha tenido su máximo punto de desarrollo en los últimos años. Actualmente, se aplica para temas como el procesamiento del lenguaje o la visión artificial.

Dentro del aprendizaje profundo se distinguen de nuevo distintos modelos de aprendizaje, como lo hacía también el aprendizaje máquina [14]:

- Aprendizaje profundo supervisado: se basa en un modelo con datos previamente etiquetados, y basa su cálculo de salida a partir del ajuste de parámetros de la red como, por ejemplo, los pesos de las conexiones entre neuronas y las funciones de pérdidas. Se aplica este tipo de aprendizaje en redes como las RNN (Recurrent Neural Network, en español redes neuronales recurrentes), CNN (Convolutional Neural Network, en español redes neuronales convolucionales) y las DNN (Deep Neural Network, en español redes neuronales profundas). Frente a otros modelos de aprendizaje, este muestra un mayor rendimiento a la hora del aprendizaje de la red.

- Aprendizaje profundo no supervisado: no hace uso de datos previamente etiquetados. Lo encontramos en las redes GAN (*Generative Adversarial Network*, en español redes generativas adversarias), dentro del campo de la reducción de dimensiones y el agrupamiento; así como también se ha implementado en las conocidas RNN. Comúnmente, este modelo se usa para tareas de agrupación.
- Aprendizaje profundo semisupervisado: en este modelo los datos no están totalmente etiquetados. Frente al modelo supervisado, muestra la capacidad de disminuir el volumen de datos etiquetados. Podemos encontrar su implementación dentro de las RNN o las GAN.
- Aprendizaje profundo por refuerzo: fue desarrollado por la empresa americana Google en el año 2013, con el conocido Google Deep Mind (modelo basado en el refuerzo con el entorno pero, a su vez, depende en gran parte del número de parámetros). No es recomendable en modelos con una gran cantidad de parámetros, ya que puede reducir notablemente la velocidad del aprendizaje.

Dentro del aprendizaje profundo se distinguen distintos tipos de redes según su estructura y funcionalidad [15]:

- Redes neuronales convolucionales (CNN): es el tipo de red más famoso dentro del aprendizaje profundo, especialmente para el uso de aplicaciones de imágenes. Este tipo de red está basado en la imitación de la anatomía neuronal humana. Su principal diferencia respecto a las redes neuronales tradicionales es el hecho de convolucionar matrices para obtener resultados, en lugar de multiplicaciones como en las redes tradicionales.

Se caracteriza por dos atributos principales: la escasez de interacciones (consigue disminuir el tamaño del núcleo del modelo frente al tamaño de la entrada, lo que consigue que la carga de uso de la memoria computacional se reduzca) y la compartición de parámetros (también conocido como pesos vinculados) entre las entradas del sistema.

Realizan tres pasos principales: convoluciones para generar activaciones lineales, ejecución de estas activaciones a través de funciones de activación para generar un mapeo a la salida y, por último, se da la agrupación para modificar así la ubicación de una salida de la red en función de sus salidas cercanas.

La primera red de este tipo que resultó exitosa fue la LeNet [16], diseñada por Yann LeCun en los años 90 del siglo pasado, llegando hasta el desarrollo de la LeNet-5, la cual alcanza una precisión del 99,2 %. Ulteriormente, se desarrollaron otras redes como GoogLeNet, de la empresa Google, o la conocida AlexNet [14].

La arquitectura de estas redes constan de varias capas: capa convolucional (encargada de la definición del núcleo y del cálculo convolucional), capa de agrupación (encargada del submuestreo de los mapas de características generados como resultado de las convoluciones), capa de activación (encargada de las funciones de activación, la más conocida de las funciones de activación es la ReLU [17]) y, por último, la capa totalmente conectada (típicamente es la capa final de la arquitectura CNN, encargada de la conexión de cada neurona con todas las neuronas de la capa anterior).

En la capa de salida de la red se implementan distintos tipos de funciones para el cálculo de las pérdidas.

- Redes neuronales recurrentes (RNN): se caracterizan por su tratamiento con datos secuenciales y por la compartición de parámetros. La red se organiza en bucles en lugar de capas, como ocurre en las CNN [18]. LSTM (*Long Short-Term Memory* en español memoria a corto y largo plazo) es el tipo de red RNN más común. Fueron introducidas en el año 1997 [19].
- Redes generativas adversarias (GAN): en este tipo de redes los datos de las entradas se combinan con predictores estadísticos para generar las salidas. Se caracterizan por su facilidad de entrenamiento [18]. Fueron introducidas por primera vez en el año 2014 y están basadas en el modelo de aprendizaje no supervisado. Efectúan principalmente 3 pasos: generación de números aleatorios, discriminación de los

números devolviendo probabilidades y, por último, retroalimentación para mejorar el rendimiento del modelo.

Por último, dentro del *Deep Learning* existe una técnica de aprendizaje profundo llamada *Transfer Learning* (en español, aprendizaje por transferencia), la cual se caracteriza por permitir la reutilización de un modelo entrenado para una tarea en concreto, en otra aplicación relacionada. Herramientas como TensorFlow permiten este tipo de aprendizaje [20].

2.2. Radiómica

Según Kumar et al. [21], la radiómica está centrada en la extracción de características cuantitativas procedentes de las imágenes médicas obtenidas por TAC (Tomografía Axial Computarizada), PET (Tomografía por Emisión de Positrones) o RM (Resonancia Magnética).

Existen dos tipos diferenciados de radiómica [22] [23]:

- Basada en características: previo al paso de la extracción de características, los datos a explotar requieren de una etapa de preprocesado para ofrecer una buena reproducibilidad en los resultados, para ello, se realiza una etapa de interpolación en la dimensión de los vóxeles de la imagen y una etapa de discretización de niveles de gris, garantizando una mejora en la comparación frente a distintos casos de escáneres, pacientes, etc. Además, dentro de la etapa de preprocesamiento, podemos encontrar etapas como el suavizado de la imagen o la reducción de ruido.

Una vez preparados los datos, tiene lugar la etapa de extracción, lo que permite extraer un gran número de características, todas ellas matemáticamente definidas y recogidas en distintos grupos:

- Características de forma: describen relaciones geométricas dentro de las regiones segmentadas, existiendo 2 subgrupos, donde se dividen por características 2D y 3D. Para nuestro proyecto solo son necesarias características bidimensionales debido a nuestras imágenes de entrada, por lo que podemos destacar dentro de todas las existentes:
 - Elongación.
 - Longitud del eje mayor y menor.
 - Perímetro.
 - Relación perímetro-superficie.
 - Esfericidad.
- Características de primer orden: están basadas en la distribución de los niveles de intensidad de la imagen. Permiten hallar distintos parámetros, entre los que podemos destacar:
 - Percentil 10 y 90.
 - Energía
 - Curtosis.
 - Desviación media absoluta.
 - Mediana.
 - Máximo y mínimo.
 - Uniformidad.
 - Varianza.
 - Asimetría.
- Características de segundo orden: representan la información textural de la imagen, la cual se define mediante distintos tipos de matrices, donde encontramos:

- GLRLM (Matriz de longitud de ejecución de niveles de gris): estas matrices se especializan en la cuantificación de secuencias de píxeles consecutivos que tienen igual nivel de gris dentro de la imagen. Podemos destacar las siguientes características:
 - No uniformidad de nivel de gris.
 - No uniformidad de nivel de gris normalizada.
 - Varianza de nivel de gris.
 - Énfasis a largo plazo.
 - Énfasis en la ejecución de niveles de grises bajos.
 - Énfasis en la ejecución de niveles de grises altos.

- NGTDM (Matriz de diferencia de niveles de gris vecinos): describen las diferencias entre los tonos de nivel de grises vecinos. Dentro de las características que permiten obtener, podemos destacar:
 - Ocupación.
 - Contraste.
 - Aspereza.
 - Fuerza.

- GLCM (Matriz de coocurrencia de niveles de gris): mide la frecuencia en la ocurrencia de los niveles de gris dentro de la imagen. Permiten obtener las siguientes características, entre otras:
 - Autocorrelación.
 - Diferencia de entropía.
 - Contraste.
 - Correlación.
 - Diferencia de varianza.

- GLSZM (Matriz de zonas de tamaño de niveles de gris): cuantifican las zonas de un mismo nivel de gris dentro de la imagen. Podemos destacar las siguientes características que nos proporcionan:
 - Énfasis en área pequeña.
 - Zona de entropía.
 - Zona de varianza.
 - Tamaño de zona no uniforme.
 - Énfasis en área grande.

- GLDM (Matriz de dependencia de niveles de gris): miden las dependencias existentes entre un píxel y sus vecinos, permitiendo obtener estas características, entre muchas otras:
 - Dependencia en entropía.
 - Dependencia en varianza.
 - Dependencia en no uniformidad.
 - Énfasis de baja dependencia.
 - Énfasis de alta dependencia.

- Basada en el aprendizaje profundo: se basa en el uso de redes neuronales para generar en salida aquellas características más representativas según los datos introducidos a la red. El uso de estas redes, obvia la necesidad de segmentación previa a la extracción.

Un problema que podemos encontrar en este grupo dentro de la radiómica es el sobreajuste del sistema, aplicándose, frente a ello, funciones de regularización y abandono.

A pesar de ser una tecnología muy prometedora en el campo de la investigación, aún necesita del apoyo humano, ya que es una herramienta complementaria para el diagnóstico de las imágenes. Sin embargo, ofrece una gran ayuda en la mayoría de los casos a todo aquel profesional que se dedique al campo del diagnóstico sanitario [24].

Además, su uso conjunto con herramientas clasificadoras como SVM es de gran aplicación en el campo de la detección oncológica.

Se ha comprobado en la mayoría de los casos que la radiómica basada en aprendizaje profundo ofrece una mayor reproducibilidad frente a aquella basada en el cálculo matemático de características, ya que esta última depende en mayor parte de la técnica de captación de la imagen. No obstante, aquella basada en las redes neuronales ofrece mayor generalización debido al uso de este tipo de redes.

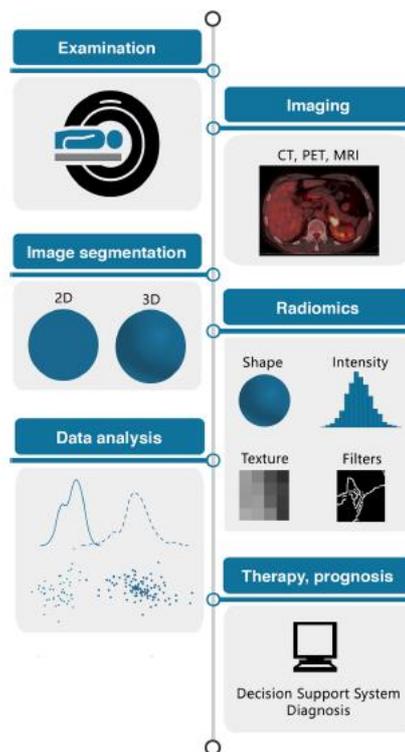


Ilustración 1. Flujo de trabajo de la radiómica [25]

En el caso concreto del cáncer de mama se obtienen biomarcadores procedentes de la imagen como las microcalcificaciones, la densidad mamaria o la genómica del tumor [26].

Entre las librerías más usadas en radiómica se encuentran PyRadiomics [23], LifeX [27], QIFE [28] o MaZda [29].

3 METODOLOGÍA

3.1. Extracción de características

La extracción de características radiómicas se realiza para obtener información, no perceptible a simple vista, de las imágenes médicas, lo que conlleva mejoras en el diagnóstico de las distintas enfermedades, debido a que, en conjunto con la información clínica del especialista, se forma un diagnóstico bastante más preciso y completo. Además, uno de sus principales usos es el desarrollo de patrones predictivos, ya que se pueden extraer indicadores que, bajo el uso de las redes neuronales, pueden desarrollar patrones de comportamiento, desarrollando la capacidad de predecir el curso de la enfermedad.

Para la extracción de características se ha optado por el lenguaje de programación “Python”, programado en el IDE (*Integrated Development Environment*, en español entorno de desarrollo integrado) “Visual Studio Code”, debido a que es un conjunto que implica sencillez a la hora de programar, aportando a su vez un gran catálogo de herramientas.

Principalmente, se ha empleado la librería “PyRadiomics” [23], la cual está orientada al uso de la investigación médica. Conlleva una infinidad de funciones matemáticas que permiten el cálculo de características de forma, niveles de grises y texturas. Adicionalmente, ofrece la ventaja de poder trabajar con imágenes procedentes de distintas técnicas de captación, como resonancias magnéticas, TAC (tomografía axial computarizada), etc.

El funcionamiento de este código se basa en un conjunto de 40 imágenes en formato DICOM [30] procedentes de mamografías, agrupadas en 2 subconjuntos, previamente etiquetados, de 20 imágenes cada una, distinguiéndose estos dos grupos en malignas y benignas, según muestren o no, masas tumorales benignas o malignas. Las imágenes proceden de la base de datos pública CBIS-DDSM [5].

Además de importar la librería principal para este código, se han importado otras como: “PyDicom” [31], para la lectura y tratamiento de imágenes formato DICOM; librería “Os” [32], para el manejo del sistema operativo; “Scipy.io” [33], para el almacenamiento de los resultados en archivos .mat y, por último, se ha usado también la librería “Pandas” [34], para el manejo de estructuras de datos.

```
import os
from radiomics import featureextractor
import pydicom as dicom
from scipy.io import savemat
import pandas as pd
```

Ilustración 2. Importación de librerías

Para la extracción de características, se han ido extrayendo grupo a grupo (FIRST ORDER, SHAPE, GLRLM, GLSZM, GLCM, GLDM Y NGTDM) y forzando al extractor a realizar una extracción en 2D, para así tener mejor manejo y sencillez en su posterior análisis estadístico. Así nos aseguramos que la extracción es correcta debido a que el factor 2D es necesario tenerlo en cuenta cuando partimos de imágenes bidimensionales como las mamografías. Esta etapa del código se ha realizado con la implementación de funciones de la librería PyRadiomics.

```
malicious_extractor=featureextractor.RadiomicsFeatureExtractor(force2D=True)
malicious_extractor.disableAllFeatures()
malicious_extractor.enableFeatureClassByName('gldm')
```

Ilustración 3. Extractor de radiomics (para el caso del grupo 'gldm')

Una vez activado y configurado el extractor de características, se realiza la lectura de las imágenes DICOM (lectura de imágenes benignas y malignas), lo que genera una lista del conjunto de todas estas características con la ayuda del método "append" de Python [35].

Cuando ya tenemos extraídas todas las características y generadas las listas con sus valores para cada una de las imágenes, el último paso de la extracción es su almacenamiento en un archivo .mat. En este momento se hace uso de la librería Pandas, convirtiendo las listas en diccionarios para facilitar su guardado con el método "savemat".

```
features=pd.DataFrame(list)
features_diccionario=features.to_dict(orient='list')
savemat('ngtdm_features.mat',features_diccionario)
```

Ilustración 4. Manipulación de lista y almacenado en archivo ".mat"

En este punto, ya tenemos generados los distintos archivos .mat, para que, a continuación, con la ayuda del programa Matlab, se realice su análisis estadístico, lo cual resulta verdaderamente interesante debido a que, de la infinidad de características obtenidas, solo una pequeña cantidad de ellas son las que realmente se convierten en parámetros indicadores y útiles para la detección de la enfermedad.

Con todo este código se consigue automatizar la extracción de las características de este conjunto de 40 imágenes de cáncer de mamas. Estos datos se podrían destinar para su uso en la investigación médica de patrones característicos o para la detección de la enfermedad.

Además de las características radiómicas, existen otras librerías que permiten la extracción de características a partir de imágenes médicas, como la librería "Scikit-Image" de Python [36], conocida como "Skimage".

La citada librería está orientada al procesamiento de imágenes, permitiendo la extracción de características de forma para nuestro caso en concreto. Esta ayuda es proporcionada por el módulo conocido como "Measure", el cual realiza una extracción basándose en regiones conectadas a partir de imágenes totalmente etiquetadas. Por ello, antes de realizar el cálculo de las características, hay que realizar la extracción de las regiones conectadas y etiquetadas.

```
I = np.array(image.pixel_array)
I_mask = np.array(mask.pixel_array)
I_mask = label(I_mask)
regions = regionprops(I_mask, intensity_image=I)
```

Ilustración 5. Extracción de regiones

A continuación, una vez se obtienen las regiones consideradas, se realiza la extracción de las siguientes características para cada una de ellas:

- Área (representa el área de la región).
- Perímetro crofton (representa el perímetro de la región).
- Compacidad (mide cómo de compacta es la región)

$$\text{Compacidad} = \frac{\text{Área}}{\text{Perímetro}^2}$$

- Orientación (mide la orientación de la región calculando el ángulo entre el primer eje y el eje mayor de la región).
- Intensidad media (representa el valor de la intensidad media comprendida en la región).
- Excentricidad (está referida a la relación existente entre la distancia focal y la longitud del eje mayor de la región).
- Solidez (mide la solidez de la región relacionando los píxeles de la región con los píxeles de la imagen).
- Momentos:
 - Momento 0 0 (relativo al cálculo del área).
 - Momento 1 1 (relativo al cálculo del centroide, basándose en la correlación entre píxeles).
 - Momento 2 2 (relativo al cálculo del centroide, basándose en la difusión de la intensidad en el espacio x-y).
 - Momento 3 3 (relativo al cálculo del centroide, basándose en la información que proporcionan las asimetrías y las curvaturas en la distribución de la intensidad).
 - Momento 1 3 (aporta información sobre la distribución de la intensidad y también sobre su variación espacial).

```
for prop in regions:
    a=prop.area
    p=prop.perimeter_crofton
    d=a/p**2
    shape_features[i,0]=d
    shape_features[i,1]=prop.orientation
    shape_features[i,2]=prop.mean_intensity
    shape_features[i,3]=prop.eccentricity
    shape_features[i,4]=prop.solidity
    moments=prop.weighted_moments_central
    shape_features[i,5]=moments[0,0]
    shape_features[i,6]=moments[1,1]
    shape_features[i,7]=moments[2,2]
    shape_features[i,8]=moments[3,3]
    shape_features[i,9]=moments[1,3]
i=i+1
```

Ilustración 6. Propiedades extraídas de las regiones

3.2. Selección de características

Una vez extraídas las características radiómicas del conjunto de imágenes, la siguiente etapa en el flujo de trabajo es la selección de estas características, ya que todas ellas no serán determinantes para la clasificación de las imágenes en malignas o benignas.

Para ello, se realiza en primer lugar un cribado estadístico basado en la teoría del P-Value, para ver así la significancia de las características. Este valor estadístico se define como la probabilidad de que un valor calculado sea posible, dando por cierta la hipótesis nula; si se obtiene un resultado por debajo del umbral establecido, se intuye que estamos ante un caso estadísticamente significativo debido a que es bastante improbable que el resultado se haya obtenido por azar [37].

En nuestro caso, partiendo de dos conjuntos de datos en los que recogemos las características extraídas para imágenes de masas tumorales benignas y malignas, tenemos la hipótesis nula de que ambos conjuntos de datos pertenezcan a la misma distribución estadística. En caso de ser así, y de obtener muchos pares de conjuntos de datos, las restas de las medias de cada par de datos deberían pertenecer a otra distribución, la cual tendría media cero.

Además, en función del objetivo que se pretenda estudiar con las muestras extraídas y del tamaño muestral que se tenga, se utilizará una prueba estadística u otra, entre las que destacan:

- t de Student [38]: muy útil para comparar las medias de dos muestras de datos, dando como resultado, en caso de quedar por debajo del umbral, que ambas muestras son significativamente diferentes entre ambas. Esta prueba se suele aplicar para un volumen de muestras pequeño y distribuciones normales en el conjunto de muestras. En Matlab se usaría el comando “ttest2” para su cálculo. Su fórmula es la siguiente:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Donde “t” representa el valor estadístico que buscamos, “ \bar{x} ” representa la media muestral, “ μ ” es la media de la población, “s” es la desviación estándar de la muestra y “n” el tamaño muestral [39].

- Chi Cuadrado [40]: se usa para datos. Es una prueba muy apta para variables categóricas. Un punto clave de este estadístico es el conocimiento del tamaño de la muestra a evaluar, ya que esta prueba es bastante sensible a este aspecto a la hora de generar sus resultados. Su fórmula es la siguiente:

$$x^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k}$$

Donde “ x^2 ” es el estadístico para calcular, “ O_k ” es la frecuencia observada y “ E_k ” es la frecuencia esperada teóricamente [41].

- Regresión lineal [42]: se usa para estudiar el comportamiento de cómo cambia una variable en función de otra. Su fórmula es la siguiente:

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Donde “Y” es el estadístico para calcular, “ β_0 ” y “ β_1 ” son los coeficientes de regresión, “x” es la variable independiente y “ ε ” es un término relacionado al error [43].

No todo el mundo es partidario de este cribado estadístico, sin embargo, hoy en día se sigue aplicando en la investigación biomédica. Un tema algo controvertido siempre ha sido la elección de este umbral, el cual históricamente ha sido 0.05; pero en 2018, el economista americano Daniel J. Benjamin, propuso reducirlo de 0.05 a 0.005 para aumentar la precisión en los resultados. Actualmente, este umbral no es generalizable, depende del conjunto de datos al que se esté refiriendo [44].

En nuestro caso, para cada característica extraída hemos procedido a calcular este p-value entre las imágenes benignas e imágenes malignas con la ayuda del comando de Matlab “ttest2” [45], usando la prueba estadística t de Student.

```
[h, p_value]=ttest2(benignas,malignas);
```

Ilustración 7. Comando de Matlab “ttest2”

Este cribado se realiza bajo un determinado umbral. Para nuestro caso, lo probamos con tres umbrales distintos: 0.1, 0.05, 0.01; para así evaluar el funcionamiento del sistema bajo distintos umbrales específicos.

Con todas las características extraídas y cribadas, puede darse igualmente la obtención de un volumen muy grande que acepte este criterio, lo que conllevaría a la necesidad de eliminar aquellas características altamente correlacionadas entre sí. Para ello, se necesitaría un código en Matlab en el cual se calcule la matriz de correlación de las características, con el comando `corrcoef` [46], y así todas aquellas que presenten un alto nivel de correlación con alguna otra, lo que llevaría a reducir el volumen de características altamente significativas.

Finalmente, una vez obtenido el subconjunto de características más discriminatorias, la última etapa sería la clasificación de estas características, la cual se puede realizar usando algún clasificador como SVM, o haciendo uso de las redes neuronales.

3.3. Clasificación de características

Finalmente, una vez extraídas todas aquellas características de las imágenes que tenemos como entrada y seleccionadas aquellas más representativas bajo el criterio del cribado estadístico, y el criterio opcional de eliminación por alta correlación, pasamos a la etapa final del proyecto, la clasificación de las imágenes en función de las características seleccionadas.

Este paso es crucial en el proceso de análisis, ya que nos permite clasificar entre benignas o malignas, comprobando así si las características obtenidas son realmente significativas para la detección de esta patología.

La clasificación se puede realizar a través de muchos métodos diferentes como pueden ser tanto el uso de redes neuronales como el uso de clasificadores prediseñados. En nuestro caso, hemos optado por el uso del conocido clasificador SVM con ayuda de correlación cruzada.

Este clasificador está basado en la técnica de clasificación a partir de clases agrupadas y distanciadas en el espacio [47]. Se tienen espacios de características mapeados, y el SVM trata de encontrar el límite que maximice el margen entre las clases, este límite recibe el nombre de “hiperplano”. Maximizar el margen es un problema de programación cuadrática, el cual, con la aplicación de operadores lagrangianos, se puede resolver. Al igual que el mapeo de características se puede resolver con la aplicación de funciones kernels [48].

Dentro de estos clasificadores se distinguen distintos casos:

- Caso linealmente separable: ante este caso, la búsqueda del hiperplano óptimo se realiza mapeando el espacio de característica a una dimensión mayor.

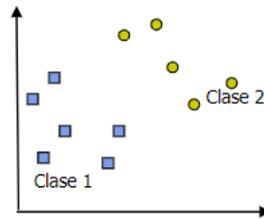


Ilustración 8. Caso linealmente separable [47]

- Caso no linealmente separable: ante conjuntos de datos no linealmente separables, el problema de búsqueda de un hiperplano óptimo se resuelve por programación cuadrática y aplicando lagrangianos.

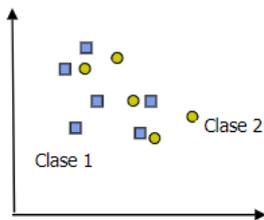


Ilustración 9. Caso linealmente no separable [47]

En este proyecto, hemos hecho uso de la máquina SVM con la técnica de validación cruzada [49], así se ha dividido el conjunto de datos en subconjuntos, llamados “*folds*”, dedicando dentro de cada *fold*, un conjunto destinado a pruebas y el otro destinado al entrenamiento de la máquina. Es un proceso reiterativo, a fin de hacer “aprender” a la máquina. Hemos optado, debido al bajo volumen de características seleccionadas para clasificar, por usar 20 *folds*, realizando así 20 particiones en el conjunto de datos.

El SVM ha sido crucial en la investigación biomédica en el campo de la clasificación de imágenes oncológicas [50]. Los primeros estudios de su aplicación se remontan a inicios del año 2000, cuando se aplicó para la clasificación de distintos tipos de leucemia, aportando datos numéricos sobre la información genética, lo que resultó dar un alto porcentaje de acierto, con 38 pacientes en entrenamiento y 34 en prueba. Posteriormente, se trabajó en la mejora del método para esta aplicación, consiguiendo llevarlo a un porcentaje de errores nulo.

Además de para la leucemia, se ha aplicado en otros tipos de cánceres como el de colón o el de mama [50], el cual es de mayor interés para nuestro estudio.

4 EVALUACIÓN DEL ALGORITMO

Tras extraer, seleccionar y clasificar las características procedentes de las imágenes mamográficas, se realiza el análisis de aquellos parámetros que evalúan la eficacia del algoritmo diseñado. Esta evaluación parte del cálculo de una serie de parámetros:

- TP (*True Positive*, en español verdadero positivo): referido a cuando la predicción y los valores reales son positivos.
- FP (*False Positive*, en español falso positivo): referido a cuando la predicción toma un valor positivo, mientras que el valor real no lo es.
- TN (*True Negative*, en español verdadero negativo): referido a cuando la predicción y los valores reales son negativos.
- FN (*False Negative*, en español falso positivo): referido a cuando la predicción toma un valor negativo, mientras que el valor real no lo es.

Estos cuatros parámetros son los que forman la conocida matriz de confusión, la cual es la herramienta que se usa para evaluar la eficacia de los algoritmos basados en aprendizaje supervisado [51]. Esta herramienta trata de evaluar las prestaciones del sistema, comparando los valores reales detectados por profesionales en la materia, los cuales forman la conocida “Ground-Truth”, y los valores predichos por la IA.



VALORES PREDICCIÓN	Verdaderos positivos	Falsos Positivos
	Falsos Negativos	Verdaderos Negativos
	VALORES REALES	

Ilustración 10. Matriz de confusión

Con estos cuatro parámetros ya es posible la evaluación del algoritmo de clasificación con la ayuda de las siguientes fórmulas:

- Sensitivity, en español sensibilidad: porcentaje de pacientes con la enfermedad que han recibido un resultado positivo en cuanto a detección de la enfermedad.

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- Specificity, en español especificidad: porcentaje de pacientes sin la enfermedad que ha recibido un resultado negativo en cuanto a detección de la enfermedad.

$$\text{Spec} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

- PPV (Positive Predictive Value, en español valor predictivo positivo): muestra la probabilidad de que el paciente tenga la enfermedad si el algoritmo dice que es positiva la detección.

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- NPV (Negative Predictive Value, en español valor predictivo negativo): muestra la probabilidad de que el paciente no tenga la enfermedad si el algoritmo dice que es negativa la detección.

$$\text{PPN} = \frac{\text{TN}}{\text{TN} + \text{FN}}$$

5 RESULTADOS

Para finalizar con este proyecto, tras las fases de extracción, selección y posterior clasificación, se muestran los resultados obtenidos, separando estos resultados en dos grandes conjuntos: resultados de las fases de extracción y selección, y resultados de la fase de clasificación.

5.1. Resultados de las fases de extracción y selección

Tras las fases de extracción, hemos extraído un total de 102 características entre los 7 grupos radiómicos existentes. Sin embargo, todas estas características extraídas no son de utilidad para este estudio, por lo que se ha aplicado el criterio del P-Value a todas estas características, obteniendo un total de 7 características que satisfacen el criterio del umbral del 0.1, 4 características que satisfacen el umbral del 0.05; y, por último, ninguna característica que satisfaga el umbral del 0.01, el cual es el más restrictivo de los que hemos aplicado.

Todas las características aplicadas pertenecen al grupo de las características de primer orden, las cuales definen parámetros estadísticos obtenidos a partir de la distribución en los niveles de grises a través de los píxeles, dentro de la imagen.

Las 7 características extraídas y seleccionadas bajo el umbral del 0.1 del P-Value son:

- Percentil 90: mide el estadístico “Percentil 90” [52] a partir del conjunto de datos.
- Energía: mide la magnitud de los valores de los píxeles en una imagen; si se obtiene un valor alto, quiere decir que la suma de los cuadrados de estos valores es alta [23].

$$\text{Energía} = \sum_{i=1}^{N_p} (X(i))^2$$

Siendo “X(i)” el conjunto de muestras de las intensidades.

- Energía total: a diferencia de la característica de energía convencional, esta está destinada al análisis de imágenes en 3D, por lo que para nuestro estudio no parece demasiado significativa.
- Máximo: mide el máximo en el conjunto de datos [23].

$$\text{Máximo} = \max(X)$$

- Desviación absoluta media (MAD, *Mean Absolute Deviation* en inglés): refleja la distancia media de los valores de intensidad respecto al valor medio de la imagen [23].

$$\text{MAD} = \frac{1}{N_p} \sum_{i=1}^{N_p} |X(i) - \bar{X}|$$

- Mediana: mide la intensidad de nivel de gris situado en la mitad del rango dinámico de los valores de los píxeles de la ROI.

- Media de la raíz cuadrada (RMS, *Root Mean Squared* en inglés): refleja la raíz cuadrada media de todos los niveles de intensidad al cuadrado.

$$\text{RMS} = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i))^2}$$

Estas son las características que satisfacen el umbral de 0.1 para el P-Value. Las características que satisfacen el umbral del 0.05 son el percentil 90, el máximo, la desviación absoluta media y el error cuadrático medio y, por último, para el umbral del 0.01, ninguna característica lo satisface.

Toda esta información que nos aportan las características seleccionadas está contenida en el ámbito de la distribución de niveles de intensidad dentro de la imagen la cual nos da información sobre la textura y la distribución de los valores.

En cuanto a las características que aportan información sobre la malignidad en masas tumorales mamarias, como ya se ha mencionado anteriormente, las distorsiones en la morfología mamaria o los aumentos en la densidad del tejido, pueden ser determinantes a la hora de analizar y dictaminar un análisis preciso sobre estos cuerpos.

Para ello, en este trabajo se ha realizado el cálculo de características de forma, a partir de la librería de Radiomics y mediante el módulo “Measure” de la librería Skimage, además de características de textura también con Radiomics.

Las características de primer orden extraídas y posteriormente seleccionadas, muestran unos histogramas, entre muestras benignas y malignas, con la siguiente forma:

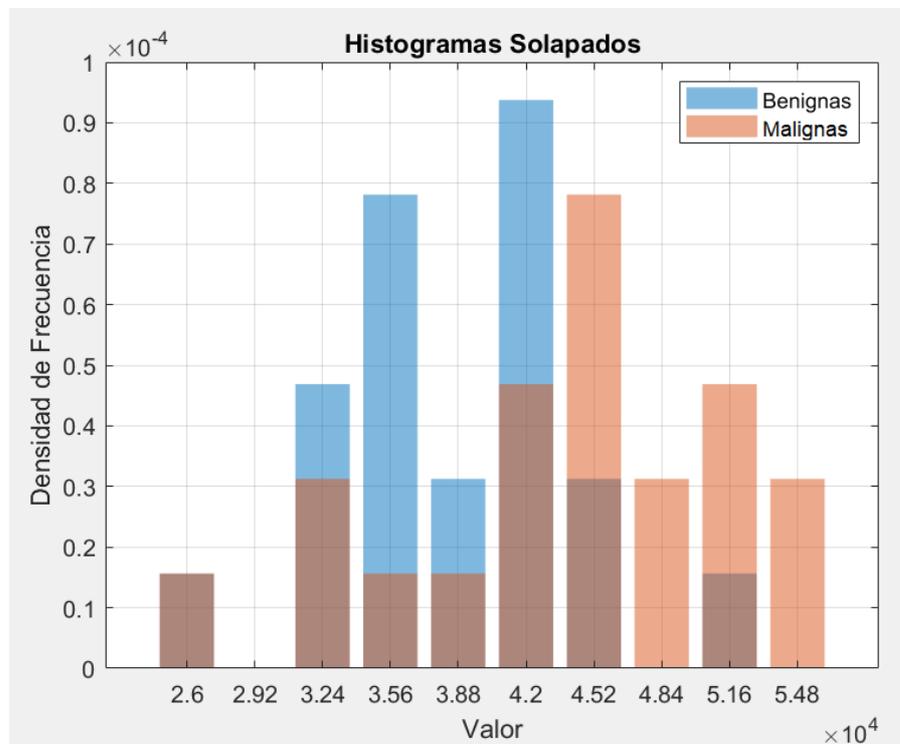


Ilustración 11. Histograma “Percentil 90”

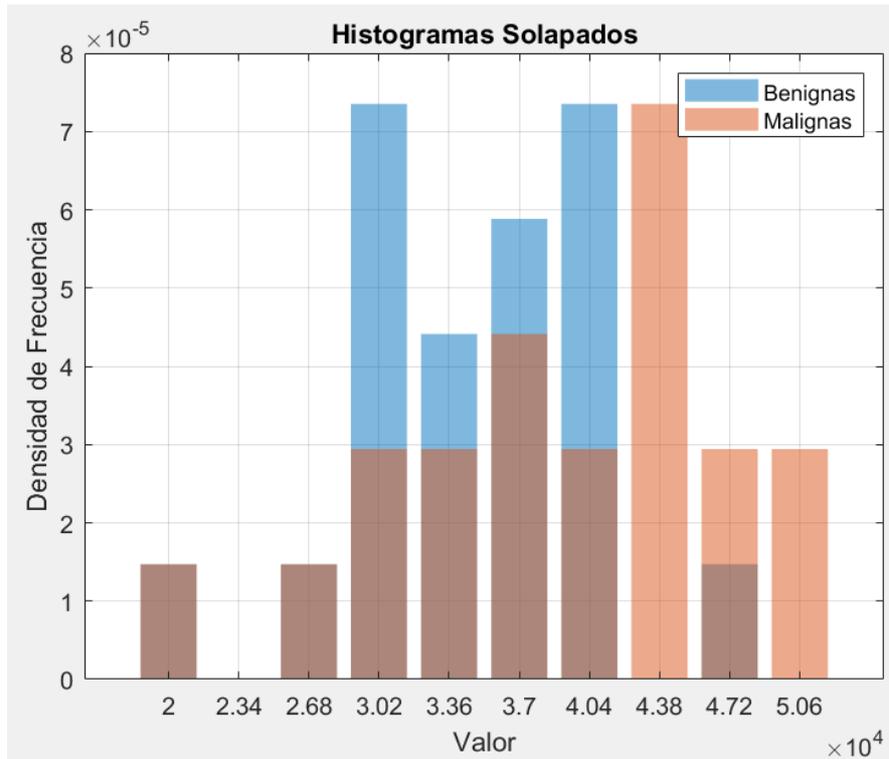


Ilustración 12. Histograma “Mediana”

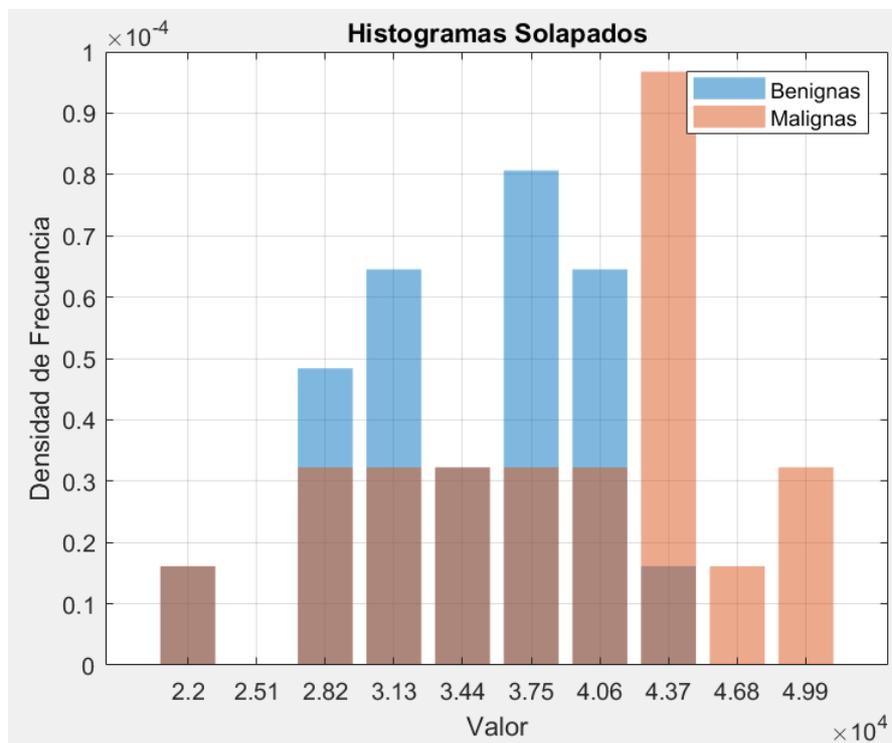


Ilustración 13. Histograma “RMS”

Para analizar estos histogramas, hemos usado el parámetro de medición de la distancia de Bhattacharyya [54], el cual dentro del campo del procesamiento digital de imágenes, mide la similitud entre distribuciones, a partir de sus histogramas; lo que nos daría información sobre si son realmente significativas esas características o no.

Las características de primer orden seleccionadas muestran un amplio rango de distancias:

- Percentil 90: 0.2085.
- Energía y energía total: 0.2863.
- Máximo: 0.5020.
- MAD: 0.0642.
- Mediana: 0.2451.
- RMS: 0.1792.

Vemos en general cómo estas características presentan mayoritariamente distancias en el rango de 0.1 a 0.3, donde observamos 5 características dentro de este rango, una por debajo de este y otra que supera el rango. Por lo regular, aquellas características que están seleccionadas como altamente significativas, deberían presentar un valor alto para esta distancia, debido a la gran diferencia que debería existir entre sus histogramas.

Sin embargo, las características de forma, tanto de *radiomics* como de *measure*, y las de textura, muestran histogramas con el siguiente patrón:

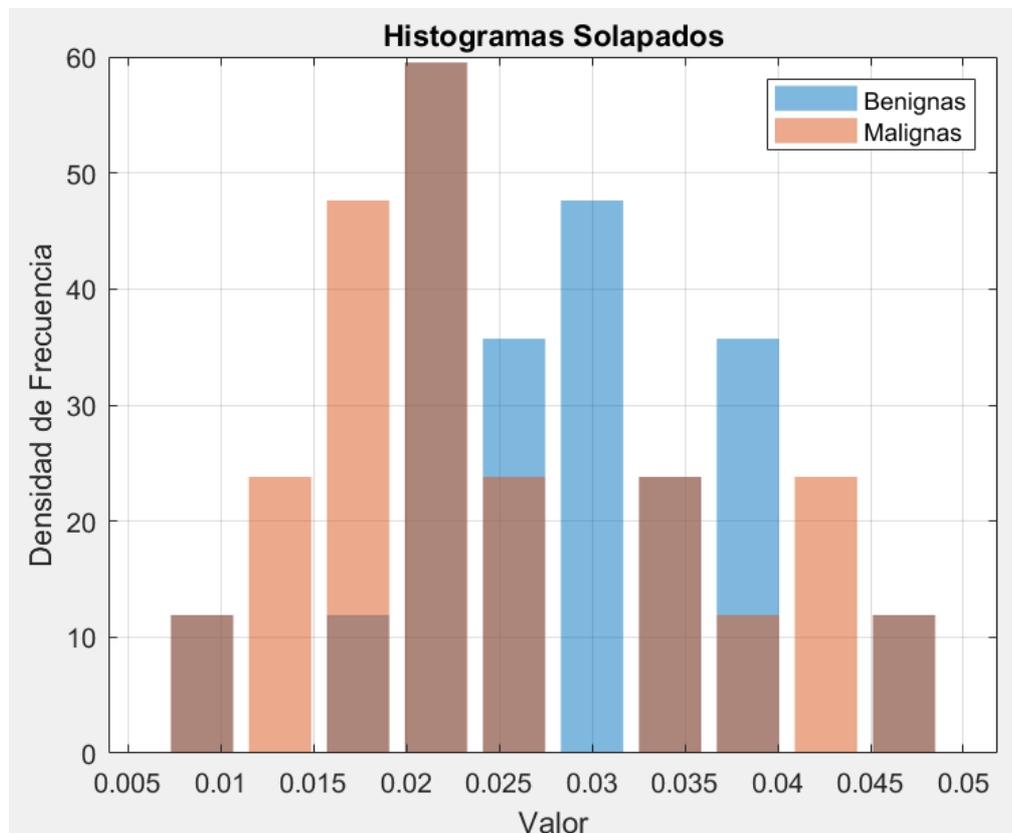


Ilustración 14. Histograma “Ratio Perímetro Superficie”

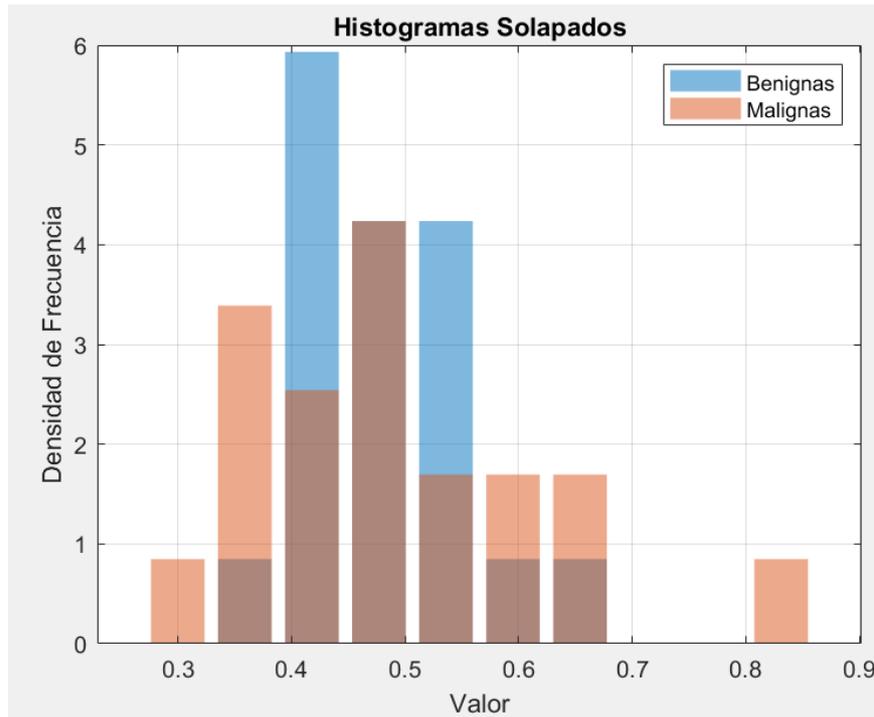


Ilustración 15. Histograma "Esfericidad"

Estos dos últimos histogramas muestran características bidimensionales de forma que, a priori, se esperaba que hubieran sido significativas, ya que nos da información morfológica sobre las masas. Sin embargo, estas características de *radiomics* muestran que de las 9 características que existen para este grupo, 8 siguen manteniendo la tendencia a quedarse dentro del rango de distancias de 0.1 a 0.3 y una se queda por debajo del rango; esto refleja una tendencia parecida con respecto a las características seleccionadas, mostrando como la mayoría de las características caen en el rango de 0.1 a 0.3, de igual forma para las características de primer orden seleccionadas.

A continuación, se muestran histogramas de características bidimensionales de forma obtenidos a partir de "measure" y, por último, histogramas que reflejan aquellas características matriciales de textura.

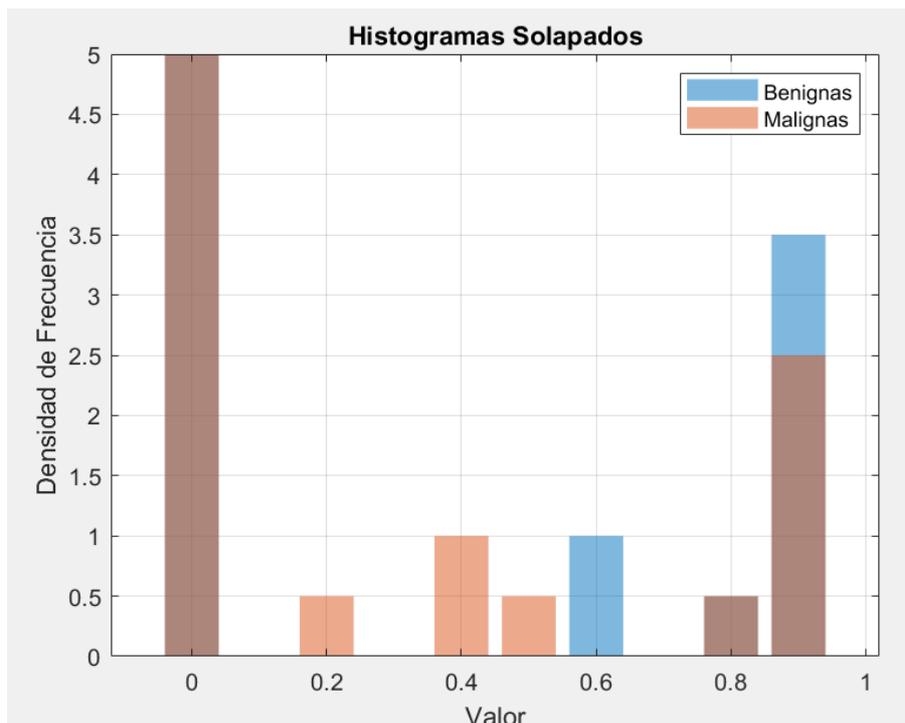


Ilustración 16. Histograma "Excentricidad"

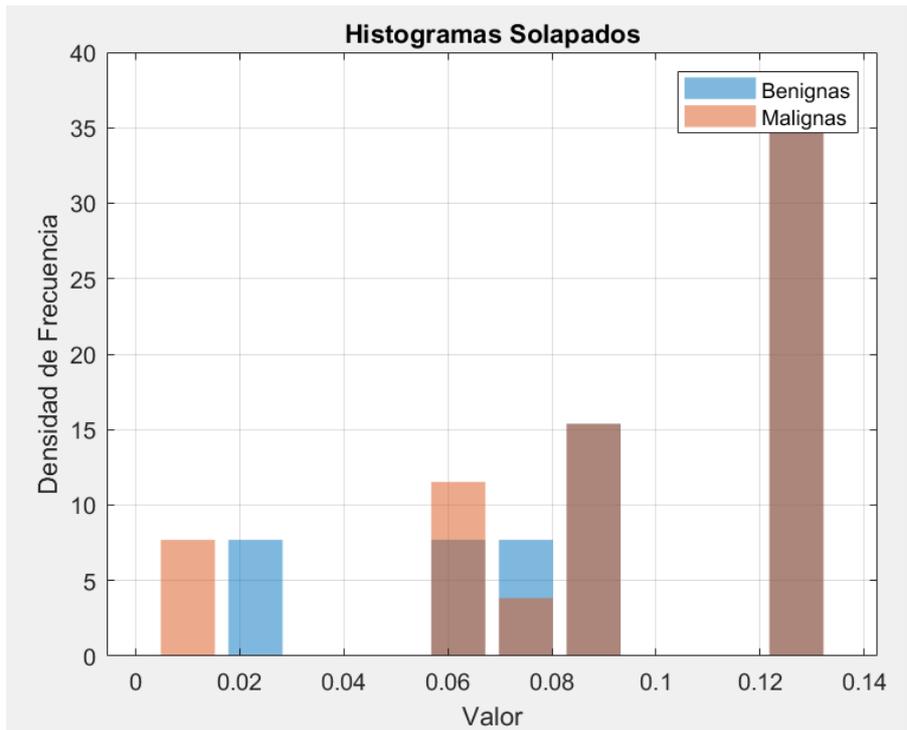


Ilustración 17. Histograma “Compacidad”

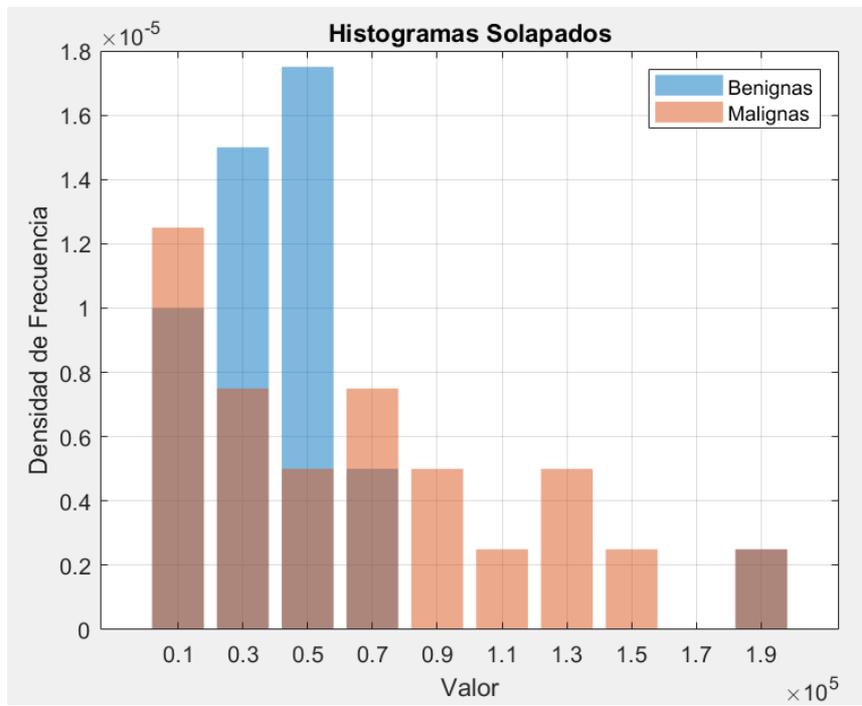


Ilustración 18. Histograma “GLDM No dependencia en uniformidad”

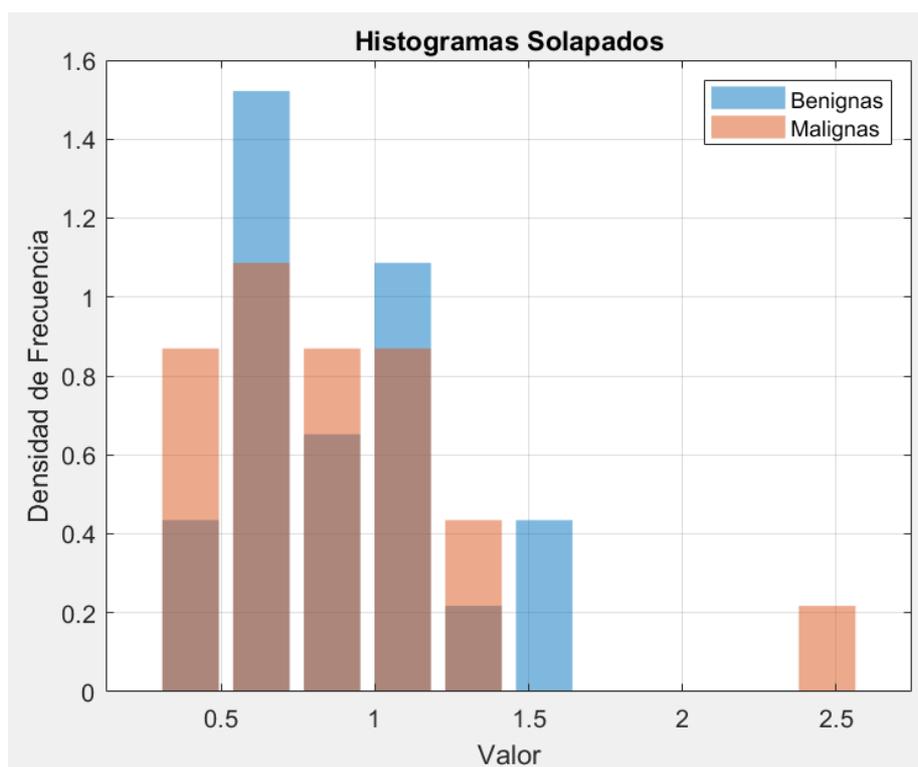


Ilustración 19. Histograma “NGTDM Contraste”

En general, el hecho de que aquellas características seleccionadas deban presentar una distancia alta respecto a aquellas no seleccionadas, no se cumple; la mayoría de las características mantienen una misma tendencia en cuanto a los valores de las distancias, conservando todas ellas la mayoría de sus características dentro del rango de distancias de 0.1 a 0.3.

- Características de textura GLCM: 24 características en total. 16 características dentro del rango de 0.1 a 0.3, 7 por debajo del rango y 1 lo supera.
- Características de textura GLRLM: 16 características en total. 9 características dentro del rango de 0.1 a 0.3, 5 por debajo del rango y 2 lo superan.
- Características de textura GLSZM: 16 características en total. 7 características dentro del rango de 0.1 a 0.3, 7 por debajo del rango y 2 lo superan.
- Características de textura NGTDM: 5 características en total. 5 características dentro del rango de 0.1 a 0.3.

Sin embargo, existen dos grupos de características donde esta tendencia de tener la mayoría de sus características en el rango de distancias de 0.1 a 0.3 no lo cumplen:

- Características de textura GLDM: 14 características en total. 6 características dentro del rango de 0.1 a 0.3, 7 por debajo del umbral y 1 lo supera.
- Características de forma (modulo “measure”): 12 características en total. 1 característica dentro del rango de 0.1 a 0.3 y 11 por debajo del rango.

Por lo tanto, estos dos últimos grupos de características, en concreto las características de forma extraídas con el módulo “measure” de Skimage, obtienen distancias acorde a lo que se esperaba, debido a que presenta una gran población de características dentro del rango de las distancias de Bhattacharyya relativamente bajas.

5.2. Resultados de la fase de clasificación

Por último, tras realizarse la extracción de todas las características a partir del conjunto de 40 imágenes iniciales, continuando con la selección de aquellas que parecen ser significativas, bajo el criterio del P-Value, se aplica la clasificación mediante una máquina SVM, y se obtienen unos resultados, un tanto esperados, ya que aquellas características que, a priori, suenan como más significativas para el análisis de un cuerpo tumoral, no han sido seleccionadas, por lo que no se esperan unos resultados de clasificación con mucha precisión.

Para esta clasificación, se usa un número de *folds*, relativamente alto, ya que se conoce que el sistema de clasificación SVM parte de un conjunto de características relativamente escaso, por lo que se pretende dotar al sistema de robustez y precisión, aumentando este número de pliegues, forzando al sistema a usar la gran mayoría del conjunto de datos para las fases de entrenamiento y de prueba.

Sin embargo, este modelo no muestra unos resultados tan negativos como se podría esperar de antemano, ya que se obtiene un alto valor en la especificidad, lo que asegura una baja tasa de falsos positivos, indicando con alta probabilidad, el diagnóstico acertado para pacientes que no padezcan tumores malignos. Además, presentan un valor predictivo positivo con un valor, relativamente, alto; lo que indica que en caso de predecir a un paciente la presencia de un cuerpo tumoral maligno en su organismo, existen altas probabilidades de no errar en este diagnóstico.

Por otra parte, el modelo carece de precisión en otros aspectos, ya que se muestra una baja sensibilidad, por lo que es altamente probable identificar un cuerpo tumoral y no acertar con su malignidad. Asimismo, este modelo también muestra una gran debilidad con su valor predictivo negativo; en caso de detectar un tumor benigno, es posible que no esté en lo cierto, lo que puede conllevar a la futura detección de cuerpos malignos en el tumor previamente diagnosticado como benigno, lo que no es deseable en una aplicación clínica.

Estas son las fortalezas y debilidades que muestra nuestro sistema, sin embargo, es bastante mejor de lo que se podría esperar, viendo la escasez y rareza en las características de las que partimos para construir el sistema de clasificación.

En general, como se puede observar, este sistema tiende a clasificar las masas tumorales en benignas. Este hecho comúnmente se puede deber al desbalance de datos, el cual se da en conjuntos donde los datos benignos son mayores inicialmente que los datos malignos a analizar. Este es un hecho característico porque en nuestro conjunto de datos no sufrimos ese desbalance, sin embargo, los clasificadores suelen tener un umbral ajustable, el cual determina el punto sobre el que se apoyan para clasificar entre las clases existentes. Ajustar este umbral podría haber sido una posible mejora a la hora de obtener mejores resultados en cuanto a la clasificación del sistema.

A continuación, se muestra una tabla resumen con todos aquellos parámetros útiles para la clasificación.

<u>Verdaderos Positivos</u>	7
<u>Verdaderos Negativos</u>	18
<u>Falsos Positivos</u>	2
<u>Falsos Negativos</u>	13
<u>Sensibilidad</u>	0.35
<u>Especificidad</u>	0.9
<u>Valor Predictivo Positivo (PPV)</u>	0.78
<u>Valor Predictivo Negativo (NPV)</u>	0.58

Tabla 1. Parámetros de clasificación

6 CONCLUSIONES

En este trabajo de fin de grado, hemos explorado el novedoso tema de la aplicación de la radiómica para el análisis automático de mamografías. Nos hemos centrado en el punto de vista de la extracción de las características, la selección de las más significativas y, por último, en la clasificación de las imágenes de las que se parte, para ver cómo de útil es nuestro sistema.

El principal objetivo ha sido el detectar qué características son altamente significativas, para poder diferenciar entre tumores benignos y tumores malignos. A pesar de no obtener los resultados esperados, hemos obtenido otros resultados, que, como mínimo, han sido de gran interés. Partiendo de un enfoque inicial donde se esperaban obtener características de textura y de forma para realizar el sistema de clasificación, hemos pasado a un sistema de clasificación formado al cien por cien por características basadas en niveles estadísticos sobre los niveles de grises que presentan los píxeles de la imagen. Esto indica que las diferencias entre los niveles de grises presentes en las imágenes, para nuestro conjunto de partida, presentan rasgos característicos en la diferenciación de tejidos benignos y malignos. Esto se puede deber a la homogeneidad morfológica entre las distintas imágenes que estamos analizando, o quizás se puede deber al reducido volumen de imágenes de las que partimos.

El hecho de que aquellas características seleccionadas hayan sido de primer orden, puede deberse a muchos aspectos, pero el hecho de que los niveles de grises presentes en los cuerpos tumorales hayan salido como atributo característico para definir la malignidad del propio tumor, puede decirnos que la propia composición del tejido de este cuerpo sea, quizás, un punto para tener en cuenta a la hora de determinar su malignidad o benignidad.

En conclusión, con este trabajo hemos comprobado el alto potencial que muestran también aquellas características de primer orden en el análisis de estos tumores en las mamas.

Este trabajo además de contribuir al avance en el diagnóstico a partir de las imágenes médicas en oncología, contribuye a la necesidad del desarrollo de técnicas, cada vez más precisas y efectivas, para los distintos casos clínicos, ya que, como hemos podido observar en este proyecto, se pueden obtener resultados no esperados y, a su vez, de alto interés para otros casos clínicos.

REFERENCIAS

- [1] V. Uroz-Martínez, I. Moreno-Díaz, A. Sánchez-Gil, M. Ballesta-Ruiz, J. H. Gómez-Gómez, O. Monteagudo-Piqueras, ... y J. Granados-Ortega, "Mortalidad por cáncer de mama en la mujer: evolución en España y la Región de Murcia en el período 1981-2021," 2024.
- [2] E. F. Blasco and M. A. Pérez, "Incidencia y mortalidad en el cáncer de mama. Especial referencia a España," 2024.
- [3] M. Muñoz, M. Fernández, M. Navarro, y S. Sainz de la Cuesta, "Clasificación radiológica y manejo de las lesiones mamarias", *Clínica e Investigación en Ginecología y Obstetricia*, vol. 37, no. 5, pp. 203-213, 2010.
- [4] O. Díaz, A. Rodríguez-Ruiz, A. Gubern-Mérida, R. Martí, and M. Chevalier, "¿Son los sistemas de inteligencia artificial una herramienta útil para los programas de cribado de cáncer de mama?," *Radiología*, vol. 63, no. 3, pp. 236-244, 2021.
- [5] Base de datos. <https://www.cancerimagingarchive.net/collection/cbis-ddsm/> (19/06/2024)
- [6] J. R. Vidal and O. R. Vidal, "Aplicaciones de la inteligencia artificial en la medicina," *Revista Peruana de Investigación en Salud*, vol. 6, no. 3, pp. 131-133, 2022.
- [7] D. Hinstroza Ramírez, "El Machine Learning a través de los tiempos, y los aportes a la humanidad," 2018.
- [8] S. L. Andresen, "John McCarthy: father of AI," *IEEE Intelligent Systems*, vol. 17, no. 5, pp. 84-85, 2002.
- [9] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685-695, 2021.
- [10] G. Rebalá, A. Ravi, and S. Churiwala, "Machine learning definition and basics," in *An Introduction to Machine Learning*, 1st ed., pp. 1-17, 2019.
- [11] B. Mahesh, "Machine learning algorithms - a review," *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381-386, 2020.
- [12] E. M. Rojas, "Machine Learning: análisis de lenguajes de programación y herramientas para desarrollo," *Revista Ibérica de Sistemas e Tecnologías de Informação*, no. E28, pp. 586-599, 2020.
- [13] L. Rouhiainen, *Inteligencia artificial*, Madrid: Alienta Editorial, pp. 20-21, 2018.
- [14] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, and L. Farhan, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, pp. 1-74, 2021.
- [15] M. P. Hosseini, S. Lu, K. Kamaraj, A. Slowikowski, and H. C. Venkatesh, "Deep learning architectures," in *Deep Learning: Concepts and Architectures*, pp. 1-24, 2020.
- [16] Y. Le Cun, L. D. Jackel, B. Boser, J. S. Denker, H. P. Graf, I. Guyon, D. Henderson, R. E. Howard, and W. Hubbard, "Handwritten digit recognition: applications of neural network chips and automatic learning," *IEEE Communications Magazine*, vol. 27, no. 11, pp. 41-46, 1989.
- [17] Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan, and Z. Liu, "Dynamic ReLU," in *European Conference on Computer Vision*, Cham: Springer International Publishing, pp. 351-367, 2020.
- [18] G. Chassagnon, M. Vakalopolou, N. Paragios, and M. P. Revel, "Deep learning: definition and perspectives for thoracic imaging," *European Radiology*, vol. 30, pp. 2021-2030, 2020.
- [19] P. G. Ullán, R. L. González, and M. A. González, "Redes LSTM para la predicción de series temporales," *Avances en Informática y Automática*, p. 131.

- [20] L. Torrey and J. Shavlik, "Transfer learning," in Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI Global, pp. 242-264, 2010.
- [21] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, ... and R. J. Gillies, "Radiomics: the process and the challenges," Magnetic Resonance Imaging, vol. 30, no. 9, pp. 1234-1248, 2012.
- [22] P. Lohmann, K. Bousabarah, M. Hoevens, and H. Treuer, "Radiomics in radiation oncology—basics, methods, and limitations," Strahlentherapie und Onkologie, vol. 196, no. 10, pp. 848-855, 2020.
- [23] PyRadiomics. <https://pyradiomics.readthedocs.io/en/latest/index.html> (20/06/2024)
- [24] Dra. Raquel Pérez López, Instituto oncológico del Vall d'Hebron. <https://seram.es/la-radiomica-eje-fundamental-en-la-investigacion-contr-el-cancer-que-esta-transformando-la-forma-de-abordar-su-lucha-segun-seram/> (21/06/2024).
- [25] Imagen Radiomics. Van Timmeren, J. E., Cester, D., Tanadini-Lang, S., Alkadhi, H., & Baessler, B. (2020). Radiomics in medical imaging—"how-to" guide and critical reflection. Insights into imaging, 11(1), 91.
- [26] Mamografía. <https://cbseram.com/2020/01/15/radiogenomica-del-cancer-de-mama/> (21/06/2024).
- [27] LifeX. <https://www.lifexsoft.org/> (20/06/2024)
- [28] S. Echegaray, S. Bakr, D. L. Rubin, and S. Napel, "Quantitative Image Feature Engine (QIFE): an open-source, modular engine for 3D quantitative feature extraction from volumetric medical images," Journal of Digital Imaging, vol. 31, pp. 403-414, 2018.
- [29] MaZda. <http://www.eletel.p.lodz.pl/programy/mazda/> (20/06/2024)
- [30] Formato DICOM. <https://clinic-cloud.com/blog/formato-dicom-que-es-estandar-imagenes-medicas> (25/06/2024)
- [31] Librería PyDicom. <https://pydicom.github.io/> (23/06/2024)
- [32] Librería OS. <https://docs.python.org/es/3.10/library/os.html> (23/06/2024)
- [33] Scipy.io. <https://docs.scipy.org/doc/scipy/reference/io.html> (23/06/2024)
- [34] Pandas. <https://aprendeconalf.es/docencia/python/manual/pandas/> (23/06/2024)
- [35] Método append, Python. <https://www.ionos.es/digitalguide/paginas-web/desarrollo-web/python-append/> (25/06/2024)
- [36] Scikit-Image. <https://scikit-image.org/> (23/07/2024)
- [37] P-Value. https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1139-76322017000500014 (25/06/2024)
- [38] Prueba T de Student. <https://www.questionpro.com/blog/es/prueba-t-de-student/> (25/06/2024)
- [39] Fórmula T de Student. <https://estadisticaeninvestigacion.wordpress.com/distribucion-t-de-student/> (25/06/2024)
- [40] Prueba Chi Cuadrado. <https://www.questionpro.com/blog/es/prueba-de-chi-cuadrado-de-pearson/> (25/06/2024)
- [41] Fórmula Chi Cuadrado. <https://www.datacamp.com/es/tutorial/chi-square-test-in-spreadsheets> (25/06/2024)
- [42] Prueba de regresión lineal. [https://www.questionpro.com/blog/es/regresion-lineal/#:~:text=La%20regresi%C3%B3n%20lineal%20es%20una,otra%20variable%20\(la%20independiente\).](https://www.questionpro.com/blog/es/regresion-lineal/#:~:text=La%20regresi%C3%B3n%20lineal%20es%20una,otra%20variable%20(la%20independiente).) (25/06/2024)
- [43] Fórmula de regresión lineal. <https://evidenciasenpediatria.es/articulo/7940/regresion-lineal-simple> (25/06/2024)
- [44] G. Di Leo and F. Sardanelli, "Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach," European Radiology Experimental, vol. 4, pp. 1-8, 2020.

- [45] Comando ttest2. <https://es.mathworks.com/help/stats/ttest2.html> (25/06/2024)
- [46] Comando corrcoef. <https://es.mathworks.com/help/matlab/ref/corrcoef.html> (25/06/2024)
- [47] Betancourt, G. A. "Las máquinas de soporte vectorial (SVMs)." *Scientia et Technica*, vol. 1, no. 27, 2005.
- [48] Funciones kernel. <https://scikit-learn.org/stable/modules/svm.html#svm-kernels> (26/06/2024)
- [49] Validación cruzada. https://qu4nt.github.io/sklearn-doc-es/modules/cross_validation.html (09/07/2024)
- [50] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics," *Cancer Genomics & Proteomics*, vol. 15, no. 1, pp. 41-51, 2018.
- [51] Matriz de confusión. <https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/> (09/07/2024)
- [52] Percentil 90. <https://lamenteesmaravillosa.com/percentiles/> (11/07/2024).
- [53] Información tumoral mamas. E. Ortega-García, J. A. Molina, A. C. Ruiz, F. J. González, y J. C. Santiago, "Factores predictivos de malignidad en las microcalcificaciones de mama y lesiones asociadas," *Cirugía Española*, vol. 80, no. 4, pp. 242-247, 2006. (11/07/2024).
- [54] Distancia de Bhattacharyya. https://es.frwiki.wiki/wiki/Distance_de_Bhattacharyya (11/07/2024).

Código extracción con Radiomics.

```
import os
from radiomics import featureextractor
import pydicom as dicom
from scipy.io import savemat
import pandas as pd
import SimpleITK as sitk

# El proceso para la extracción de las características en las imágenes benignas
sólo variaría el cambiar el directorio de "malignas" a "benignas"
main_folder='C:/malignas'
subdirectories = [os.path.join(main_folder, name) for name in
os.listdir(main_folder) if os.path.isdir(os.path.join(main_folder, name))]
malicious_extractor=featureextractor.RadiomicsFeatureExtractor(force2D=True)
malicious_extractor.disableAllFeatures()
# La próxima línea de código varía en función del grupo radiómico que se
pretende extraer
malicious_extractor.enableFeatureClassByName('shape')
list=[]
for sub in subdirectories:
    malicious_image_path = os.path.join(sub, 'original.dcm')
    malicious_mask_path = os.path.join(sub, 'mask.dcm')
    image=dicom.dcmread(malicious_image_path).pixel_array
    mask=dicom.dcmread(malicious_mask_path).pixel_array
    mask //= mask.max()
    image_sitk = sitk.GetImageFromArray(image)
    mask_sitk = sitk.GetImageFromArray(mask)
    result = result = malicious_extractor.execute(image_sitk, mask_sitk)
    shape = {key: value for key, value in result.items() if 'shape2D' in key}
    list.append(shape)

features=pd.DataFrame(list)
features_diccionario=features.to_dict(orient='list')
savemat('shape_features_malignas.mat',features_diccionario)
```

Código extracción con Measure.

```
import os
import pydicom as dicom
from scipy.io import savemat
import pandas as pd
import numpy as np
from skimage.measure import regionprops, label

main_folder_benignas = 'C:/benignas'
subdirectories_benignas = [os.path.join(main_folder_benignas, name) for name in
os.listdir(main_folder_benignas) if
os.path.isdir(os.path.join(main_folder_benignas, name))]
long_benignas=len(subdirectories_benignas)

main_folder_malignas = 'C:/malignas'
subdirectories_malignas = [os.path.join(main_folder_malignas, name) for name in
os.listdir(main_folder_malignas) if
os.path.isdir(os.path.join(main_folder_malignas, name))]
long_malignas=len(subdirectories_malignas)

shape_features= np.zeros((long_benignas+long_malignas,10))

i = 0

for sub in subdirectories_benignas:
    path=os.path.join(main_folder_benignas,sub)
    image_path = os.path.join(path, 'original.dcm')
    mask_path = os.path.join(path, 'mask.dcm')

    image = dicom.dcmread(image_path)
    mask = dicom.dcmread(mask_path)

    I = np.array(image.pixel_array)
    I_mask = np.array(mask.pixel_array)
    I_mask = label(I_mask)
    regions = regionprops(I_mask, intensity_image=I)
    for prop in regions:
        a=prop.area
        p=prop.perimeter_crofton
        d=a/p**2
        shape_features[i,0]=d
        shape_features[i,1]=prop.orientation
        shape_features[i,2]=prop.mean_intensity
        shape_features[i,3]=prop.eccentricity
        shape_features[i,4]=prop.solidity
        moments=prop.weighted_moments_central
        shape_features[i,5]=moments[0,0]
        shape_features[i,6]=moments[1,1]
```

```

        shape_features[i,7]=moments[2,2]
        shape_features[i,8]=moments[3,3]
        shape_features[i,9]=moments[1,3]
    i=i+1

for sub in subdirectories_malignas:
    image_path = os.path.join(sub, 'original.dcm')
    mask_path = os.path.join(sub, 'mask.dcm')

    image = dicom.dcmread(image_path)
    mask = dicom.dcmread(mask_path)

    I = np.array(image.pixel_array)
    I_mask = np.array(mask.pixel_array)
    I_mask = label(I_mask)
    regions = regionprops(I_mask, intensity_image=I)
    for prop in regions:
        a=prop.area
        p=prop.perimeter_crofton
        d=a/p**2
        shape_features[i,0]=d
        shape_features[i,1]=prop.orientation
        shape_features[i,2]=prop.mean_intensity
        shape_features[i,3]=prop.eccentricity
        shape_features[i,4]=prop.solidity
        moments=prop.weighted_moments_central
        shape_features[i,5]=moments[0,0]
        shape_features[i,6]=moments[1,1]
        shape_features[i,7]=moments[2,2]
        shape_features[i,8]=moments[3,3]
        shape_features[i,9]=moments[1,3]
    i=i+1

features_n = np.zeros((long_benignas+long_malignas,10))
for j in range(10):
    max=np.max(shape_features[:,j])
    min=np.min(shape_features[:,j])
    features_n[:,j] = (shape_features[:,j]-min)/(max-min)

features_benignas = features_n[:long_benignas, :]
features_malignas = features_n[long_benignas:, :]
features_benignas_T = features_benignas.T
features_malignas_T = features_malignas.T
savemat('features_benignas.mat', {'features_benignas': shape_features[0:20,:]})
savemat('features_malignas.mat', {'features_malignas': shape_features[20:40,:]})

```

Código selección con P-Value.

```
% Aquí se deben ir cargando los distintos archivos
donde están agrupadas las
% matrices con los datos radiómicos extraídos
data_shapeBenignas = load('features_benignas.mat');
data_shapeMalignas = load('features_malignas.mat');

A = data_shapeBenignas.features_benignas;
B = data_shapeMalignas.features_malignas;

[num_filas, num_muestras_A] = size(A);

for i = 1:num_filas
    [h, p_value] = ttest2(A(i, :), B(i, :));
    fprintf('P-value entre fila %d de matrix1 y
fila %d de matrix2: %.4f\n', i, i, p_value);
end
```

Código clasificación con SVM.

```
import numpy as np
import os
from scipy.io import loadmat

data=loadmat('C:/MATRIZ_RESULTADO_SELECCION_TRASPUESTA_cerocomauno.mat')
matrix=data['matriz_resultante_traspuesta']
X=matrix[:,0:7]
Y=matrix[:,7].astype(int)

y_true_all = []
y_pred_all = []

from sklearn import svm
from sklearn.model_selection import KFold
from sklearn.metrics import confusion_matrix
clf = svm.SVC()
kf = KFold(n_splits=20)
for i, (train_index, test_index) in enumerate(kf.split(X)):
    print(f"Fold {i}:")
    print(f"  Train: index={train_index}")
    print(f"  Test:  index={test_index}")
    Xtrain=X[train_index,:]
    Ytrain=Y[train_index]
    Xtest=X[test_index,:]
    Ytest=Y[test_index]
    clf.fit(Xtrain,Ytrain)
    y_pred=clf.predict(Xtest)
    y_true_all.extend(Ytest)
    y_pred_all.extend(y_pred)
    print(y_pred)
    print(Ytest)

y_true_all = np.array(y_true_all)
y_pred_all = np.array(y_pred_all)
cm = confusion_matrix(y_true_all, y_pred_all, labels=[1, 2])
TP = cm[1, 1]
TN = cm[0, 0]
FP = cm[0, 1]
FN = cm[1, 0]
sensitividad = TP / (TP + FN)
especificidad = TN / (TN + FP)
valor_predictivo_positivo = TP / (TP + FP)
valor_predictivo_negativo = TN / (TN + FN)
print(f"Verdaderos Positivos (TP): {TP}")
print(f"Verdaderos Negativos (TN): {TN}")
print(f"Falsos Positivos (FP): {FP}")
print(f"Falsos Negativos (FN): {FN}")
```

```
print(f"Sensibilidad: {sensibilidad}")
print(f"Especificidad: {especificidad}")
print(f"Valor Predictivo Positivo (PPV): {valor_predictivo_positivo}")
print(f"Valor Predictivo Negativo (NPV): {valor_predictivo_negativo}")
```