



Algoritmo para reducir características radiómicas en la clasificación de nódulos pulmonares

Rachel Abreu-Llanes, Marlen Perez-Diaz, Yusely Ruiz-Gonzalez, Deborah Galpert-Cañizares, Ruben Orozco-Morales

RESUMEN / ABSTRACT

El cáncer de pulmón constituye una de las principales causas de mortalidad a nivel mundial. La Tomografía Computarizada (TC) es la técnica más utilizada para el diagnóstico temprano y la evaluación de la respuesta al tratamiento. En este contexto, la Radiómica ha emergido para extraer automáticamente un gran volumen de características cuantitativas de imágenes digitales. No obstante, el desafío fundamental reside en la gestión de la redundancia y la inestabilidad de estas características, lo que compromete la fiabilidad y la generalización. En este estudio, se estableció un sistema basado en la reducción de características Radiómicas y Aprendizaje Automático. Se propuso un enfoque estadístico basado en el cálculo del Coeficiente de Correlación de Pearson entre cada característica y la anotación de los datos para lograr una reducción sensible y eficiente de los atributos, minimizando la redundancia y el costo computacional. Se empleó un conjunto de 275 nódulos de la base de datos LIDC-IDRI anotada. Se extrajo un grupo inicial de 102 características y se redujeron a solo 5 para la clasificación binaria y 3 para la clasificación de cinco grados de malignidad. Se entrenaron y validaron diez modelos de Support Vector Machine (SVM) y Random Forest (RF). En ambos experimentos, los mejores modelos lograron valores de sensibilidad, especificidad, precisión y exactitud superiores al 90%, superando los resultados alcanzados por diferentes enfoques reportados en la literatura científica para la misma base de datos. Se concluye que la metodología de reducción propuesta contribuye a un clasificador más robusto y fiable con un costo computacional excepcionalmente bajo.

Palabras claves: Radiómica, Nódulos pulmonares, Tomografía computarizada, Aprendizaje automático, Reducción de características

Lung cancer is one of the leading causes of death. Computed tomography is the most widely used technique to diagnose and evaluate responses to treatment. Computer-aided diagnosis systems have proven effective in assisting radiologists in their decision making. Radiomics aims to automatically extract features from digital images and develop models to predict lesion phenotypes non-invasively. However, managing the redundancy and instability of these features remains a significant challenge, compromising model reliability and generalizability. In this study, a system based on reduced Radiomics-Machine Learning was established. An approach based on the calculation of the Pearson correlation coefficient between each radiomic feature and the database annotation was proposed to achieve sensitive and efficient feature reduction, avoiding redundancies and reducing computational cost compared to other algorithms in the state of the art. A set of 275 nodules from the annotated LIDC-IDRI database was used. A group of 102 radiomic features was extracted and reduced to just 5 for binary classification and 3 for classification into five degrees of malignancy. Ten models of Support Vector Machine and Random Forest were trained and validated. In both experiments, the best models achieved sensitivity, specificity, precision, and accuracy values above 90%, surpassing the values achieved by different approaches in scientific literature for the same database. It is concluded that the proposed feature reduction methodology contributes to a more robust and reliable classifier with exceptionally low computational cost.

Keywords: Radiomics, Lung nodules, Computer Tomography, Machine Learning, Feature reduction

Algorithm to reduce radiomic features for classification of lung nodules.

Recibido: 01/2026 Aceptado: 02/2026

1. -INTRODUCCIÓN

El cáncer de pulmón se mantiene como una de las principales causas de muerte a nivel global, lo que subraya la necesidad crítica de herramientas de diagnóstico y pronóstico más precisas y oportunas [1]. En este escenario, la tomografía computarizada (TC) es la técnica de imagenología dominante para la detección y evaluación de las lesiones pulmonares. Un nódulo pulmonar se define típicamente como una opacidad de forma aproximadamente redondeada, con un diámetro de hasta 3 cm. No obstante, la variabilidad morfológica, textural y característica de estas lesiones dificulta la determinación precisa del grado de malignidad por parte del radiólogo. Por ello, el apoyo de herramientas informáticas, como los Sistemas de Diagnóstico Asistido por Computadora (CAD), resulta esencial para mejorar la detección y clasificación de las lesiones [2].

La práctica médica actual se orienta hacia la denominada "Medicina de Precisión", que se basa en el análisis exhaustivo de datos individuales de los pacientes. Dentro de esta tendencia, la radiómica ha emergido como una disciplina clave, cuyo objetivo es extraer automáticamente características cuantitativas (atributos) de imágenes médicas digitales con los cuales desarrollar modelos predictivos no invasivos para el fenotipo de la lesión [3]. El proceso de clasificación radiómica se divide generalmente en cinco etapas principales: (a) adquisición y reconstrucción de la imagen, (b) segmentación y representación de la imagen, (c) extracción y calificación de características, (d) gestión de la base de datos e intercambio de datos, y (e) análisis computacional [4].

Si bien en Radiómica se definen vectores de atributos de alta calidad que sirven como entrada para los algoritmos de clasificación de aprendizaje automático (ML), la construcción de modelos robustos enfrenta un desafío significativo: la gestión de la incertidumbre. Esta incertidumbre se origina en múltiples fuentes, como las variaciones en los protocolos de adquisición de imágenes, los algoritmos de reconstrucción de la TC, y de manera crucial, la variabilidad inter-observador en la segmentación manual de los nódulos [5,6].

Esta variabilidad se propaga inevitablemente a la etapa de extracción de características. Los sistemas típicos de cálculo radiómico generan un número muy elevado de características, a menudo superando el centenar (102 en este estudio) y llegando incluso a 750 por imagen en otros trabajos [7–12]. Como se aprecia en los mismos, no existe un consenso evidente entre los autores sobre cuántas y cuáles características son óptimas, llevando a muchos a calcular todas las disponibles en la plataforma utilizada. Este enfoque introduce un alto grado de redundancia e inestabilidad en el vector de rasgos, lo que compromete seriamente la fiabilidad y la capacidad de generalización del modelo resultante y aumenta innecesariamente el costo computacional sin que aumente la precisión en los resultados. Por esta razón, la selección de un vector pequeño de atributos, pero altamente informativo y robusto, es imperativa para construir un modelo preciso y eficiente.

La detección de redundancias en la contribución de las características a la clasificación es un problema recurrente en la literatura. La selección de un vector de atributos pequeño, pero altamente informativo y robusto, es imperativa para construir un modelo preciso y eficiente.

La revisión de la literatura sobre la clasificación de nódulos pulmonares revela una variedad de enfoques para abordar el problema de la alta dimensionalidad de las características radiómicas. Es fundamental agrupar estos enfoques para establecer la posición y novedad de la metodología propuesta. Una aproximación común para reducir la dimensionalidad implica el uso de modelos de aprendizaje automático para identificar las características más relevantes, o bien, filtros basados en la estabilidad de las características. Por ejemplo, un estudio [7] utilizó nódulos de 593 pacientes de la base de datos del *Lung Imaging Database Consortium* (LIDC-IDRI), extrayendo 150 características radiómicas. Para la reducción, se empleó un modelo RF para identificar los atributos con alta correlación con los resultados de clasificación obtenidos por una SVM. Si bien este enfoque logró un rendimiento aceptable (ACC 82.7%, sensibilidad 79.5%, especificidad 86.7%), el proceso es computacionalmente complejo, requiriendo un número elevado de iteraciones para converger a los mejores valores.

En [4] se utilizaron 97 imágenes de la base de datos LIDC-IDRI. Dado que el modelo radiómico implica una red neuronal convolucional y el número de datos disponibles es bajo, se realizó un aumento de datos mediante deformación espacial. Con estos datos, se entrenó un secuenciador radiómico basado en redes neuronales convolucionales profundas. A continuación, se utilizó un clasificador de árbol binario. Los resultados fueron: precisión: 77.52 % y sensibilidad: 79.06 %. La limitación de este tipo de método es que para que el entrenamiento sea eficaz y sin sobreajuste, se requiere una gran cantidad de datos y un tiempo de cálculo considerable.

En general, la limitación fundamental del aprendizaje profundo (DL) en aplicaciones médicas es la considerable necesidad de grandes volúmenes de datos para realizar un aprendizaje significativo y sin sesgos. Dado que obtener suficientes datos en Medicina no es una tarea trivial, los enfoques basados en vectores de características y ML tradicional siguen siendo más plausibles y eficaces en escenarios de datos escasos.

Otro enfoque [9] utilizó 219 características extraídas de imágenes del *National Lung Screening Trial*. Este trabajo se centró en un filtrado de estabilidad, seleccionando solo 23 características que se mantenían constantes a lo largo de los datos. Aunque

el filtrado por estabilidad es crucial para la robustez, el modelo RF resultante solo alcanzó un ACC del 80%. Esto demuestra que la estabilidad, por sí sola, no garantiza la máxima relevancia predictiva para el resultado clínico.

Otros investigadores [10,11] han recurrido a métodos estadísticos y de regularización para manejar la alta dimensionalidad. El algoritmo *Least Absolute Shrinkage and Selection Operator* (LASSO) se ha utilizado para la selección de atributos en diferentes contextos. En [11] 103 características radiómicas de 72 imágenes de LIDC-IDRI fueron reducidas con LASSO para alimentar un modelo SVM. De manera similar, otro trabajo [12] utilizó LASSO para reducir 385 características de 294 casos de una base de datos privada, logrando clasificaciones superiores al 80% mediante regresión lineal. Una alternativa en [11] fue utilizar la prueba de suma no paramétrica de Wilcoxon, que logró reducir un conjunto inicial de 750 características a solo 4 para clasificar 75 nódulos a partir de una base de datos privada, con métricas de éxito superiores al 84%.

Existe una necesidad clara de desarrollar metodologías de reducción de características que sean simultáneamente altamente eficientes, de bajo costo computacional y capaces de seleccionar vectores de atributos robustos que conduzcan a un rendimiento diagnóstico superior. El enfoque propuesto en el presente trabajo se diferencia de los métodos complejos (RF basado en rasgos, LASSO) y de los enfoques dependientes de redes neuronales convolucionales (CNN) al emplear una técnica estadística simple univariada (correlación de Pearson), que garantiza que solo los atributos con una relación fuerte y directa con el resultado clínico permanezcan en el sistema, lo que mitiga inherentemente la redundancia y mejora el rendimiento general.

Para abordar específicamente el desafío de la redundancia y la inestabilidad de las características, las contribuciones científicas fundamentales que se hacen en el presente trabajo son un método de reducción de características estadístico simple el cual tiene una doble ventaja: (1) mitiga el impacto de características inciertas o no informativas al retener solo aquellas con una correlación fuerte y estable con la anotación clínica; y (2) construye un modelo de clasificación inherentemente más robusto y generalizable con un costo computacional notablemente bajo.

El presente artículo se estructura de la siguiente manera: en la sección 1 se realiza una revisión de trabajos relacionados, en la sesión 2 se detallan los materiales y métodos utilizados, en la sección 3 se presentan los resultados, en la sección 4 se discuten los resultados y en la sesión 5 se presentan las conclusiones y líneas futuras de trabajo.

2.-MATERIALES Y MÉTODOS

2.1.- BASE DATOS Y PREPROCESAMIENTO

Se utilizó la base de datos LIDC-IDRI, disponible en línea <https://www.cancerimagingarchive.net/collection/lidc-idri/> y anotada por expertos, reconocida por ser un estándar en la investigación de nódulos pulmonares. El análisis se centró en los escaneos de tomografía de los primeros 100 pacientes de dicha base de datos, constituidos por una muestra de 61 hombres y 39 mujeres, con una edad promedio de 66 años. Se empleó un conjunto total de 1165 cortes 2D que contenían 275 nódulos pulmonares, distribuidos por tamaño de la siguiente manera: 4 menores de 3 mm, 209 entre 3 y 10 mm, 57 entre 10 y 30 mm, y 5 mayores de 30 mm. Para la segmentación de cada nódulo, fue necesario obtener primero una imagen binaria corte por corte. Se probaron varios métodos de segmentación basados en umbrales, seleccionándose finalmente los métodos de umbral de máxima verosimilitud y el de mínimo error debido a sus resultados satisfactorios [12,13] y a que no introducirían la incertidumbre asociada al empleo de segmentación manual por diversos humanos.

Posteriormente, la región de interés (ROI) de cada nódulo fue seleccionada manualmente por un especialista altamente calificado en imágenes torácicas. Esta selección se comparó, utilizando la plataforma *PyIidic*, con las segmentaciones realizadas por los cuatro radiólogos expertos que anotaron la base de datos.

Una vez segmentado en 2D cada nódulo, fue reconstruido en 3D mediante la creación de un archivo *multiframe* que contenía todos los cortes 2D de la lesión en el mismo conjunto de datos. A los píxeles de la región de interés del nódulo 3D se les asignó un valor de 1 (blanco), y a los correspondientes al fondo, un valor de 0 (negro).

Las anotaciones de malignidad proporcionadas por los radiólogos en la base de datos presentan 5 grados de malignidad, siguiendo el estándar Lung-RADS, donde los grados 1, 2 y 3 se consideran con baja probabilidad de malignidad y los grados 4 y 5 se consideran sospechosos. Para obtener una etiqueta única, se promediaron las cuatro anotaciones y el valor resultante se redondeó al entero más cercano, el cual se utilizó como etiqueta final.

2.2.-EXTRACCIÓN DE CARACTERÍSTICAS

La extracción de características se llevó a cabo utilizando la librería *PyRadiomics*, manteniendo el formato DICOM de las imágenes. Se creó un archivo CSV para almacenar los nombres de las imágenes, las máscaras de segmentación y las

anotaciones de malignidad. Se obtuvo un conjunto inicial de 102 características radiómicas, distribuidas en las siguientes categorías: 14 características de forma, 18 características estadísticas de primer orden, 24 de la Matriz de concurrencia de nivel de gris (GLCM), 14 de la matriz de diferencia de nivel de gris (GLDM), 16 de la matriz de tamaño de zona de nivel de gris (GLSZM) y 16 de la matriz de longitud de ejecución de nivel de gris (GLRLM) [14,15].

Para reducir la dimensionalidad del vector de características, se implementó el método de selección basado en el Coeficiente de Correlación de Pearson. Este coeficiente se calculó entre la anotación de la base de datos (ya sea binaria 0/1 para el primer experimento o de 1 a 5 grados para el segundo) y el valor de cada característica radiómica.

El criterio de selección fue estricto: solo se seleccionaron aquellas características que presentaron un coeficiente de correlación de Pearson superior a 0.60, lo cual se considera una correlación fuerte para un intervalo de confianza del 95% [16]. En este tipo de análisis la escala plantea que de 0.0 a 0.1 la correlación (modularmente) es nula o muy débil, de 0.1 a 0.3, la correlación se considera débil, de 0.3 a 0.5 se considera moderada y de 0.5 a 1.0 se considera correlación alta o fuerte. Como se aprecia, se asumió un criterio de estar por encima del mínimo valor de correlación fuerte, para que ninguna característica de las seleccionadas estuviese en el margen moderada-fuerte, lo cual implicaría un aumento del tiempo de cómputo sin que el modelo gane en precisión de forma significativa. La Fig. 1 muestra la distribución de las 102 características para los umbrales modulares de 0 a > 0.6 para ambos experimentos.

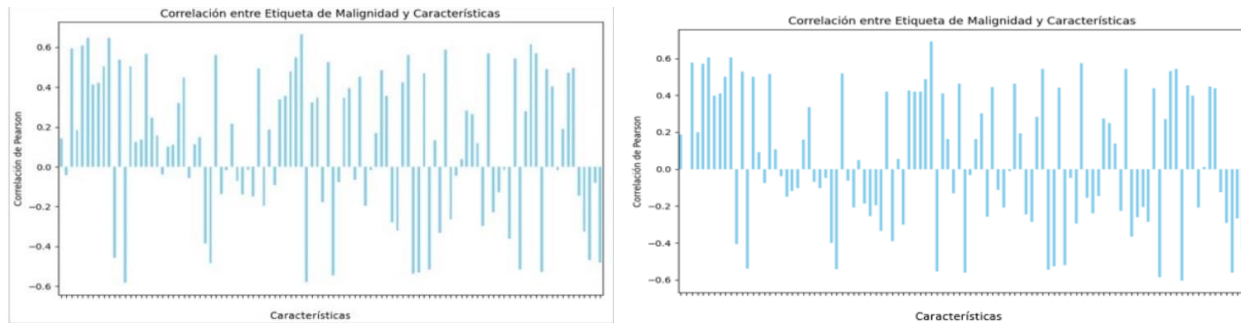


Figura 1

Correlación entre las características y las etiquetas para 2 grados de malignidad (izquierda) y 5 grados de malignidad (derecha).

El resultado del paso anterior fue obtener un vector de 5 características para el experimento binario (Tabla 1) y un vector de 3 características para el experimento no binario (Tabla 2). El uso exclusivo de estos vectores reducidos como entrada para los clasificadores eliminó la redundancia y simplificó significativamente la complejidad computacional del sistema.

Tabla 1
Características con Correlación de Pearson ($\rho > 0.60$) ($p < 0.05$) para Clasificación Maligna/Benigna (Etiquetado Binario)

Característica	ρ
original_shape_Maximum2DDiameterColumn	0.608
original_shape_Maximum2DDiameterRow	0.647
original_shape_MinorAxisLength	0.646
original_glcm_Imc1	0.664
original_gldm_DependenceEntropy	0.613

Tabla 2
Características con Correlación de Pearson ($\rho > 0.60$) ($p < 0.05$) para Clasificación de Cinco Grados de Malignidad (Etiquetado No Binario)

Característica	ρ
original_shape_Maximum2DDiameter Row	0.604
original_shape_MinorAxisLength	0.606
original_glcm_Imc1	0.692

Este enfoque univariado de selección actúa como un mecanismo de filtrado para la incertidumbre. Las características que no mantienen una correlación fuerte y directa con el resultado clínico son intrínsecamente inestables o redundantes, por lo que su descarte no solo reduce el vector de entrada y el costo computacional, sino que mejora la calidad de la información que alimenta a los clasificadores. Las figuras Fig. 2 y 3 muestran el procedimiento completo y el flujo de los datos.

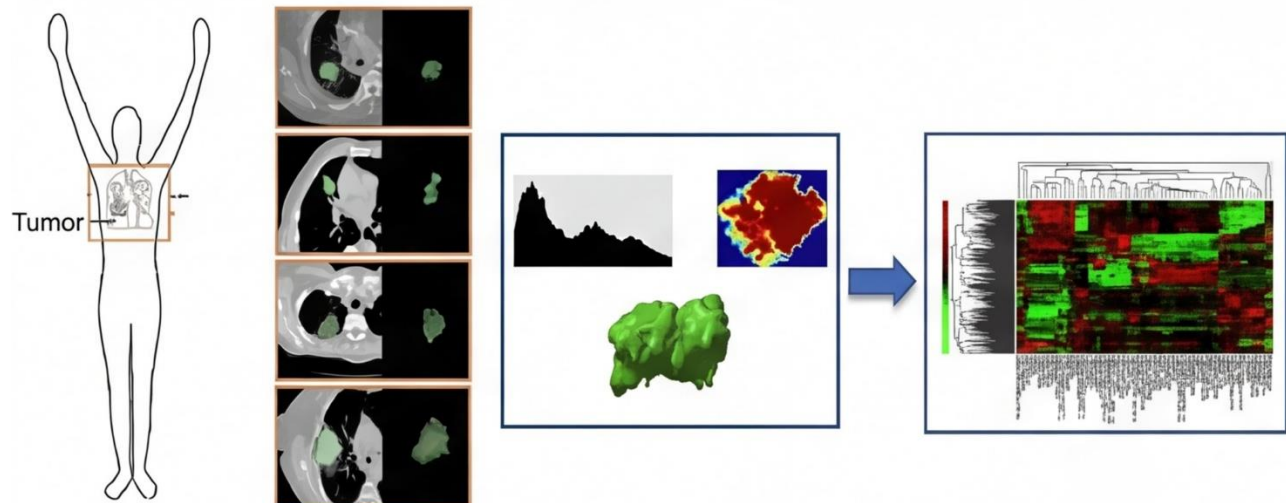


Figura 2

Procedimiento para la extracción de rasgos basados en radiómica, de izquierda a derecha se aprecia el nódulo del paciente sobre los cortes 2D de la CT, se reconstruye el nódulo en 3D y se le calculan 102 rasgos radiómicos sobre Pyradiomic

2.3-CLASIFICACIÓN DE LOS NÓDULOS

La clasificación se llevó a cabo utilizando dos algoritmos ampliamente aceptados: SVM y RF. Se empleó la librería *Scikit-Learn* de Python para su implementación. Para SVM se probaron tres tipos de *kernel*: lineal, radial y polinómicos de grados 2, 3 y 5, tomándose el parámetro de regularización $c=1$ en todos los casos, el cual lo trae implementado *Scikit-Learn* como base. El criterio de parada de las iteraciones se asumió cuando las pérdidas fuesen menores que 0.001. La beta del *kernel* radial también se tomó la que por base selecciona *Scikit-Learn*, es decir, 1. Para RF se varió solamente el número de árboles fuente (100, 50 y 20), su profundidad (*max_depth*) se ajustó a *None* y 10 para regular su complejidad.

El conjunto de datos se dividió en dos grupos: el 80% para el entrenamiento de los 10 modelos y el 20% restante para la validación. Las métricas utilizadas para validar y comparar los resultados fueron sensibilidad, precisión, exactitud (ACC) y especificidad [17].

Para el mejor modelo, se calculó la incertidumbre asociada al mismo, utilizando como métrica la Entropía de Shannon [18]. A modo de comparación de la propuesta se incluyó también correr el mejor modelo obtenido, haciendo previamente la reducción de características con un modelo LASSO, que es más complejo y no lineal.

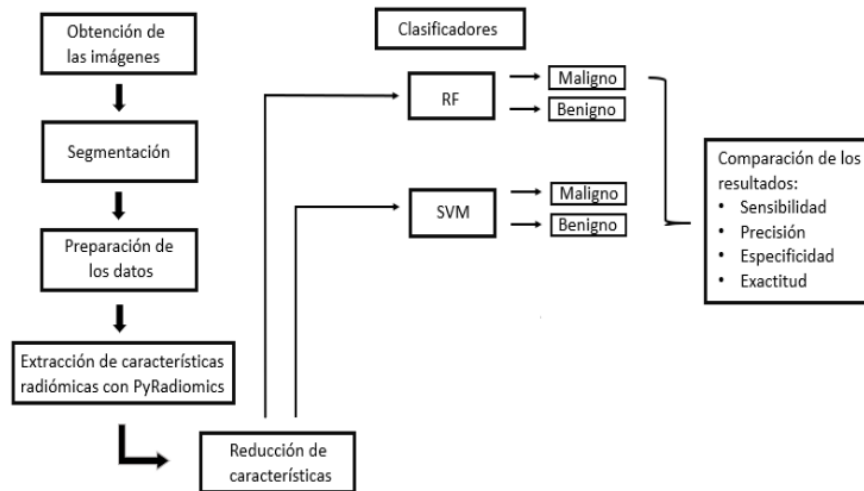


Figura 3
Diagrama de bloques del flujo de datos.

3.-RESULTADOS

El algoritmo de reducción basado en Pearson > 0.6 demostró ser altamente selectivo, reduciendo las 102 características iniciales a un vector de solo 5 o 3 atributos. El análisis de las Tablas 1 y 2 revela que las características seleccionadas con mayor correlación con la malignidad son consistentemente de forma y una característica de textura (GLCM).

La *original_shape_Maximum2DDiameterColumn* representa el diámetro máximo de la lesión. Se mide específicamente en el plano de la columna y ayuda a diferenciar entre subtipos de cáncer de pulmón o grados de agresividad, al capturar variaciones morfológicas que el ojo humano no cuantifica fácilmente. La *original_shape_Maximum2DDiameterRow* es una métrica que cuantifica la extensión física del nódulo en una orientación específica de la imagen. Ayuda a evaluar la irregularidad de los tumores. Valores altos en esta métrica suelen asociarse con lesiones más alargadas o de mayor tamaño, lo que se correlaciona en Oncología con una mayor agresividad o un estadio más avanzado del cáncer. El *original_shape_MinorAxisLength* es una característica que cuantifica el tamaño del nódulo basándose en su representación como un elipsoide. En Oncología, se utiliza para caracterizar el crecimiento tumoral. Lesiones con ejes menores desproporcionadamente grandes en comparación con el eje mayor pueden indicar un crecimiento infiltrativo o menos organizado, lo que ayuda a predecir la agresividad del cáncer. La *original_glcmm_Ic1* cuantifica la complejidad de la imagen mediante la información mutua de correlación. Es una de las métricas más potentes para caracterizar la heterogeneidad intra-tumoral, ya que se considera un biomarcador de agresividad. Sus valores específicos ayudan a predecir la respuesta al tratamiento y la supervivencia, puesto que los tumores más agresivos suelen presentar una arquitectura de niveles de gris más caótica. La *original_gldm_DependenceEntropy* es una métrica de textura avanzada perteneciente a la matriz de dependencia de niveles de gris (GLDM). Cuantifica la complejidad o el desorden en la distribución de las dependencias de los píxeles. Es un marcador clave para identificar variaciones biológicas dentro de una lesión que no son visibles al ojo humano.

En este trabajo se observó que las características: *original_shape_Maximum2DDiameterRow*, *original_shape_MinorAxisLength* y *original_glcmm_Ic1* fueron comunes a ambos tipos de clasificación (binaria y no binaria), lo que indica que el tamaño y la complejidad textural son descriptores prominentes del fenotipo tumoral en estos datos. Los resultados de rendimiento (métricas en porcentaje) para los modelos SVM en las clasificaciones binaria y no binaria se resumen en las Tablas 3 y 4.

Tabla 3
Métricas de Rendimiento (%) de los Modelos SVM Entrenados con Etiquetado Binario (5 características)

Modelo SVM	Precisión	Sensibilidad	Especificidad	ACC
Lineal kernel	92.31	92.31	97.62	96.36
Radial kernel	92.31	92.31	97.62	96.36
Polynomic G3 kernel	100	92.31	100	98.8
Polynomic G5 kernel	100	69.23	100	92.73
Polynomic G7 kernel	90.91	76.92	97.61	92.73

Tabla 4
Métricas de Rendimiento (%) de los Modelos SVM Entrenados con Etiquetado No Binario (3 características)

Modelo SVM	Precisión	Sensibilidad	Especificidad	ACC
Lineal kernel	68.75	78.57	87.80	85.45
Radial kernel	70.59	85.71	87.80	87.27
Polynomic G3 kernel	90.00	64.29	97.56	89.09
Polynomic G5 kernel	88.89	57.4	97.56	87.27
Polynomic G7 kernel	83.33	71.43	95.2	89.09

En general, el mejor desempeño se observó en la clasificación binaria. El modelo con *kernel* polinómico de grado 3 demostró el balance óptimo entre métricas, alcanzando el 100% de precisión y especificidad y una exactitud del 98.8%. Este modelo clasificó correctamente todos los casos benignos y solo tuvo un fallo en la clasificación de un nódulo maligno durante la validación. Los resultados para el clasificador RF se detallan en las Tablas 5 y 6.

Tabla 5
Métricas de Rendimiento (%) de los Modelos RF Entrenados con Etiquetado Binario (5 características)

Modelo RF	Precisión	Sensibilidad	Especificidad	ACC
n_estimators = 100 y max_depth = None	87.5	77.78	94.59	89.09
n_estimators = 50 y max_depth = None	87.5	77.78	94.59	89.09
n_estimators = 20 y max_depth = None	93.33	77.78	97.3	90.90
n_estimators = 100 y max_depth = 10	88.24	83.33	94.59	90.90
n_estimators = 50 y max_depth = 10	93.33	77.78	97.3	90.90
n_estimators = 20 y max depth = 10	84.21	88.89	91.89	90.90

Tabla 6
Métricas de Rendimiento (%) de los Modelos RF Entrenados con Etiquetado No Binario (3 características)

Modelo RF	Precisión	Sensibilidad	Especificidad	ACC
n_estimators = 100 y max_depth = None	94.44	89.47	97.22	94.55
n_estimators = 50 y max_depth = None	90	94.74	94.44	94.55
n_estimators = 20 y max_depth = None	89.47	89.47	94.44	92.72
n_estimators = 100 y max_depth = 10	85	89.47	91.67	90.91
n_estimators = 50 y max_depth = 10	85.71	94.74	91.67	92.72
n_estimators = 20 y max_depth = 10	85	89.47	91.67	90.91

Contrario a SVM, los mejores resultados del RF se obtuvieron en la clasificación no binaria (5 grados de malignidad), donde los modelos inicializados con 100 y 50 árboles alcanzaron una exactitud del 94.55%. En general el mejor modelo de todos los experimentos realizados fue el SVM polinómico grado 3. Para este modelo se obtuvo una incertidumbre de 0.4 considerada moderadamente baja [18].

Cuando se corre el mejor modelo propuesto, con la misma base de datos, pero reduciendo características por el método LASSO, como se propone en [10,12], se obtuvo una sensibilidad del 77 %, especificidad 83 %, ACC 81 % y precisión 72 %, valores todos por debajo de los obtenidos con la actual propuesta. En este caso la reducción fue de 102 a 18 características.

Un resultado crucial es la baja complejidad del modelo propuesto. El costo computacional de entrenar en general cada modelo (tanto SVM como RF) fue de solo 5 segundos en un entorno de hardware modesto (CPU: Intel® Core™ i5-4310M 2.70 GHz, RAM: 8 GB DDR4). Esta eficiencia es una consecuencia directa de la reducción drástica de la dimensionalidad, lo que hace que el sistema sea eficiente y escalable.

4.-DISCUSIÓN

La metodología de reducción de características basada en el Coeficiente de correlación de Pearson demostró ser altamente efectiva, lo que responde directamente a la necesidad de gestionar la redundancia de atributos. Al exigir una correlación fuerte (> 0.60) con la etiqueta de malignidad, el algoritmo elimina inherentemente aquellas características que son inestables o que aportan solo ruido, lo que se traduce en un vector de entrada concentrado y robusto.

Esta simplicidad metodológica es, paradójicamente, su principal fortaleza. Mientras que métodos más complejos como LASSO [10,11] intentan un enfoque multivariado para la selección de características y son más costosos computacionalmente; para este problema de clasificación de nódulos, el enfoque univariado propuesto demostró ser más efectivo y eficiente computacionalmente, al concentrarse únicamente en la relevancia directa de cada atributo con peso real con respecto a la anotación. La eliminación de características que no cumplen con el umbral de correlación previene la introducción de redundancias que podrían obstaculizar el rendimiento de los clasificadores. LASSO, al ser un modelo no lineal, requiere de más datos para maximizar la relación entre simplicidad y rendimiento. Si se reduce más el número de rasgos con este método, baja el rendimiento, si se aumenta puede introducir sobreajuste.

Hoy en día, los hospitales que trabajan con enfoque moderno y medicina de precisión, en sus departamentos de Radiología incluyen tres lectores, dos humanos y una inteligencia artificial (IA). Allí donde el ojo humano entrenado no ve hallazgos clínicos significativos, por limitaciones del sistema visual, o incluso limitaciones del monitor de visualización o pobre calidad de imagen, pero la IA reporta un patrón anormal, el caso será consecuentemente revisado. Muchas veces esa alerta oportuna cambia completamente el manejo del paciente y contribuye de forma temprana a salvarle la vida.

Hay estudios que plantean que los nódulos pulmonares, de pequeño tamaño (pocos mm) o con una localización compleja en el tórax, pasan a menudo desapercibidos [2], sin embargo, son la base del cáncer de pulmón. Por otra parte, la clasificación

en maligno o benigno, solo por simple visualización imagenológica es aún más compleja y subjetiva. Esto implica que hay que acudir a un invasivo procedimiento de biopsia. Ante estas circunstancias, los sistemas de detección automatizada mediante IA se han convertido en una herramienta muy útil para ayudar en este problema.

Los resultados mostrados en las Tablas 3 – 6, reportan valores por encima del umbral de visibilidad humana, que para nódulos menores de 3 mm y difícil localización en la región torácica, se reporta del 65 % en Radiología [2]. Particularmente los mejores modelos presentaron un buen balance entre todas las métricas, de modo que una no mejora a expensas de empeorar otra, lo cual es un aspecto a favor de encontrar modelos con estabilidad y bajo nivel de incertidumbre. La sensibilidad, que es el parámetro de mayor interés en el medio clínico, supera para el mejor modelo el 90 %, por lo que se arroja muy pocos falsos negativos. Por otro lado, la alta especificidad (100% con el mejor modelo SVM) y etiquetado binario, es particularmente relevante clínicamente, ya que significa que el sistema minimiza los falsos positivos, reduciendo también la clasificación errónea de nódulos benignos.

Los resultados de clasificación binaria (maligno-benigno) del mejor modelo SVM (ACC 98.8%) fueron notablemente superiores a los reportados en la literatura [19,20] para la misma base de datos LIDC-IDRI, utilizando enfoques más complejos. En la comparación con modelos basados en rasgos, los presentes resultados superan significativamente a los reportados en [7], que logró una ACC del 82.7% usando un modelo RF para seleccionar 150 características.

En la comparación con modelos que utilizan aprendizaje profundo, el rendimiento de la propuesta también superó enfoques basados en CNN como los reportados en [8,21], que obtuvieron una precisión por debajo del 80%. Esto valida que, en condiciones donde el volumen de datos es relativamente bajo, un sistema robusto de Radiómica y ML, con una selección de características eficiente, es más plausible y eficaz que modelos de DL, los cuales suelen requerir volúmenes de datos mucho mayores para evitar el sobreajuste.

En la comparación con otros métodos de reducción [9–12], a diferencia del filtrado por estabilidad (9), que solo alcanzó el 80% de ACC, el método propuesto no solo es más sencillo, sino que produce un clasificador de mayor rendimiento. Además, aunque el método con Wilcoxon [11] es tan simple como el propuesto, y logra reducir el vector de atributos también a solo 4 características, las métricas alcanzadas en el presente trabajo son superiores a las reportadas con dicho enfoque.

La Radiómica, cuando se aplica con rigor en la estandarización y selección de características, contribuye a reducir la incertidumbre diagnóstica [5,6,22]. El diagnóstico subjetivo de nódulos por parte de los radiólogos puede ser variable. Al seleccionar un pequeño conjunto de características biológicamente plausibles y fuertemente correlacionadas (como la longitud del eje menor o la correlación de nivel de gris), el método propuesto mitiga los efectos del ruido y de la inestabilidad de las características que provienen de la adquisición de la imagen o de la segmentación.

En este trabajo el valor de la incertidumbre de 0.44 se encuentra fuera de la zona de duda máxima (que es de 0.8 a 1.00) [18]. En clasificación general de 2 clases este valor se considera un nivel de confianza robusto y por tanto conduce a un modelo estable. En el contexto binario muestra que la SVM ha encontrado patrones claros en el dato evaluado. El modelo tiene una confianza del 91 % en su predicción.

Aunque la propuesta arrojó resultados prometedores, es aún un estudio de laboratorio. Requiere la optimización de todos los hiperparámetros del modelo y el entrenamiento con una base de datos de diferente origen, para probar su nivel de generalización y garantizar que esté listo para utilizarse en la práctica de rutina clínica

5.-CONCLUSIONES

La contribución metodológica presentada para reducir la dimensionalidad del problema mediante el Coeficiente de correlación de Pearson ha demostrado ser altamente efectiva. Este enfoque simple permitió eliminar redundancias, lo que resultó en una reducción drástica del costo computacional de entrenamiento, a solo 5 segundos por modelo.

Adicionalmente, la metodología de reducción de características propuesta contribuye a reducir la incertidumbre del modelo, al seleccionar explícitamente solo aquellas características robustamente correlacionadas con el resultado clínico, generando un clasificador más fiable y estable.

Se demostró que los clasificadores de ML (SVM y RF), alimentados con un número relativamente bajo de datos y un vector reducido de características radiómicas, son capaces de clasificar las clases de nódulos pulmonares en benigno y maligno con una precisión diagnóstica superior incluso a modelos de DL que requieren un gran volumen de datos.

Los valores de sensibilidad, especificidad y exactitud obtenidos en la validación superan el 90%, y en el caso del SVM polinómico de grado 3 para clasificación binaria, alcanzan una exactitud del 98.8%. Estos resultados son iguales o superiores

a los obtenidos por otros sistemas internacionales que han utilizado enfoques más complejos, entrenándose con la misma base de datos LIDC-IDRI.

Como continuidad futura se propone realizar una prueba externa con otra base de datos, para comprobar el poder de generalización de la propuesta, así como optimizar todos los hiperparámetros para los diversos modelos a fin de realizar un ajuste más fino de los mismos. Este algoritmo formará parte de la versión 2.0 del software De-Lung Nodules -Rx, registrado en CENDA (Registro 3129-12-23, 20/Dic/2023), el cual se encuentra actualmente en implementación en la práctica clínica de rutinaria en varios hospitales.

AGRADECIMIENTOS

Los autores agradecen a la Agencia de Energía Nuclear y Tecnologías Avanzadas de Cuba, código PS211LH02, Oficina de Gestión de Fondos y Proyectos Internacionales código PN223LH004-039 y el Programa Territorial del CITMA sobre Salud Humana, código PT241VC001-009, por el financiamiento parcial de este trabajo.

REFERENCIAS

1. Loverdos K, Fotiadis A, Kontogianni C, Iliopoulou M, Gaga M. Lung nodules: A comprehensive review on current approach and management. *Ann Thorac Med*. 2019;14(4):226–38.
2. Wang J, Lin L, Zhao S, Wu X, Wu S. Research progress on computed tomography image detection and classification of pulmonary nodule based on deep learning. *J Biomed Eng*. 2019 Aug 25;36(4):670–6.
3. Chen P, Li F, Wu C. Research on Intrusion Detection Method Based on Pearson Correlation Coefficient Feature Selection Algorithm. *J Phys: Conf Ser* [Internet]. 2021 Jan [cited 2026 Feb 2];1757(1). Available from: <https://doi.org/10.1088/1742-6596/1757/1/012054>
4. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012 Nov;30(9):1234–48.
5. Pavic M, Bogowicz M, Würms X, Glatz S, Finazzi T, Riesterer O, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. *Acta Oncol*. 2018 Aug;57(8):1070–4.
6. Park JE, Park SY, Kim HJ, Kim HS. Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. *Korean J Radiol*. 2019 July;20(7):1124–37.
7. Wang J, Xia Liu, Dong D, Song J, Min Xu, Yali Zang, et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016 Aug;2016:1272–5.
8. Kumar D, Chung AG, Shaifee MJ, Khalvati F, Haider MA, Wong A. Discovery Radiomics for Pathologically-Proven Computed Tomography Lung Cancer Prediction. In: Karray F, Campilho A, Cheriet F, editors. Cham: Springer International Publishing; 2017 [cited 2026 Feb 2]. p. 54–62. (Lecture Notes in Computer Science; vol. 10317). Available from: https://link.springer.com/10.1007/978-3-319-59876-5_7
9. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting Malignant Nodules from Screening CT Scans. *Journal of Thoracic Oncology*. 2016 Dec;11(12):2120–8.
10. Choi W, Oh JH, Riyahi S, Liu CJ, Jiang F, Chen W, et al. Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer. *Med Phys*. 2018 Apr;45(4):1537–49.
11. Chen CH, Chang CK, Tu CY, Liao WC, Wu BR, Chou KT, et al. Radiomic features analysis in computed tomography images of lung nodule classification. *PLoS ONE* [Internet]. 2018 Feb 5;13(2). Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5798832/>
12. Mao L, Chen H, Liang M, Li K, Gao J, Qin P, et al. Quantitative radiomic model for predicting malignancy of small solid pulmonary nodules detected by low-dose CT screening. *Quant Imaging Med Surg*. 2019 Feb;9(2):263–72.

13. Kurita T, Otsu N, Abdelmalek N. Maximum likelihood thresholding based on population mixture models. *Pattern Recognition*. 1992 Oct 1;25(10):1231–40.
14. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med*. 2020 Apr;61(4):488–95.
15. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in lung cancer screening. *Transl Lung Cancer Res*. 2021 Feb;10(2):1186–99.
16. Rodgers JL, Nicewander WA. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*. 1988 Feb;42(1):59.
17. Hossin M, Sulaiman M.N. A Review on Evaluation Metrics for Data Classification Evaluations. *IJDKP*. 2015 Mar 31;5(2):01–11.
18. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J* 1948;27(3):379–423.
19. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012 Mar;48(4):441–6.
20. Wei K, Su H, Zhou G, Zhang R, Cai P, Fan Y, et al. Potential Application of Radiomics for Differentiating Solitary Pulmonary Nodules. *OMICS J Radiol*. 2016;5(2):1000218.
21. Shi W, Hu Y, Chang G, Qian H, Yang Y, Song Y, et al. Development of a clinical prediction model for benign and malignant pulmonary nodules with a CTR $\geq 50\%$ utilizing artificial intelligence-driven radiomics analysis. *BMC Med Imaging*. 2025 Jan 17;25(1):21.
22. Lambert B, Forbes F, Doyle S, Dehaene H, Dojat M. Trustworthy clinical AI solutions: A unified review of uncertainty quantification in Deep Learning models for medical image analysis. *Artif Intell Med*. 2024 Apr;150:102830.

CONFLICTO DE INTERESES

Ninguno de los autores manifestó la existencia de posibles conflictos de intereses que debieran ser declarados en relación con este artículo.

CONTRIBUCIONES DE LOS AUTORES

Rachel Abreu-Llanes: Investigación, Metodología, Software, Visualización, Redacción- borrador original

Marlen Perez-Diaz: Adquisición de fondo, Conceptualización, Análisis formal, Recursos, Supervisión, Visualización, Redacción- borrador original

Yusely Ruiz-Gonzalez: Adquisición de fondo, Administración de proyecto, Conceptualización, Supervisión, Redacción – revisión y edición

Deborah Galpert-Cañizares: Análisis formal, Supervisión, Redacción – revisión y edición

Ruben Orozco-Morales: Conceptualización, Validación -Verificación,

AUTORES

Rachel Abreu-Llanes: Se graduó de Ingeniería en Automática en la Universidad Central “Marta Abreu” de Las Villas (UCLV) Cuba en 2023. Actualmente estudia la Maestría en Ingeniería Eléctrica en el campo de Procesamiento Digital de Señales de la Universidad Nacional Autónoma de México. Se encuentra vinculada al Laboratorio avanzado de Procesamiento de Imágenes de la Facultad de Ingeniería de dicha universidad. Su campo de investigación es el procesamiento de imágenes médicas para la detección y análisis de patologías.

Correo electrónico: racheabreul@gmail.com, ORCID [0000-0008-9336-5901](https://orcid.org/0000-0008-9336-5901)

Marlen Pérez Díaz. Se graduó de Ing. Física Nuclear en el INSTEC, obtuvo el DrC Físicas por el Tribunal Nacional de Cuba y es especialista en Física médica. Trabaja para la Universidad Central Marta Abreu de las Villas, Departamento de Automática. Villa Clara, Cuba y el Instituto de Física Nuclear de Italia (INFN). 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.

Correo electrónico: mpdiaz@infn.it , ORCID [0000-0002-3706-9154](https://orcid.org/0000-0002-3706-9154)

Yusely Ruiz-González: se graduó de Ingeniera en Telecomunicaciones y Electrónica en la Universidad Central “Marta Abreu” de Las Villas (UCLV), Cuba, en 2002 y obtuvo el grado de Dra. C. en Ingeniería Biomédica en la Universidad de Zhejiang, República Popular China, en 2009. Actualmente se desempeña como jefa del Laboratorio de Procesamiento y Análisis de Imágenes y Señales del Centro de Investigaciones de la Informática (CII), Facultad de Matemáticas, Física y Computación, UCLV, Villa Clara, Cuba. Su campo de investigación se vincula al procesamiento digital de señales e imágenes, visión computacional e inteligencia artificial.

Correo electrónico: yuselyr@uclv.edu.cu , ORCID [0000-0002-6011-1947](https://orcid.org/0000-0002-6011-1947)

Deborah Galpert-Cañizares: se graduó de Lic. en Cibernética Matemática en en la Universidad Central “Marta Abreu” de Las Villas (UCLV), Cuba, en 1992 y obtuvo el grado de Dra. C. en Ciencias Informáticas en la UCLV en 2016. Actualmente se desempeña como jefa del Laboratorio de Inteligencia Artificial del Centro de Investigaciones de la Informática (CII), Facultad de Matemáticas, Física y Computación, UCLV, Villa Clara, Cuba. Su campo de investigación se vincula a la inteligencia artificial, bioinformática y análisis de datos.

Correo electrónico: deborah@uclv.edu.cu , ORCID [0000-0002-5222-3324](https://orcid.org/0000-0002-5222-3324)

Ruben Orozco-Morales: Se graduó de Ing. Electrónico (1981) en la Universidad Central “Marta Abreu” de las Villas (UCLV), Cuba. Master en Redes y Sistemas de Telecomunicaciones (1994) por la Universidad Politécnica de Madrid, España. Doctor en Ciencias Técnicas (1998) por la UCLV. Ha trabajado en proyectos vinculados con el procesamiento digital de señales e imágenes.

Correo electrónico: rorozco@uclv.edu.cu, ORCID [0000-0002-6240-1569](https://orcid.org/0000-0002-6240-1569)



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/)