



Inteligencia artificial y Radiómica en la clasificación maligno- benigno de masas en mamas

Ariel Fernández Pirez, Marlen Pérez Díaz

RESUMEN / ABSTRACT

La mamografía es el método más fiable y preciso para la detección temprana del cáncer de mama. Sin embargo, no es tan eficaz para diferenciar entre lesiones malignas y benignas debido a la complejidad y diversidad de las lesiones, así como de la región anatómica del tórax donde se superponen muchas estructuras. Por ese motivo se recurre al procedimiento invasivo de biopsiar el tejido. En la actualidad, la radiómica se ha desarrollado como una de las posibles soluciones computacionales para ayudar a resolver al problema. Objetivo: Diseñar un sistema automatizado de diagnóstico asistido por computador (CAD), basado en radiómica y *machine learning* (ML), para clasificar masas mamarias en malignas o benignas. Método: Se extrajeron 474 características radiómicas de 181 mamografías con masas de la base de datos VinDr-Mammo, con la biblioteca PyRadiomics. Estas fueron reducidas a 4 características mediante la correlación de Pearson entre todos los pares de características, para eliminar las redundancias que no son útiles en tareas de clasificación y complejizan computacionalmente la tarea. Con este resultado se entrenaron y validaron varios modelos de inteligencia artificial *Random Forest* y *Support Vector Machine*. Resultado: El modelo con mejor desempeño fue un *Support Vector Machine* polinómico de grado 7, con el que se obtuvo una sensibilidad de 92,31% y una especificidad de 91,67% en la clasificación. Se realizó además una prueba externa con 40 mamografías de la base de datos CBIS-DDSM, obteniéndose una sensibilidad de 90,00% y una especificidad de 80,00%. Conclusión: Los resultados anteriores evidencian el poder de generalización del modelo seleccionado, así como la factibilidad del enfoque conjunto de radiómica y ML, lo que permitió diseñar un Sistema de diagnóstico asistido por computadora, para su empleo en rutina clínica.

Palabras claves: Radiómica, clasificación de masas mamarias, PyRadiomics, *Support Vector Machine*, *Random Forest*

Mammography is the most reliable and accurate method for the early detection of breast cancer. However, it is not as effective in differentiating between malignant and benign lesions due to the complexity and diversity of the lesions, as well as the anatomical region of the thorax, where many structures overlap. For this reason, the invasive procedure of tissue biopsy is often used. Currently, radiomics has been developed as one of the possible computational solutions to help solve this problem. Objective: To design an automated computer-aided diagnosis (CAD) system, based on radiomics and machine learning (ML), to classify breast masses as malignant or benign. Method: 474 radiomic features were extracted from 181 mammograms with masses from the VinDr-Mammo database, using the PyRadiomics library. These were reduced to 4 features using Pearson correlation between all pairs of features, to eliminate redundancies that are not useful in classification tasks and make the task computationally complex. With this result, several Random Forest and Support Vector Machine artificial intelligence models were trained and validated. Result: The model with the best performance was a polynomial Support Vector Machine of degree 7, with which a sensitivity of 92.31% and a specificity of 91.67% in the classification was obtained. An external test was also performed with 40 mammograms from the CBIS-DDSM database, obtaining a sensitivity of 90.00% and a specificity of 80.00%. Conclusion: The above results demonstrate the generalization power of the selected model, as well as the feasibility of the joint approach of radiomics and ML, which allowed the design of a Computer-Aided Diagnostic System for use in clinical routine.

Keywords: Radiomics, breast mass classification, PyRadiomics, Support Vector Machine, Random Forest
Artificial intelligence and radiomics in the malignant-benign classification of breast masses

Recibido: 06/2025 Aceptado: 09/2025

1. -INTRODUCCIÓN

El cáncer es una enfermedad caracterizada por el crecimiento descontrolado y la proliferación anómala de células, que pueden invadir tejidos y órganos circundantes, así como diseminarse a otras partes del cuerpo mediante un proceso conocido como metástasis [1].

Las imágenes médicas son una herramienta fundamental que proporciona un respaldo integral al proceso médico, abarcando tanto el diagnóstico como el seguimiento de la condición del paciente. Se ha comprobado que la mamografía es el método más fiable y preciso para detectar tempranamente el cáncer de mama [2]. Sin embargo, su desafío radica en que aún no es tan eficaz en la capacidad de diferenciar entre los tejidos normales y las áreas patológicas. Esto se debe a que el tejido normal y el patológico poseen coeficientes de atenuación y absorción de los rayos X muy similares. Además, las imágenes de las anomalías pueden ser extremadamente pequeñas, lo que demanda una buena definición de contrastes y una alta resolución espacial [3].

Desde la revolución de la radiografía con la tecnología digital, se han creado diversos sistemas de diagnóstico asistidos por computadora (CAD, por sus siglas en inglés). La mayoría de estos sistemas siguen un proceso que consta principalmente de cinco etapas: adquisición de imágenes, preprocesamiento, segmentación, extracción de características y clasificación [4]. Algunos CAD se enfocan únicamente en la detección, mientras que otros pueden diferenciar entre clases benigna y maligna. Cabe destacar aquí, que estos sistemas se diseñan para que actúen como un segundo lector y no sustituyen el papel del radiólogo. El CAD ayuda al especialista médico a dirigir su atención hacia áreas de posible sospecha, que son matemáticamente detectadas como patrones anómalos [4], [5].

Hoy en día, los sistemas CAD han obtenido beneficios significativos gracias al renovado enfoque de la inteligencia artificial aplicada a la imagen médica. Durante la última década, ha existido un aumento en la investigación relacionada con el aprendizaje automático (*machine learning*, ML en inglés), que es una disciplina de la inteligencia artificial.

Actualmente la investigación en el área de la salud se enfoca en lo que se conoce como "medicina de precisión". Consiste en la adopción de decisiones adaptadas para la prevención, diagnóstico y tratamiento de enfermedades. Esto se logra mediante el uso de datos individualizados de cada paciente, recolectados a través de mediciones exactas y de una integración eficiente de la información. Este enfoque da lugar al concepto de "ómica" [6]. Este término se utiliza para describir el proceso de análisis de grandes cantidades de datos complejos, con el fin de identificar características específicas o resultados precisos y extraer información valiosa [7].

En 2012, el científico holandés P. Lambin introdujo el término "radiómica", definiéndolo como la extracción de una amplia gama de características de imágenes radiográficas mediante un enfoque de alto rendimiento, con el propósito de investigar posibles correlaciones con la biología y los resultados clínicos [8]. Esta disciplina emergente se centra en desarrollar modelos para predecir los fenotipos de lesiones de forma no invasiva [9].

El concepto de radiómica se emplea en dos ámbitos: uno relacionado con el estudio de la variación genética vinculada a la respuesta ante la radiación, y otro referente a la correlación entre las características de las imágenes del cáncer y su expresión génica [10]. Esta correlación sugiere que las características de la imagen están asociadas con las firmas genéticas presentes [11].

La radiómica es un campo de investigación emergente que se enfoca en crear nuevos biomarcadores mediante el análisis de imágenes radiológicas [12]. Se parte de la premisa de que estas imágenes reflejan características subyacentes fisiopatológicas, lo que sugiere que los análisis cuantitativos pueden ofrecer información sobre la biología de las imágenes. Mediante la extracción automatizada de una amplia variedad de características cuantitativas de las imágenes, se pueden detectar de manera eficaz características sutiles, que pueden ser relevantes para el diagnóstico de enfermedades, el pronóstico y la evaluación de la respuesta al tratamiento.

La investigación sobre la clasificación de malignidad de masas en mamas se ha centrado en dos tendencias principales a nivel mundial:

1. Utilización de métodos de aprendizaje de máquinas profundo, (*deep learning*, DL en inglés), los cuales implican la construcción de una red neuronal convolucional profunda, el entrenamiento del modelo y su posterior uso para clasificar los datos [13].

2. Empleo de métodos tradicionales de *machine learning*, que se basan en la extracción de la región de interés (ROI) y su correspondiente vector de atributos o rasgos (en este caso a partir de características radiómicas), los cuales son luego utilizados para entrenar clasificadores. Tanto la selección de las características como la elección de los clasificadores tienen un impacto significativo en los resultados obtenidos [14].

Ambas corrientes, aunque abarcan diversos procedimientos técnicos, se enfocan en características de alta complejidad, extraídas de las imágenes para su aplicación en diagnóstico y predicción. En el primer caso, estas características son seleccionadas automáticamente por la red, mientras que, en el segundo, deben ser elegidas y calculadas antes de la clasificación. En el presente trabajo se sigue la segunda tendencia, siendo hasta el presente el primer estudio en Cuba en esta área del conocimiento.

Para extraer datos de rasgos importantes, una vez identificadas las áreas tumorales, es posible discernir las características de la imagen, a partir de aspectos como son la forma, la intensidad, la estructura y la textura de la lesión, además de su ubicación. Durante el proceso de extracción de características, es fundamental que la información sea informativa pero no redundante, asegurando así la reproducibilidad [15].

Estudios recientes han demostrado el potencial de la radiómica para identificar lesiones malignas en mamas. En [16] publicaron en 2007 uno de los primeros artículos que emplea características radiómicas para discriminar masas malignas de benignas. Estos autores combinaron características basadas en fractalidad, como la dimensión fractal y concavidad fractal, para etiquetar lesiones malignas y lograron un área bajo la curva sensibilidad vs. especificidad (AUC) de 0,93.

En 2019 en [17], aplicaron radiómica para evaluar cuantitativamente la textura de las lesiones mamarias obtenidas por mamografía. Se aplicó además el análisis al parénquima mamario contralateral normal, demostrando el valor añadido por la radiómica en el rendimiento de la clasificación mediante redes neuronales artificiales bayesianas, con un AUC de 0,84.

Wang et al. [18], en 2022, demuestran que es posible construir un modelo radiómico para la caracterización de masas mamográficas con un umbral de sensibilidad similar al de la biopsia. En este estudio se extrajeron 188 características, que luego fueron reducidas a 32 mediante el método de correlación de Pearson y a 14 mediante *least absolute shrinkage and selection operator* (LASSO), para luego construir el modelo con SVM. Los resultados de este trabajo muestran un AUC de 0,901, sensibilidad de 0,958 y especificidad de 0,667. El trabajo reportado en [19] soporta resultados semejantes.

2.- METODOLOGIA

2.1 Flujo de trabajo

Un esquema con los diferentes pasos para obtener el sistema CAD se incluye en la Figura 1.

El proceso comienza con el preprocesamiento de las imágenes de mamografía digital de campo completo adquiridas previamente en formato DICOM (*Digital Imaging and Communications in Medicine*), para asegurar su uniformidad y calidad antes de pasar a un proceso de segmentación de la región de la masa, en el cual se elimina la parte de la imagen no relevante para el análisis.

Una vez preprocesadas las imágenes y segmentada la región de la masa, se lleva a cabo la extracción de características radiómicas, utilizando la biblioteca PyRadiomics. Este paso convierte las imágenes en conjuntos de datos numéricos que describen propiedades físicas y texturales de las masas mamarias.

Dado que se obtienen una gran cantidad de características, es necesario aplicar técnicas de reducción de dimensionalidad, para eliminar aquellas que son redundantes o no aportan valor al proceso de clasificación. Para esto, se utiliza el coeficiente de correlación de Pearson, una medida estadística que evalúa la relación lineal entre dos variables, en este caso, entre cada par de características radiómicas, y entre cada característica y la anotación de malignidad de las masas.

A continuación, con el objetivo de lograr la mejor clasificación posible, se probaron dos algoritmos de aprendizaje automático con modelos de SVM y RF. Ambos modelos fueron entrenados y validados con los datos de características seleccionadas previamente. Para ambos algoritmos, se llevó a cabo un ajuste manual de los parámetros más influyentes en la clasificación, para optimizar el rendimiento del modelo en términos de sensibilidad, especificidad, precisión, exactitud y AUC.

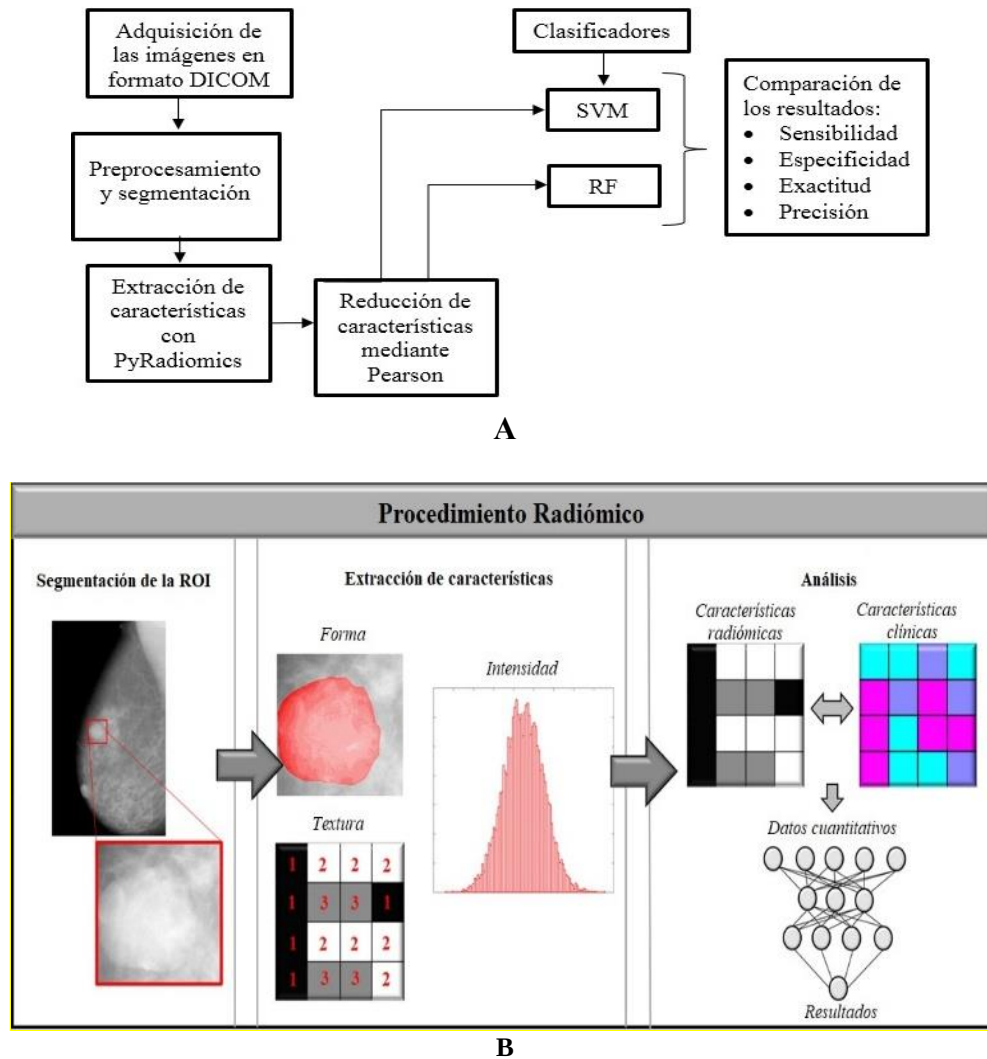


Figura 1
Flujo de trabajo con radiómica e inteligencia artificial: A- Esquema general, B- Procedimiento radiómico

Finalmente, se realizó una prueba externa, utilizando imágenes de la base de datos CBIS-DDSM (*Curated Breast Imaging Subset of DDSM*) lo que permitió verificar la capacidad de generalización del mejor modelo a nuevas imágenes provenientes de un conjunto de datos completamente independiente. Las métricas principales utilizadas para la evaluación de los modelos fueron: sensibilidad, especificidad, precisión, exactitud y AUC.

2.2 Imágenes

Para entrenamiento y validación de modelos de machine learning se utilizó la base de datos VinDr-Mammo [20]. Cada estudio contiene la anotación de 3 radiólogos expertos en Sistema BI-RAD, en escala 3 al 5, donde 3 significa probablemente benigno, 4 sospechoso y 5 maligno. Para lograr etiquetas correctas para entrenar los modelos se utilizaron los casos con media BI-RADS 3 como benignos y BI-RADS 5 como malignos. Se escogió un subconjunto de 181 imágenes en total, de las cuales 110 corresponden a masas BI-RADS 3 y 71 a BI-RADS 5. Se tomó el 80 % de los datos para entrenamiento y el 20 % restante para validar los modelos. Para la evaluación externa del mejor modelo entrenado, se empleó la base de datos de CBIS-DDSM [21], en este caso se tomaron al azar 20 casos con masas benignas y 20 con masas malignas.

A modo de apreciar visualmente las características de las imágenes de ambas bases de datos, la Figura 2 muestra una imagen patológica de cada una.

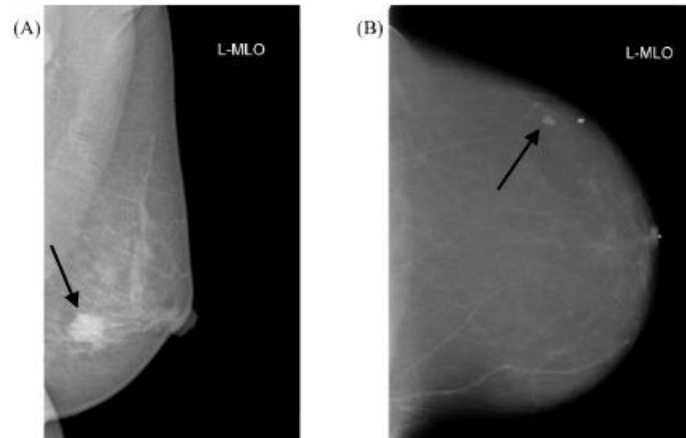


Figura 2

A Imagen de la base de datos VinDr-Mammo de una mama izquierda donde se observa una masa. B Imagen de la base de datos CBIS-DDSM de una mama izquierda donde se observa una masa

2.3 Procesamiento

El proyecto fue desarrollado en Python (versión 3.9), utilizando el entorno de desarrollo integrado PyCharm Community (versión 2024.1.4). Los modelos fueron implementados y ejecutados en una computadora personal con las siguientes especificaciones: Procesador: Intel® Core™ i3-3120M, 2,50 GHz, Memoria RAM: 8 GB.

El proceso comenzó con la lectura de las imágenes en formato DICOM y la extracción de las regiones de interés. Estas se determinaron utilizando las coordenadas provistas en las anotaciones de la base de datos, que definen un rectángulo donde se encierra la masa. Seguidamente se normalizaron las imágenes. La normalización convierte los valores de los píxeles a una escala uniforme (0-255), lo que facilita el procesamiento posterior y asegura que las imágenes tengan una intensidad de píxel comparable. Luego se aplicó un recorte de percentiles. Consiste en eliminar una porción fija de los valores extremos del histograma de intensidades de la imagen, con el objetivo de mejorar el contraste y eliminar los valores atípicos que pueden distorsionar la imagen. En este caso, se decidió eliminar el 10% más alto y más bajo de los valores de intensidad, esto quiere decir que los valores de intensidad menores que el percentil 10 del histograma se llevaron a intensidad 0 (color negro) y los valores mayores que el percentil 90 se llevaron a intensidad 255 (color blanco). Este método ayudó a reducir el ruido y estabilizó las variaciones de intensidad, mejorando así la visualización de las estructuras relevantes en la imagen.

Seguidamente se procedió a segmentar la zona de la masa mediante RGI (*Radial Gradient Index*) [22]. Se generó el contorno de cada masa y se produjo una máscara que delimita la región de la masa. Las máscaras se convirtieron a formato binario, donde los píxeles pertenecientes a las regiones de interés se establecieron en 1 y el resto en 0. Finalmente, estas máscaras fueron guardadas en formato DICOM para mantener la consistencia y estructura con los datos originales. El resultado de estos pasos puede ser observado en la Figura 3.

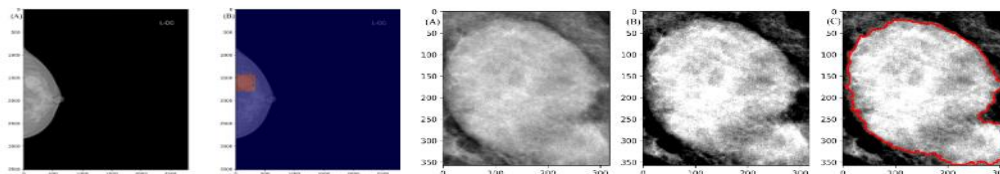


Figura 3

Flujo de pre-procesamiento. De izquierda a derecha: Imagen original, selección de la región de interés que contiene la masa, recorte y zoom de la masa, imagen de la masa procesada digitalmente, masa segmentada con contorno delimitado.

2.4 Extracción de características radiómicas

Para la extracción de características radiómicas, se empleó la herramienta PyRadiomics [23], un paquete de software desarrollado para Python y diseñado para obtener un amplio conjunto de características cuantitativas a partir de imágenes médicas. Este proceso permite transformar las imágenes en datos numéricos que pueden ser analizados mediante algoritmos de *machine learning*.

El procedimiento comenzó con la configuración del extractor de características, utilizando un archivo YAML (*Yet Another Markup Language*) [24]. Este archivo especifica los parámetros necesarios para definir las características a extraer, los métodos de preprocesamiento de las imágenes y otros ajustes específicos de los parámetros de extracción.

Inicialmente, el software lee un archivo tipo CSV. Este contiene las anotaciones de las imágenes de mamografía e incluye la identificación de cada imagen y su respectiva máscara segmentada. Luego se iteró sobre cada fila del *Data Frame* (DF) obtenido del archivo tipo CSV, para cargar cada imagen y su máscara binaria con la biblioteca “SimpleITK”, que es compatible con el formato DICOM de las imágenes empleadas, para ser posteriormente procesadas por el extractor de características de PyRadiomics.

El extractor aplica una serie de algoritmos para calcular diversas características radiómicas de cada imagen, entre ellas: las de primer orden, las de forma y las de textura (GLCM, GLRLM, GLSZM, NGTDM, y GLDM) [25]. Los resultados de la extracción de características se almacenaron en una lista, que luego se convirtió en un DF de la biblioteca “Pandas”. Este DF facilitó la manipulación y análisis de los datos extraídos, para finalmente, guardarlos en un archivo CSV, que se utilizó como entrada a los algoritmos de *machine learning* en las etapas posteriores del estudio.

Se extrajeron un total de 474 características radiómicas, divididas en 9 características de forma, 18 características estadísticas de primer orden, 24 características de GLCM, 16 características de GLRLM, 16 características de GLSZM, 14 características de GLDM y 5 características de NGTDM. Las características de 1er orden y las de textura se calcularon a su vez 5 veces, para incluir no solo a la imagen original sino a cada imagen procesada por 5 filtros Wavelet.

2.5 Reducción de dimensión en el problema computacional

El proceso de reducción de dimensionalidad sirve para conservar únicamente las características que aporten utilidad significativa al modelo de clasificación. Para ello, se empleó el cálculo del coeficiente de correlación de Pearson. En este caso, se evaluó la correlación entre todos los pares de características radiómicas y, además, entre cada característica y la etiqueta de malignidad de las masas mamarias (anotación media realizada por los 3 expertos). Una vez calculados los coeficientes de correlación para todos los pares de características, se estableció un umbral de 0,60 (este valor es el umbral de correlación fuerte) para eliminar las características redundantes. Para seleccionar cuál descartar, se priorizó la que presentaba una correlación más baja con la etiqueta de malignidad, manteniendo así la característica que mostrara una relación más significativa con la clasificación de masas como benignas o malignas. Al reducir el número de características a aquellas no redundantes y más representativas para la clasificación, se mejora el proceso de entrenamiento de los clasificadores. Esto no solo reduce el tiempo de cómputo, sino que también minimiza el riesgo de ruido en los datos y aumenta la eficiencia del modelo al centrarse en las características de mayor valor predictivo.

De las 474 características iniciales, se seleccionaron como no redundantes *original_shape2D Elongation*, *original_shape2D Sphericity*, *wavelet-LL_glcm_Joint Entropy* y *wavelet-LL_ngtdm_Busyness* [25].

2.6 Implementación del *machine learning*, entrenamiento y evaluación de modelos

En este trabajo se implementaron 12 modelos de clasificación supervisada a partir de *Support Vector Machine* (SVM) y *Random Forest* (RF). Estos clasificadores fueron elegidos por su capacidad de manejar problemas de clasificación complejos y su eficiencia en el manejo de datos que pueden tener alta dimensionalidad a partir de características radiómicas. En los 6 SVM, para implementar modelos de clasificación, se probó el núcleo lineal, el polinómico y el radial (RBF), a partir de la experiencia acumulada en trabajos anteriores del grupo de trabajo [26]. Para entrenar los modelos, se cargaron las características radiómicas y las anotaciones de malignidad. Los datos fueron divididos en un 80% para entrenamiento y un

20% para prueba, utilizando la función “train_test_split” de la biblioteca “scikit-learn”. Los modelos fueron entrenados cambiando manualmente varios parámetros, como el parámetro de penalización C, que controla el equilibrio entre maximizar el margen y minimizar el error de clasificación, el tipo de núcleo y el grado del polinomio, en caso de utilizar un núcleo polinómico.

En el caso de los *Random Forest* es un algoritmo de aprendizaje supervisado basado en la creación de múltiples árboles de decisión. Cada árbol en el bosque toma una decisión basada en una muestra aleatoria de las características y datos. Se realizaron múltiples ensayos de prueba y error para determinar los hiperparámetros del modelo (número de nodos padres, número de nodos hijos y profundidad del bosque). En total se probaron también 6 modelos de RF.

Todos modelos fueron evaluados mediante cinco métricas: sensibilidad, especificidad, precisión, exactitud y AUC (área bajo la curva ROC sensibilidad vs. especificidad) [27], lo que permitió medir no solo cuántas predicciones eran correctas, sino también cuán bien los modelos detectaban correctamente las masas malignas y benignas. El mejor modelo global obtenido fue sometido entonces a una prueba externa con datos de otra BD no utilizada durante el entrenamiento y la validación y se evaluó con las mismas métricas.

Las expresiones de cálculo de las métricas son las siguientes.

$$\text{Sensibilidad} = \frac{VP}{VP+FN} \quad (1)$$

$$\text{Especificidad} = \frac{VN}{VN+FP} \quad (2)$$

$$\text{Precisión} = \frac{VP}{VP+FP} \quad (3)$$

$$\text{Exactitud} = \frac{VP+VN}{FP+FN+VP+VN} \quad (4)$$

donde: VP es verdadero positivo, VN es verdadero negativo, FP es falso positivo y FN es falso negativo. Todos estos valores se toman de la respuesta de cada modelo con relación a la anotación médica que aparece en ambas BD.

3.- RESULTADOS Y DISCUSIÓN

Las Tablas 1 y 2 muestran los resultados de los 12 modelos de IA conformados para clasificación con *machine learning* (6 RF y 6 SVM). Los resultados de los experimentos con el clasificador *Random Forest* indican que la configuración de los hiperparámetros impactó significativamente en su rendimiento. El modelo que de forma general presenta mejor desempeño es el primero, con 500 estimadores y sin restricción de profundidad. Los resultados de los experimentos SVM indican que la configuración de los hiperparámetros impactó también significativamente en su rendimiento. Las configuraciones polinómicas con mayor grado presentaron el mejor desempeño en cuanto a sensibilidad, aunque se perdió poder de clasificación en la clase negativa. Por tanto, basado en lo anterior, se considera que las mejores configuraciones de los SVM probados se obtuvieron con los *kernels* polinómicos de grado 7 y 9. La elección del mejor entonces se decantó por el grado 7, por poseer menor costo computacional.

El RF con 500 estimadores fue un modelo también de buen desempeño, pero el modelo SVM polinómico de grado 7, presentó su misma sensibilidad, y lo supera, además, en todas las demás métricas. Por esta razón, este el modelo se sometió a una prueba externa, con una data de diferente origen a la utilizada para entrenarlo y validarlo.

Los resultados de la prueba externa fueron satisfactorios, logrando una sensibilidad del 90,00%, una especificidad del 80,00% y una exactitud del 85,00%. Es importante señalar que, a pesar de la calidad inferior de las imágenes en CBIS-DDSM (mamografías digitalizadas de placas fílmicas) en comparación con las imágenes de campo completo de alta resolución de la base VinDr-Mammo (utilizadas en el entrenamiento), el modelo mantuvo un rendimiento no aleatorio. Esto sugiere que el modelo seleccionado para diseñar el CAD posee poder de generalización para la clasificación de masas mamarias.

El éxito en esta prueba externa sugiere que las estrategias de pre-procesamiento y selección de características aplicadas, incluyendo la reducción de ruido y el ajuste de contraste, resultaron efectivas en la minimización de la superposición entre clases y en la mejora de la claridad de los datos, lo que contribuyó a una robusta generalización frente a imágenes de origen diferente a las incluidas en el entrenamiento del modelo.

El costo computacional del mejor modelo fue de solo 5 segundos con las prestaciones computacionales utilizadas.

Tabla 1
Resultados de los modelos *Random Forest*

Modelo RF	Sensibilidad %	Especificidad %	Precisión %	Exactitud %	AUC
<i>n_estimators</i> = 500 y <i>max_depth</i> = None	92,31	87,50	80,00	89,19	0,946
<i>n_estimators</i> = 100 y <i>max_depth</i> = None	92,31	83,33	75,00	86,49	0,944
<i>n_estimators</i> = 20 y <i>max_depth</i> = None	92,31	83,33	75,00	86,49	0,934
<i>n_estimators</i> = 500 y <i>max_depth</i> = 1	84,62	95,83	91,67	91,89	0,946
<i>n_estimators</i> = 100 y <i>max_depth</i> = 1	84,62	95,83	91,67	91,89	0,946
<i>n_estimators</i> = 20 y <i>max_depth</i> = 1	84,62	87,50	78,57	86,49	0,934

Tabla 2
Resultados de los modelos *Support Vector Machine*

Modelo SVM	Sensibilidad %	Especificidad %	Precisión %	Exactitud %	AUC
Kernel Lineal	84,62	87,50	78,57	86,49	0,965
Kernel Radial	84,62	95,83	91,66	91,89	0,978
Kernel Polinómico Grado 3	84,62	95,83	91,66	91,89	0,968
Kernel Polinómico Grado 5	84,62	95,83	91,66	91,89	0,966
Kernel Polinómico Grado 7	92,31	91,67	85,71	91,89	0,968
Kernel Polinómico Grado 9	92,31	91,67	85,71	91,89	0,971

Comparando los presentes resultados con enfoques similares en la literatura científica [16]- [19] se puede afirmar que el modelo propuesto muestra un rendimiento comparable con el estándar internacional. Si se tienen en cuenta solamente los resultados de la validación interna, el modelo propuesto exhibe un rendimiento general superior a sus homólogos de la literatura analizada. Por ejemplo, los experimentos realizados en [18], que son los que ofrecen una mayor sensibilidad, alcanzando valores superiores al 95%, lo hacen a expensas de reducir la especificidad hasta un 66.7 %. Todos los demás estudios analizados presentan métricas, cercanas e inferiores a las obtenidas con el modelo propuesto. Las principales causas pueden ir desde la calidad de la segmentación de las imágenes y el método de selección de características, hasta la elección del clasificador y sus parámetros [28].

4.- CONCLUSIONES

Esta investigación confirma que el empleo de características radiómicas extraídas de imágenes mamográficas es eficaz para clasificar masas mamarias en benignas o malignas a través de modelos de *machine learning*. Estas características cubren una amplia variedad de atributos, varios de los cuales son determinantes para la clasificación. Se comprobó que clasificadores de ML, como RF y SVM, entrenados con estas características radiómicas, logran una buena clasificación de masas, incluso luego de entrenarse con un conjunto de datos relativamente reducido. Además, el mejor modelo escogido demostró una buena capacidad de generalización ante datos provenientes de fuentes distintas a las del entrenamiento inicial. Los valores obtenidos de sensibilidad y especificidad en prueba externa del mejor modelo entrenado superan el 80 %, alcanzando valores superiores al 90 % en validación interna, lo que coloca estos resultados en un nivel comparable o incluso superior a los que se logran con otros sistemas CAD profesionales internacionales.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado gracias la Oficina de Gestión de Fondos y Proyectos Internacionales bajo el código PN223LH004-039.

REFERENCIAS

1. B. S. Chhikara and K. Parang. Global Cancer Statistics 2022: the trends projection analysis. *Chemical Biology Letters* 10(1): 451–451, Jan. 2023.
2. S. Pacilè, J. Lopez, P. Chone, T. Bertinotti, J. M. Grouin, and P. Fillard. Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. *Radiology: Artificial Intelligence* 2(6): 1–9, Nov. 2020. doi: 10.1148/ryai.2020190208.
3. E. L. Henriksen, J. F. Carlsen, I. M. M. Vejborg, M. B. Nielsen, and C. A. Lauridsen. The efficacy of using computer-aided detection (CAD) for detection of breast cancer in mammography screening: a systematic review, *Acta Radiologica*, 60(1). SAGE Publications Inc., p.: 13–18, Jan. 2019. doi: 10.1177/0284185118770917.
4. J. D López-Cabrera, L. A. López Rodríguez, and M. Pérez-Díaz. Classification of breast cancer from digital mammography using deep learning. *Inteligencia Artificial* 23(65): 56–66, May 2020. doi: 10.4114/intartif.vol23iss65pp56-66.
5. S. Parvathavarthini, N. K. Visalakshi, and S. Shanthi. Breast cancer detection using crow search optimization based intuitionistic fuzzy clustering with neighborhood attraction. *Asian Pacific Journal of Cancer Prevention* 20(1): 157–165, Jan. 2019. doi: 10.31557/APJCP.2019.20.1.157.
6. B. B. Misra, C. Langefeld, M. Olivier, and L. A. Cox. Integrated omics: Tools, advances and future approaches, *Journal of Molecular Endocrinology*, 62(1). Bioscientifica Ltd, p.: R21–R45, Jan. 01, 2019. doi: 10.1530/JME-18-0055.
7. I. Subramanian, S. Verma, S. Kumar, A. Jere, and K. Anamika. Multi-omics Data Integration, Interpretation, and Its Application, *Bioinformatics and Biology Insights*, 14. SAGE Publications Inc., Jan. 31, 2020. doi: 10.1177/1177932219899051.
8. P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. Van Stiphout, P. Granton, *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer* 48(4): 441–446, Mar. 2012. doi: 10.1016/j.ejca.2011.11.036.
9. P. B. Dean. Radiomics and Breast Cancer Management, *Academic Radiology*, 29(12). Elsevier Inc., p.: 1783–1785, Dec. 2022. doi: 10.1016/j.acra.2022.09.012.
10. A. S. Tagliafico, M. Piana, D. Schenone, R. Lai, A. M. Massone, and N. Houssami. Overview of radiomics in breast

- cancer diagnosis and prognostication. *Breast* 49: 74–80, Feb. 2020. doi: 10.1016/j.breast.2019.10.018.
11. A. Conti, A. Duggento, I. Indovina, M. Guerrisi, and N. Toschi. Radiomics in breast cancer classification and prediction, *Seminars in Cancer Biology*, 72. Academic Press, p.: 238–250, Jul. 01, 2021. doi: 10.1016/j.semcancer.2020.04.002.
 12. T. Pang, J. H. D. Wong, W. L. Ng, and C. S. Chan. Deep learning radiomics in breast cancer with different modalities: Overview and future, *Expert Systems with Applications*, 158. Elsevier Ltd, Nov. 15, 2020. doi: 10.1016/j.eswa.2020.113501.
 13. F. Lubinus Badillo, C. A. Rueda Hernández, B. Marconi Narváez, and Y. E. Arias Trillos. Redes neuronales convolucionales: un modelo de Deep Learning en imágenes diagnósticas. Revisión de tema. *Revista colombiana de radiología* 32(3): 5591–5599, Sep. 2021. doi: 10.53903/01212095.161.
 14. E. Pérez-Careta, J. R. Guzmán- Sepúlveda, J. M. Lozano-García, M. Torres-Cisneros, and R. Guzmán-Cabrera. Classification of medical images using machine learning, *Dyna (Spain)*, 97(1). p.: 35–38, 2022. doi: 10.6036/10117.
 15. S. H. Lee, H. Park, and E. S. Ko. Radiomics in breast imaging from techniques to clinical applications: A review, *Korean Journal of Radiology*, 21(7). Korean Society of Radiology, p.: 779–792, Jul. 2020. doi: 10.3348/kjr.2019.0855.
 16. R. M. Rangayyan and T. M. Nguyen. Fractal analysis of contours of breast masses in mammograms. *Journal of Digital Imaging* 20(3): 223–237, Sep. 2007. doi: 10.1007/s10278-006-0860-9.
 17. H. Li, K. R. Mendel, L. Lan, D. Sheth, and M. L. Giger. Digital mammography in breast cancer: Additive value of radiomics of breast parenchyma. *Radiology* 291(1): 15–20, Apr. 2019. doi: 10.1148/radiol.2019181113.
 18. G. Wang, D. Shi, Q. Guo, H. Zhang, S. Wang, and K. Ren. Radiomics Based on Digital Mammography Helps to Identify Mammographic Masses Suspicious for Cancer. *Frontiers in Oncology* 12(April): 1–10, 2022. doi: 10.3389/fonc.2022.843436.
 19. N. Mao, P. Yin, Q. Wang, M. Liu, J. Dong, X. Zhang, *et al.* Added Value of Radiomics on Mammography for Breast Cancer Diagnosis: A Feasibility Study. *Journal of the American College of Radiology* 16(4): 485–491, Apr. 2018. doi: 10.1016/j.jacr.2018.09.041.
 20. H. T. Nguyen, H. Q. Nguyen, H. H. Pham, K. Lam, L. T. Le, M. Dao, *et al.* VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography. *Scientific Data* 10(1)2023. doi: 10.1038/s41597-023-02100-7.
 21. R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin. Data Descriptor: A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data* 4(1): 1–9, Dec. 2017. doi: 10.1038/sdata.2017.177.
 22. M. A. Kupinski and M. L. Giger. Automated seeded lesion segmentation on digital mammograms. *IEEE Transactions on Medical Imaging* 17(4): 510–517, 1998. doi: 10.1109/42.730396.
 23. AIM Harvard. Welcome to pyradiomics documentation., *Read the Docs*, 2022. <https://pyradiomics.readthedocs.io/en/latest/> (accessed Sep. 27, 2024).
 24. YAML. The Official YAML Web Site, <https://yaml.org/>, 2022. <https://yaml.org/> (accessed Sep. 27, 2024).
 25. A. Abbasian Ardakani, N. J. Bureau, E. J. Ciaccio, and U. R. Acharya. Interpretation of radiomics features—A pictorial review, *Computer Methods and Programs in Biomedicine*, 215. Elsevier, p.: 106609, Mar. 01, 2022. doi: 10.1016/j.cmpb.2021.106609.
 26. J. A. Castro, M. Pérez Díaz, and R. Morales. Detection of lung nodules using support vector machine. *itegam- Journal of Engineering and Technology for Industrial Applications (ITEGAM-JETIA)* 10(48): 69–74, Jul. 2024. doi: 10.5935/jetia.v10i48.1202.
 27. L. Breiman. Random forests. *Random Forests*, 1–122. *Machine Learning* 45(45): 5–32, Oct. 2001. doi: 10.1023/A:1010933404324/METRICS.
 28. R. P. Mathews. CAD in Medical Imaging : A Review of Current Trends and Future Directions. *IJERT* 8(06): 199–213, Jun. 2019. doi: 10.17577/IJERTV8IS060243.

CONFLICTO DE INTERESES

Los autores declaran que no presentan ningún conflicto de interés con la entidad en la que laboran o con otra entidad, persona o grupo de personas, respecto a la presente investigación.

CONTRIBUCIONES DE LOS AUTORES

Ariel Fernández Pirez participa en: Conceptualización, curación de datos, análisis formal, desarrollo de tareas de investigación, implementación de la metodología, programación de software, validación de resultados y visualización de estos.

Marlen Pérez Díaz participa en: Conceptualización, análisis formal, adquisición de fondos, diseño metodológico, administración del proyecto, gestión de recursos, supervisión de todos los procesos, visualización, revisión de resultados, redacción del borrador original y edición del manuscrito.

AUTORES

Ariel Fernández Pirez: Ing. en Automática. UEB CEDAI Cienfuegos, Cuba. arielfdez26@gmail.com, <https://orcid.org/0009-0008-4372-9156>. Los principales intereses de investigación son la automatización de procesos a partir de los métodos de *Machine Learning* así como el análisis y procesamiento de grandes volúmenes de datos.

Marlen Pérez Díaz. Ing. Física Nuclear, DrC Físicas, especialista en Física médica. Universidad Central Marta Abreu de las Villas, Departamento de Automática. Villa Clara, Cuba. mperez@uclv.edu.cu, <https://orcid.org/0000-0002-3706-9154> Los intereses de investigación son en métodos de *Machine Learning* para detección de patologías a partir de imágenes médicas, análisis y procesamiento de imágenes médicas digitales.



Esta revista se publica bajo una [Licencia Creative Commons Atribución-No Comercial-Sin Derivar 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)