

Implementación de MobileNetV2 para detección de cáncer de pulmón mediante la inteligencia artificial para el cómputo móvil

MobileNetV2 implementation for lung cancer detection through artificial intelligence for mobile computing

O. D. Arellano-Magdaleno¹*, M. A. Ramos-Corchado¹, V. Muñoz-Jiménez¹
(*oarellanomagdaleno@gmail.com)

Recibido: 29 de abril de 2026 Aceptado: 15 de junio de 2026

RESUMEN

El cáncer de pulmón es uno de los principales problemas de salud a nivel mundial reportando 1.18 millones de fallecimientos al año. Frecuentemente, esta enfermedad se diagnostica en etapas avanzadas, lo que reduce las posibilidades de un tratamiento exitoso. Entre los principales factores de riesgo se encuentra el consumo de tabaco; sin embargo, cuando se identifica de manera oportuna, es posible mejorar el pronóstico y prolongar la supervivencia de los pacientes. Los métodos tradicionales de diagnóstico se basan principalmente en técnicas de imagen, como la radiografía de tórax y la tomografía computarizada (CT), cuya interpretación requiere experiencia clínica para reconocer alteraciones asociadas con etapas iniciales de la enfermedad. Además, en

regiones con acceso limitado a personal médico especializado, el análisis oportuno de estos estudios puede verse afectado, dificultando un diagnóstico temprano. En este contexto, el uso de inteligencia artificial mediante algoritmos como MobileNetV2 permite implementar herramientas de apoyo capaces de identificar patrones asociados con cáncer de pulmón de forma rápida y precisa. En este estudio se obtuvo una exactitud del 91.0 %, resultado comparable con el desempeño de arquitecturas más robustas reportadas en la literatura.

Palabras clave: *cáncer de pulmón, tomografía, red neuronal convolucional, inteligencia artificial, MobileNetV2*

¹Facultad de Ingeniería, Universidad Autónoma del Estado de México.



ABSTRACT

Lung cancer is one of the leading global health problems, causing approximately 1.18 million deaths each year. This disease is frequently diagnosed at advanced stages, reducing the likelihood of successful treatment. Tobacco consumption is among the main risk factors; however, when detected at an early stage, lung cancer can be treated, improving prognosis and extending patient survival. Traditional diagnostic methods are primarily based on imaging techniques, such as chest radiography and computed tomography (CT), whose interpretation requires clinical expertise to recognize abnormalities associated with the early stages of the disease. Furthermore, in regions with limited access to specialized medical personnel, the timely analysis of these studies may be affected, hindering early diagnosis. In this context, the use of artificial intelligence through algorithms such as MobileNetV2 enables the implementation of support tools capable of identifying lung cancer-related patterns rapidly and accurately. In this study, an accuracy of 91.0 % was achieved, a result comparable to the performance of more robust architectures reported in the literature.

Keywords: *lung cancer, tomography, convolutional neural network, Artificial Intelligence, MobileNetV2*

INTRODUCCIÓN

El cáncer de pulmón es una enfermedad caracterizada por el crecimiento anormal de células en el tejido pulmonar del tracto respiratorio [1]. Esta patología representa un problema de salud pública a nivel mundial debido a que provoca aproximadamente 1.18 millones de fallecimientos cada año [2]. Actualmente, se distinguen diferentes tipos de cáncer de pulmón. De acuerdo con Cajape-González et al. [1], aproximadamente el 85 % de los casos corresponden al carcinoma de células no pequeñas, mientras que el porcentaje restante incluye adenocarcinomas, carcinomas de células grandes y carcinomas escamosos. Los carcinomas son tumores malignos que se originan en tejidos epiteliales y representan alrededor del 90 % de los casos de cáncer en humanos [3].

El cáncer de pulmón suele diagnosticarse en etapas avanzadas debido a que sus manifestaciones clínicas iniciales tienden a ser silenciosas. Entre los principales factores de riesgo se encuentran el consumo de tabaco, la contaminación atmosférica y la exposición a sustancias nocivas. Para su evaluación clínica, esta enfermedad se clasifica mediante el sistema Tumor, Ganglio y Metástasis (TNM), el cual permite establecer pronósticos precisos y orientar la atención médica adecuada [4].

Aunque el cáncer de pulmón puede ser controlado si se detecta de manera temprana, los métodos tradicionales de diagnóstico, como las radiografías de tórax y las tomografías computarizadas (CT), requieren tiempo y dependen en gran medida de la experiencia del especialista [5]. Además, en regiones

con acceso limitado a personal médico especializado, el análisis oportuno de estos estudios puede verse afectado, dificultando un diagnóstico temprano. En este contexto, el diagnóstico asistido por inteligencia artificial (IA) surge como una alternativa para mejorar los procesos de análisis de imágenes médicas. Los modelos de aprendizaje automático son capaces de procesar grandes volúmenes de información e identificar patrones complejos presentes en las imágenes de forma rápida y consistente. Diversos estudios reportados en la literatura han abordado esta problemática mediante técnicas de aprendizaje automático y aprendizaje profundo. Entre las estrategias más utilizadas se encuentran las redes neuronales convolucionales (CNN), la transferencia de aprendizaje y el procesamiento digital de imágenes (PID), alcanzando resultados competitivos en términos de precisión diagnóstica. Estas características favorecen el desarrollo de soluciones accesibles para entornos con infraestructura tecnológica limitada y acceso restringido a personal médico especializado. Además, la Organización Mundial de la Salud (OMS) proyecta un incremento en los casos de cáncer de pulmón durante los próximos años, lo que refuerza la necesidad de desarrollar herramientas que contribuyan a una detección más oportuna de la enfermedad [1]. En este trabajo se emplea la IA para desarrollar una CNN orientada a la detección de cáncer de pulmón en imágenes CT. Se utiliza la arquitectura MobileNetV2, seleccionada por su equilibrio entre precisión y eficiencia computacional. Gracias al uso de convoluciones separables en profundidad y bloques residuales invertidos, esta arquitectura reduce significativamente el número de parámetros y operaciones requeridas en comparación con redes más complejas, favoreciendo su implementación en dispositivos con recursos limitados [6]. La originalidad de la propuesta radica en el uso de una arquitectura ligera capaz de alcanzar un desempeño competitivo en tareas de clasificación de cáncer de pulmón, en comparación con arquitecturas más complejas, facilitando su aplicación en entornos con infraestructura tecnológica restringida.

TRABAJOS RELACIONADOS

Patil et al. [5] proponen un enfoque basado en CNN para la detección de cáncer de pulmón a partir de imágenes de CT. El estudio considera etapas de preprocesamiento y entrenamiento de modelos de aprendizaje profundo para diferenciar imágenes con y sin presencia de cáncer. El desempeño se evalúa mediante métricas como exactitud, sensibilidad, especificidad y F1-score. Los resultados muestran que las CNN pueden distinguir eficazmente entre imágenes con presencia de cáncer e imágenes sin evidencia de cáncer, evidenciando su potencial como herramienta de apoyo al diagnóstico temprano.

Mohamed et al. [7] proponen un método para la detección de cáncer de pulmón en imágenes CT mediante un modelo híbrido que combina CNN con el algoritmo metaheurístico Ebola Optimization Search Algorithm (EOSA). El enfoque integra PID para la extracción de características y clasificación de casos normales, benignos y malignos, mientras que EOSA optimiza los parámetros de la red neuronal. Utilizando el conjunto de datos IQ-OTH/NCCD, el modelo alcanzó una exactitud del 93.21 %,

sensibilidad de 90.71 %, especificidad de 1.0, precisión de 1.0 y F1-score de 92.72 %.

Raut et al. [8] presentan un sistema para la detección de cáncer de pulmón basado en PID y aprendizaje automático aplicado a imágenes de CT. El método incluye etapas de preprocesamiento para facilitar la identificación de regiones sospechosas de tumor. Posteriormente, las características extraídas son utilizadas por un árbol de decisión C4.5 para clasificar las imágenes como cancerosas o normales. El sistema obtuvo una exactitud aproximada del 78 %.

Hamim et al. [9] presentan un estudio para la detección de cáncer de pulmón mediante el análisis de datos clínicos estructurados. El enfoque emplea algoritmos de aprendizaje automático para clasificación binaria, incluyendo Máquinas de Vectores de Soporte (SVM), Regresión Logística, Gradient Boosting y CatBoost. Como parte del preprocesamiento, se realiza la gestión de valores nulos, la codificación de variables categóricas y el submuestreo aleatorio para corregir el desbalance de clases. El modelo SVM obtuvo una exactitud del 95 %.

Tiwari et al. [10] proponen un enfoque basado en aprendizaje profundo para la detección y clasificación de nódulos pulmonares en imágenes de CT. El método combina una arquitectura CMixNet tridimensional con Faster R-CNN para la detección de nódulos y Gradient Boosting Machine (GBM) para su caracterización. Asimismo, emplea una arquitectura tipo U-Net para la segmentación de regiones de interés y la reducción de falsos positivos. Utilizando el conjunto de datos LIDC-IDRI, el sistema alcanzó una sensibilidad del 94 % y una especificidad del 91 %.

Ribeiro y Sousa [11] presentan un estudio para la clasificación de cáncer de pulmón utilizando los conjuntos de datos IQ-OTH/NCCD y LIDC-IDRI. El enfoque combina arquitecturas CNN como AlexNet, ResNet-50 y EfficientNet con la Transformada Wavelet (coif5) y el filtro CLAHE para el preprocesamiento de imágenes. Los modelos alcanzaron una exactitud promedio de 96.12 %.

Wankhade y Vigneshwari [12] proponen un método híbrido basado en 3D-CNN, RNN y U-Net para la detección de cáncer de pulmón utilizando imágenes de CT del conjunto de datos LUNA16. El modelo alcanzó una exactitud del 95 %, demostrando la efectividad del aprendizaje profundo para la identificación de nódulos pulmonares malignos. La Tabla 1 presenta un resumen de los principales trabajos relacionados en el periodo comprendido entre 2021 y 2025.

De acuerdo con los trabajos analizados y con el estudio de Bocanegra-Valencia et al. [13], las CNN constituyen la técnica más utilizada para la detección de cáncer de pulmón mediante imágenes CT, alcanzando resultados competitivos en términos de precisión. No obstante, la mayoría de las investigaciones reportadas se enfocan en mejorar el desempeño de clasificación mediante arquitecturas profundas o modelos híbridos con un elevado costo computacional, tales como ResNet, EfficientNet, 3D-CNN o combinaciones de CNN con algoritmos de optimización y segmentación avanzada. En

Tabla 1. Resumen de trabajos relacionados sobre la detección de cáncer de pulmón.

Autor	Técnicas utilizadas	Resultados	Bases de datos
Patil et al., (2025)	Normalización de imágenes, Rotación, Escalado, Volteo de imágenes, CNN	Exactitud: 92 %, Precisión: 94 %, Sensibilidad: 90 %, Especificidad: 95 %, F1-score: 93 %	NLST
Ribeiro y Sousa (2024)	AlexNet, ResNet-50, EfficientNet, Transformada Wavelet y filtro CLAHE	Exactitud: 96.12 %	IQ-OTH/NCCD y LIDC-IDRI
Mohamed al., (2023)	Transformada Wavelet, CLAHE, CNN, EOSA, EOSA-CNN	Exactitud: 93.21 %, Precisión: 1.0, Sensibilidad: 90.71 %, Especificidad: 1.0, F1-score: 92.72 %	IQ-OTH/NCCD
Wankhade y Vigneshwari (2023)	Segmentación, Normalización, 3D-CNN + RNN	Exactitud: 95 %, Sensibilidad: 87 %, Especificidad: 90 %	LIDC-IDRI y LUNA16
Hamim et al., (2023)	Eliminación de valores nulos, Label Encoding, Random Under-Sampling, SVM, Regresión Logística, KNN, Naive Bayes, Árbol de Decisión, Random Forest, Gradient Boosting, XGBoost, CatBoost, Grid Search	Exactitud: 95 %, Precisión: 0.94, Sensibilidad: 1.00, F1-score: 95 %	Registros médicos reales numéricos y categóricos
Tiwari et al., (2022)	Normalización, Downsampling, Zero-centering, Segmentación pulmonar, U-Net 3D CNN (CMixNet), Faster R-CNN, GBM	Sensibilidad: 94 %, Especificidad: 91 %	LIDC-IDRI
Raut et al., (2021)	Canny edge detection, Sobel edge detection, Árbol de decisión C4.5	Exactitud: ~78 %	CT scans de pulmones de 61 pacientes

contraste, pocos estudios analizan la viabilidad de implementar modelos ligeros para entornos con recursos computacionales limitados o con acceso restringido a personal médico especializado. En este contexto, la presente investigación propone una metodología basada en MobileNetV2, una arquitectura diseñada para dispositivos con capacidad de procesamiento reducida, complementada con técnicas de PID para la clasificación de estudios CT en tres categorías: hallazgos benignos, malignos y pulmones sin evidencia de cáncer. De esta manera, la propuesta busca combinar un desempeño competitivo en la clasificación con una menor demanda computacional, favoreciendo su implementación en sistemas de cómputo móvil y plataformas con infraestructura limitada.

MATERIALES Y MÉTODOS

Para el desarrollo del modelo de clasificación de cáncer de pulmón se integraron técnicas de PID y aprendizaje predictivo. Estas técnicas permitieron mejorar la calidad de las imágenes de entrada, extraer características relevantes y entrenar una red neuronal capaz de distinguir entre 3 categorías de CT. La investigación presentada en este artículo se estructura en dos etapas principales:

1. Preprocesamiento de imágenes: aplicación de métodos de PID para mejorar la calidad del conjunto de datos (dataset) que es utilizado como mecanismos de entrenamiento de la CNN.
2. Aprendizaje predictivo: implementación de una red neuronal MobileNetV2 que permite clasificar las CT en tomografías sin evidencia de cáncer, hallazgos benignos y malignos.

Preprocesamiento del conjunto de datos

El dataset utilizado para entrenar la red MobileNetV2 está compuesto por imágenes médicas de tomografía computarizada, por lo que se aplicaron técnicas de PID para mejorar su calidad. Posteriormente, se dividió en 80 % para entrenamiento y 20 % para validación, empleando este último subconjunto para evaluar el desempeño del modelo y calcular las métricas reportadas. Las etapas del preprocesamiento aplicadas a las imágenes se describen a continuación.

Conversión a escala de grises

En la etapa de conversión a escala de grises, a cada píxel de la imagen se le asigna un valor de intensidad comprendido entre 0 y 255 [14]. Las tonalidades rojas (R), verde (G) y azul (B) se combinan para obtener una tonalidad de gris de acuerdo con la Ec. (1):

$$y = (R \times 0,3) + (G \times 0,59) + (B \times 0,11) \quad (1)$$

Filtro de desenfoque Gaussiano

Es una técnica de suavizado que difumina la imagen mediante la convolución del mapa de bits con una gaussiana 2D, la cual se define en la Ec. (2):

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

Donde: x e y son las coordenadas del píxel; σ representa la desviación estándar de la gaussiana. El desenfoque permite reducir detalles y ruido presentes en la imagen, facilitando la detección de contornos [15].

Segmentación OTSU

Es un método automático de umbralización que selecciona el umbral óptimo mediante la maximización de la varianza interclase entre el fondo y el objeto de interés [16]. De esta manera, aumenta la diferencia estadística entre ambas regiones y se minimizan los errores de clasificación de píxeles, mejorando la precisión de la segmentación [17].

a. Probabilidad de cada nivel de gris, Ec. (3):

$$B(i) = \frac{a(i)}{M \times N} \quad (3)$$

Donde: $a(i)$ es el número de píxeles con nivel de gris i , y $M \times N$ es el tamaño de la imagen.

b. Proporción de fondo y primer plano, Ec. (4) y Ec. (5):

$$h_0 = \sum_{i=0}^t B(i) \quad (4)$$

$$h_1 = \sum_{i=t+1}^{L-1} B(i) \quad (5)$$

c. Media de niveles de gris por clase, Ec. (6) y Ec. (7):

$$p_0 = \frac{\sum_{i=0}^t i \cdot B(i)}{h_0} \quad (6)$$

$$p_1 = \frac{\sum_{i=t+1}^{L-1} i \cdot B(i)}{h_1} \quad (7)$$

d. Media global de la imagen, Ec. (8):

$$\rho_T = \sum_{i=0}^{L-1} i \cdot B(i) \quad (8)$$

e. Varianza inter-clase, Ec. (9):

$$\sigma_B^2 = h_0(\rho_0 - \rho_T)^2 + h_1(\rho_1 - \rho_T)^2 \quad (9)$$

f. Selección del umbral óptimo, Ec. (10):

$$t^* = \arg \max \sigma_B^2(t) \quad (10)$$

Normalización Min-Max

Es una técnica de preprocesamiento que transforma los valores de un conjunto de datos a un rango específico, comúnmente entre 0 y 1 [18]. Esta normalización se define como la Ec. (11):

$$v_{\text{norm}_i} = \frac{v_i - v_{\text{mín}}}{v_{\text{máx}} - v_{\text{mín}}} \quad (11)$$

donde v_i representa el valor original, $v_{\text{mín}}$ y $v_{\text{máx}}$ son los valores mínimo y máximo del dataset.

Erosión y Dilatación

La dilatación morfológica genera un efecto de expansión o ampliación en las regiones de la imagen. Por su parte, la erosión morfológica produce un efecto de adelgazamiento o contracción de las regiones, eliminando píxeles en sus contornos [7]. El proceso de erosión y dilatación se define como la Ec. (12) y la Ec. (13):

$$Y = A \ominus B = \{x, y \mid (B)_{xy} \subseteq A\} \quad (12)$$

$$Y = A \oplus B = \{y : B(y) \cap A \neq \emptyset\} \quad (13)$$

Donde: Y es la imagen binaria resultante, B es el operador morfológico (elemento estructurante) y A es la imagen original por procesar.

Filtro CLAHE

El filtro CLAHE actúa en pequeñas regiones de la imagen, aplicando la técnica dentro de cada bloque [11] [19]. Sus etapas son las siguientes:

- División de la imagen en bloques.
- Cálculo del histograma local para cada bloque.
- Aplicación del límite de recorte al histograma.
- Transformación del histograma ajustado en una función de distribución acumulada (CDF).
- Mapeo de nuevas intensidades para los píxeles con base en la CDF.
- Interpolación entre bloques adyacentes para suavizar la transición.
- Reconstrucción de la imagen con el contraste local mejorado.

Para estos pasos se considera la realización de la normalización del histograma, dada por la Ec. (14):

$$H'(i) = \min\{H(i), clip_limit\} + \frac{\Sigma(H(i) - clip_limit)}{N} \quad (14)$$

donde $H(i)$ es el valor del histograma en la intensidad i , $clip_limit$ es el valor máximo permitido, y N es el número total de *bins* del histograma. Además, se realiza el mapeo de intensidades utilizando la Ec. (15):

$$I'(x, y) = CDF(I(x, y)) \times (L - 1) \quad (15)$$

donde $I(x, y)$ es la intensidad original del píxel, $I'(x, y)$ es la nueva intensidad del píxel, CDF es la función de distribución acumulada del histograma ajustado, y L es el número de niveles de intensidad.

Aprendizaje predictivo

En esta etapa se implementa un modelo basado en redes neuronales convolucionales con el objetivo de identificar patrones relevantes en las imágenes previamente procesadas.

Transformada Wavelet

Es una técnica utilizada para descomponer señales, como imágenes, en componentes de diferentes frecuencias. En el procesamiento de imágenes, permite analizar y procesar detalles a distintas escalas,

capturando tanto características globales como locales de la imagen [20]. Para el análisis bidimensional, se emplea la transformada de wavelet discreta 2D (2D-DWT), la cual descompone la imagen mediante filtros pasa-baja y pasa-altas en las direcciones horizontal y vertical, generando 4 subbandas denominadas LL, LH, HL y HH, que representan la aproximación y los detalles de alta frecuencia de la imagen [21]. Matemáticamente, esta descomposición se define como la Ec. (16), Ec. (17), Ec. (18) y Ec. (19):

$$LL(k, l) = \sum_{m=0}^{\frac{N}{2}-1} \sum_{n=0}^{\frac{N}{2}-1} x(2m, 2n) h(k-m) h(l-n) \quad (16)$$

$$LH(k, l) = \sum_{m=0}^{\frac{N}{2}-1} \sum_{n=0}^{\frac{N}{2}-1} x(2m, 2n+1) g(k-m) h(l-n) \quad (17)$$

$$HL(k, l) = \sum_{m=0}^{\frac{N}{2}-1} \sum_{n=0}^{\frac{N}{2}-1} x(2m+1, 2n) h(k-m) g(l-n) \quad (18)$$

$$HH(k, l) = \sum_{m=0}^{\frac{N}{2}-1} \sum_{n=0}^{\frac{N}{2}-1} x(2m+1, 2n+1) g(k-m) g(l-n) \quad (19)$$

donde (k, l) representan las coordenadas de los coeficientes de salida en el dominio transformado, $x(m, n)$ representa los valores de los píxeles de la imagen de entrada, $h(k)$ y $g(k)$ representan los filtros de escalamiento y wavelet, respectivamente, y $\frac{N}{2}$ representa la reducción del tamaño del dominio transformado debido al proceso de submuestreo.

Red Neuronal MobileNetV2

Es una CNN propuesta por Google en 2018, diseñada para dispositivos móviles y entornos con recursos limitados. Esta arquitectura reduce significativamente el número de operaciones y el consumo de memoria sin afectar de manera considerable la precisión del modelo [22]. MobileNetV2 introduce dos conceptos principales: la convolución separable en profundidad y la estructura residual invertida. La primera descompone la convolución tradicional en operaciones *depthwise* y *pointwise*, una convolución *depthwise* (DW), que aplica un filtro independiente a cada canal de entrada, y una convolución *pointwise* (PW) de tamaño 1×1 , que combina linealmente los canales para generar nuevas características. Esta estrategia reduce la complejidad computacional respecto a la convolución tradicional. Considerando un mapa de características de entrada de tamaño $w_{in} \times h_{in}$ con M canales, un kernel de tamaño $k \times k$ y N canales de salida. MobileNetV2 emplea un kernel de 3×3 , la reducción del cálculo es aproximadamente

entre 8 y 9 veces menor que en la convolución estándar [22].

La arquitectura recibe imágenes de entrada de $224 \times 224 \times 3$ píxeles y se compone de una capa convolucional inicial con 32 filtros, seguida por 19 bloques residuales invertidos (bottlenecks) y una convolución pointwise final que genera mapas de características de tamaño $7 \times 7 \times 1280$. La Figura 1 muestra la arquitectura general de MobileNetV2 y la organización de sus bloques [23].

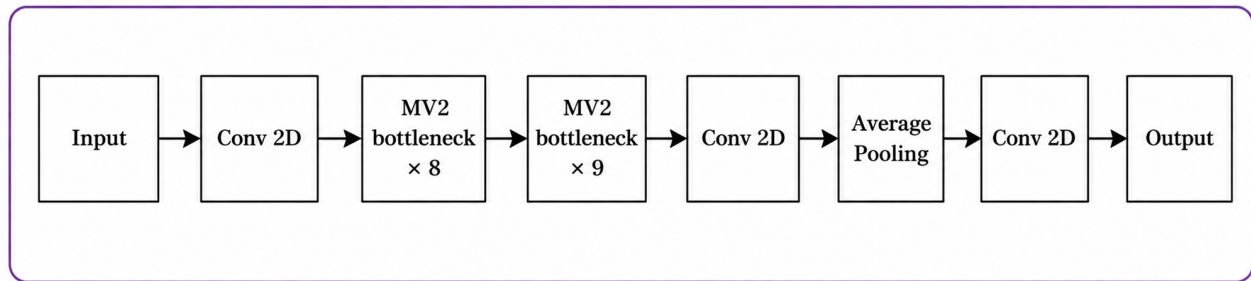


FIGURA 1. *Arquitectura general de MobileNetV2 [24].*

METODOLOGÍA

Se propone la siguiente metodología, basada en CRISP-DM (Cross Industry Standard Process for Data Mining), la cual integra distintas fases para el desarrollo sistemático de modelos de aprendizaje automático [25], orientados a la detección de cáncer de pulmón en imágenes CT mediante una CNN basada en MobileNetV2. El esquema metodológico se presenta en la Figura 2.

Recopilación de los datos

En el estudio se utilizaron dos datasets. El primero corresponde al Hospital de Enseñanza de Oncología de Iraq / Centro Nacional de Enfermedades del Cáncer (IQ-OTH/NCCD), compuesto por 1097 imágenes de CT correspondientes a 110 casos distribuidos en tres clases: normal (55 casos), benigno (15 casos) y maligno (40 casos) [26]. El análisis de la distribución de imágenes evidenció un desbalance entre las clases, particularmente en la categoría benigna, que contaba con 120 imágenes frente a 416 de la clase normal y 561 de la clase maligna. Para reducir este desbalance se incorporó el conjunto de datos SPIE-AAPM-NCI Lung Nodule Classification Challenge (LUNGx) [27].

La incorporación de LUNGx permitió incrementar la cantidad de imágenes benignas de 120 a 371. La inclusión de estudios provenientes de pacientes distintos aportó una mayor diversidad clínica y una mayor

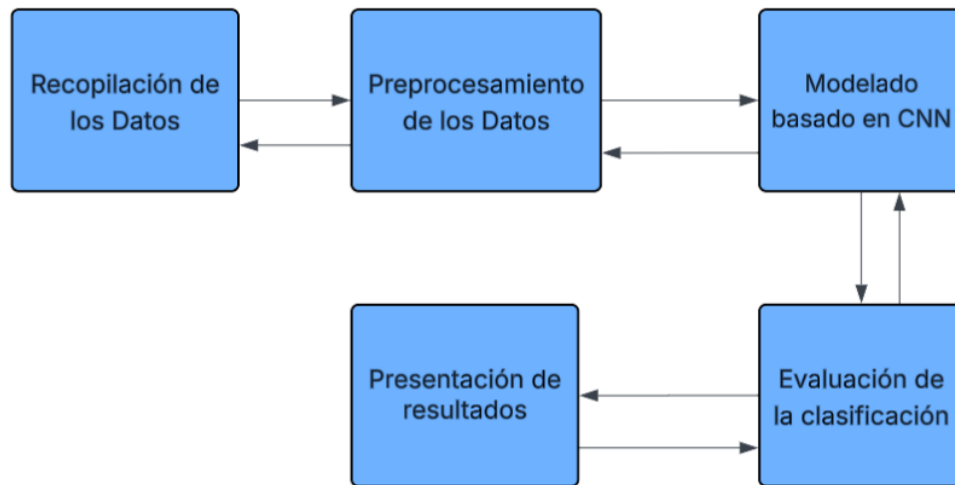


FIGURA 2. Estrategia metodológica utilizada para la detección de cáncer de pulmón.

variabilidad en las condiciones de adquisición, favoreciendo la capacidad de generalización del modelo (ver Figura 3). Ambos conjuntos de datos son de acceso público y están disponibles en las referencias [26, 27].

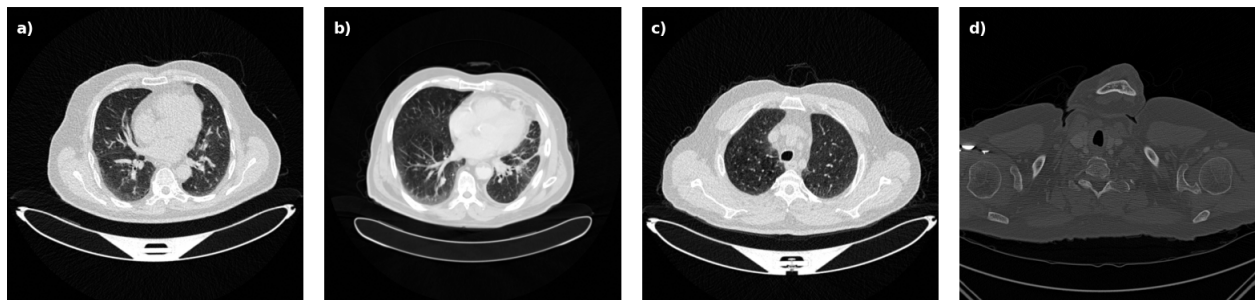


Figura X. Ejemplos de imágenes pulmonares utilizadas en el estudio: a) caso benigno, b) caso maligno, c) caso normal y d) caso benigno correspondiente al nuevo conjunto de datos.

FIGURA 3. Ejemplos de imágenes pulmonares utilizadas en el estudio. Las subfiguras a), b) y c) corresponden al dataset IQ-OTH/NCCD, que incluye casos normal, benigno y maligno, respectivamente. La subfigura d) pertenece al dataset LUNGx, correspondiente a casos benignos.

Preprocesamiento de los datos

Se extraen características visuales para la identificación de patrones asociados al cáncer de pulmón mediante técnicas de PID, como se muestra en la Figura 4.

En esta etapa se realizó el PID para mejorar la calidad de los datasets y resaltar características relevantes para la clasificación automática. El preprocesamiento se implementó en Python utilizando las librerías OpenCV, PyWavelets, Scikit-image y Pydicom. El flujo inicia con imágenes en escala

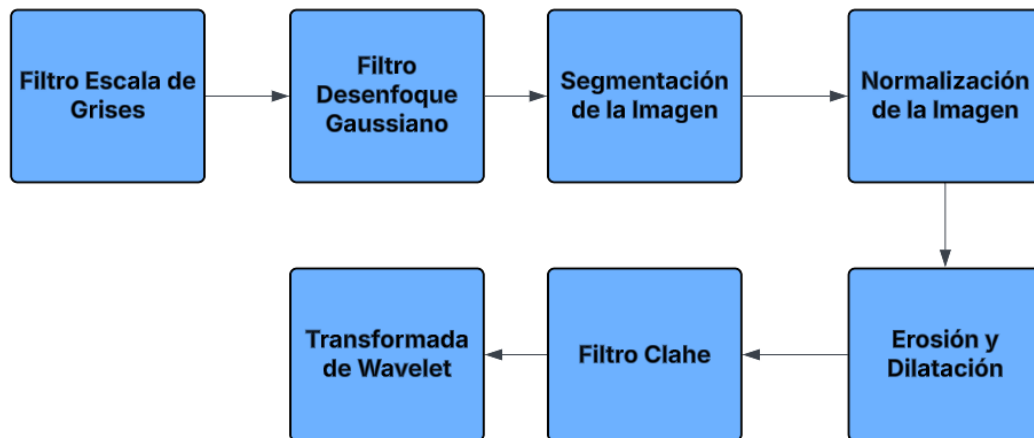


FIGURA 4. *Preprocesamiento de los datasets utilizando PID.*

de grises. Debido a que algunas tomografías pueden encontrarse almacenadas en formato RGB, se realizó una conversión para trabajar con un único canal de información y reducir la complejidad computacional. Posteriormente, se aplicó un desenfoque Gaussiano para reducir ruido, seguido de segmentación mediante el método de Otsu, normalización del dataset y operaciones morfológicas para corregir discontinuidades en la imagen. Después, se utilizó CLAHE para mejorar el contraste local. Finalmente, se aplicó la Transformada Discreta Wavelet bidimensional (2D-DWT) utilizando la función wavelet biortogonal bior1.3 y se seleccionó la sub-banda LL, correspondiente a la aproximación de baja frecuencia, como entrada para el entrenamiento de la red neuronal encargada de clasificar las imágenes en hallazgos benignos, malignos y pulmones sin evidencia de cáncer.

Modelado

Se empleó la arquitectura MobileNetV2 mediante transferencia de aprendizaje para la clasificación de imágenes CT en tres categorías: hallazgos benignos, malignos y pulmones sin evidencia de cáncer. Sobre la arquitectura base se incorporaron una capa GlobalAveragePooling2D, una capa densa de 128 neuronas con activación ReLU, una capa Dropout y una capa final Dense, descritas a continuación.

1. Capa GlobalAveragePooling2D, utilizada para reducir la dimensionalidad de los mapas de características y disminuir el riesgo de sobreajuste [28].
2. Capa densa de 128 neuronas con activación ReLU, encargada de aprender representaciones complejas de los datos con una alta eficiencia computacional [29].
3. Capa Dropout con tasa de 0.3, empleada como técnica de regularización para reducir el sobreajuste y mejorar la capacidad de generalización del modelo [30].

4. Capa final Dense con activación softmax, encargada de la clasificación de imágenes [29].

Si bien MobileNetV2 permite diferentes configuraciones, en este trabajo se empleó una configuración previamente reportada en la literatura, debido a que el objetivo principal fue validar la aplicabilidad de esta arquitectura para la clasificación de cáncer de pulmón en entornos con infraestructura limitada.

La Figura 5a muestra la evolución de la precisión durante el entrenamiento de la CNN, observándose convergencia en aproximadamente 15 épocas y una precisión cercana al 90 %, mientras que la precisión de validación se mantiene por encima del 70 %. La Figura 5b presenta una disminución progresiva de la pérdida y un comportamiento estable durante el entrenamiento, evidenciando una adecuada capacidad de generalización del modelo.

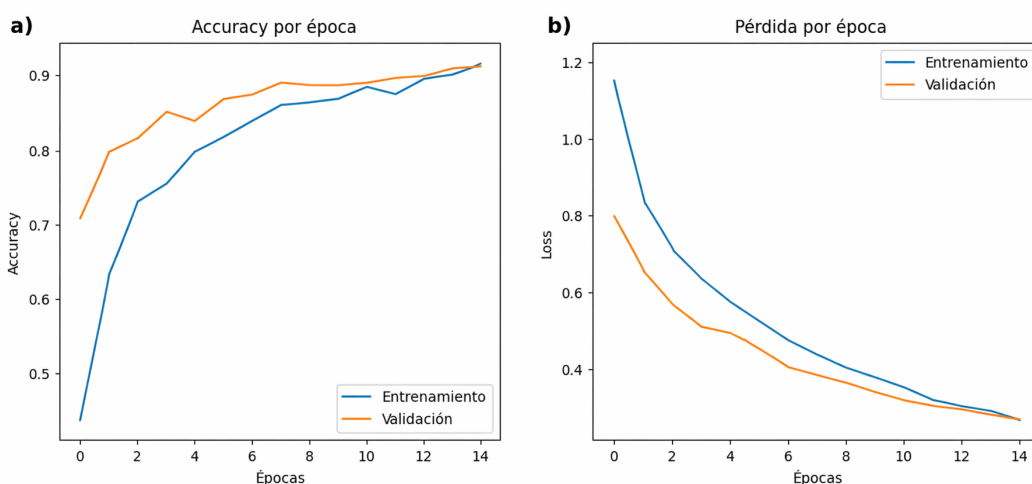


FIGURA 5. Curvas de entrenamiento y validación del modelo: a) exactitud (accuracy) y b) pérdida (loss).

Evaluación

Para validar el modelo se utilizan métricas estándar empleadas en problemas de clasificación multiclase: exactitud, precisión, sensibilidad y *F1-score* a partir de la matriz de confusión [31].

La exactitud representa la proporción de predicciones correctas respecto al total de instancias. La precisión mide la exactitud de las predicciones positivas, la sensibilidad evalúa la capacidad del modelo para identificar correctamente los casos positivos y *F1-score* corresponde a la media armónica entre precisión y la sensibilidad proporcionando una medida equilibrada del desempeño del modelo.

EXPERIMENTACIÓN Y RESULTADOS

Se implementó una CNN basada en MobileNetV2 con el objetivo de validar su utilización en sistemas computacionales con recursos limitados para la detección del cáncer de pulmón. En las Tablas 2 y 3

se presentan la matriz de confusión y las métricas de clasificación obtenidas. La clase benigno alcanzó una precisión de 0.98, sensibilidad de 0.86 y F1-score de 0.92, mostrando un equilibrio adecuado entre precisión y sensibilidad. Por su parte, la clase maligno obtuvo una precisión de 0.93, sensibilidad de 0.94 y F1-score de 0.94, evidenciando una adecuada capacidad para identificar casos malignos. Finalmente, la clase normal alcanzó una precisión de 0.80, sensibilidad de 0.94 y F1-score de 0.86, lo que indica una adecuada capacidad para identificar imágenes sin evidencia de cáncer.

Tabla 2. *Clasificación de la red MobileNetV2 sobre el conjunto de prueba.*

	Benigno	Maligno	Normal
Benigno	119	3	16
Maligno	2	97	4
Normal	1	4	79

Tabla 3. *Métricas obtenidas para el dataset durante la clasificación.*

	Precision	Recall	F1-score
Benigno	0.98	0.86	0.92
Maligno	0.93	0.94	0.94
Normal	0.80	0.94	0.86

El modelo propuesto se comparó con el trabajo de Mohamed et al. [7], quienes utilizaron el mismo conjunto de datos IQ-OTH/NCCD. Como se observa en la Tabla 4, el modelo EOSA-CNN reportó una exactitud global de 93.21 %, mientras que la propuesta basada en MobileNetV2 alcanzó una exactitud de 91.0 %. Sin embargo, el análisis por clase muestra diferencias importantes entre ambos enfoques. En la categoría benigno, MobileNetV2 obtuvo un F1-score de 0.92 frente a 0.4151 reportado por EOSA-CNN. De manera similar, para las clases maligno y normal se obtuvieron F1-scores de 0.94 y 0.86, respectivamente, superando los valores reportados por Mohamed et al. [7].

Estos resultados muestran que, aunque la exactitud global es ligeramente inferior, MobileNetV2 presenta un comportamiento más equilibrado entre las tres clases, especialmente en la categoría benigna, que corresponde a la clase con menor número de muestras. Asimismo, el uso de transferencia de aprendizaje y aumento de datos contribuye a mejorar la capacidad de generalización del modelo y reducir el sesgo hacia las clases mayoritarias.

CONCLUSIONES

En este trabajo se presentó una metodología para la clasificación de cáncer de pulmón a partir de CT mediante técnicas de PID y aprendizaje profundo. La propuesta combinó etapas de preprocesamiento,

Tabla 4. Comparación de resultados entre el modelo propuesto y el trabajo de Mohamed et al. [7].

	Arquitectura	Exactitud (%)	F1 Benigno	F1 Maligno	F1 Normal
Mohamed et al. [7]	EOSA-CNN	93.21	0.4151	0.8947	0.8384
Propuesta	MobileNetV2	91.00	0.92	0.94	0.86

entre las que destacan CLAHE y la Transformada Discreta Wavelet, con la arquitectura MobileNetV2 para clasificar las imágenes en pulmones sin evidencia de cáncer, hallazgos benignos y malignos.

Los resultados obtenidos muestran que la propuesta alcanzó una exactitud del 91.0 %, evidenciando que MobileNetV2 constituye una alternativa competitiva frente a arquitecturas más complejas reportadas en la literatura para la clasificación de cáncer de pulmón. Asimismo, la comparación realizada con otros estudios permitió identificar que, aunque algunas arquitecturas profundas alcanzan valores de exactitud superiores, MobileNetV2 representa una alternativa atractiva debido a su menor complejidad computacional y su potencial aplicación en entornos con recursos tecnológicos limitados.

Como trabajo futuro se plantea la integración del modelo en sistemas de diagnóstico asistido por computadora, implementar la herramienta propuesta en el cómputo móvil, la ampliación del modelo para la clasificación de patologías pulmonares más específicas, la evaluación de su impacto clínico en escenarios reales y la validación de los resultados por especialistas en oncología y radiología pulmonar.

DECLARACIÓN DE CONTRIBUCIÓN DE AUTORES Y COLABORADORES

Osmar D. Arellano: investigación, conceptualización, metodología, software, experimentación y redacción. **Marco A. Ramos:** investigación, conceptualización, metodología, revisión, validación y redacción. **Vianney Muñoz:** investigación, conceptualización, metodología, revisión, validación y redacción.

REFERENCIAS

- [1] A. L. Cajape-González, J. L. Ramírez-Vélez, and E. T. Fuentes-Sánchez, “Prevalencia y factores de riesgos del cáncer de pulmón: una enfermedad letal y silenciosa,” *MQRInvestigar*, vol. 7, no. 1, pp. 2904–2928, 2023.
- [2] S. B. T. Puente, D. G. Vera, Y. La O Jimenez, V. C. Torres, and M. L. Santana, “Necesidad de cuidados continuos en personas con cáncer de pulmón avanzado,” *Revista Cubana de Enfermería*, vol. 39, 2023.
- [3] S. M. Alzahrani, H. A. Al Doghaither, and A. B. Al-Ghafari, “General insight into cancer: an

overview of colorectal cancer (review),” *Molecular and Clinical Oncology*, vol. 15, no. 6, p. 271, Dec. 2021.

- [4] A. M. N. Dolz, D. Á. Matos, L. C. Toledo, M. V. M. Peña, and A. G. Nazario, “Algunas especificidades en torno al cáncer de pulmón,” *Revista Cubana de Medicina Militar*, vol. 50, no. 1, p. e0210725, 2021.
- [5] P. Rajaram, B. Balkhande, A. Bhattad, N. B. Pokale, T. S. Bhosale, and S. Patil, “Implementation of convolutional neural networks for lung cancer detection from CT scans,” *Frontiers in Health Informatics*, vol. 13, no. 2, pp. 230–246, 2024.
- [6] O. Ukwandu, H. Hindy, and E. Ukwandu, “An evaluation of lightweight deep learning techniques in medical imaging for high precision COVID-19 diagnostics,” *Healthcare Analytics*, vol. 2, p. 100096, 2022.
- [7] M. T. I. A. Mohamed, O. N. Oyelade, and A. E. Ezugwu, “Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm,” *PLOS ONE*, vol. 18, no. 8, p. e0285796, 2023.
- [8] S. Raut, S. Patil, and G. Shelke, “Lung cancer detection using machine learning approach,” *International Journal of Advance Scientific Research and Engineering Trends (IJASRET)*, vol. 6, no. 1, pp. 47–51, Jan. 2021.
- [9] M. A. Hamim, F. M. Tanmoy, U. Asha, M. Haq, M. Alam, and B. Ghosh, “Lung cancer detection using machine learning approach,” in *Proc. IEEE Int. Conf. Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2023, pp. 2676–2682.
- [10] L. Tiwari, V. Awasthi, R. Patra, R. Miri, H. Raja, and N. Bhaskar, “Lung cancer detection using deep convolutional neural networks,” in *Advances in Intelligent Systems and Computing*, 2022, pp. 373–385.
- [11] M. Ribeiro and P. Sousa, “Classification of lung cancer in medical imaging using convolutional networks with CLAHE preprocessing and wavelet transform,” in *Anais do XV Workshop de Sistemas de Informação (WSIS)*, Rio Paranaíba, MG, Brazil, 2024, pp. 45–50.
- [12] S. Wankhade and V. S. Vigneshwari, “A novel hybrid deep learning method for early detection of lung cancer using neural networks,” *Healthcare Analytics*, vol. 3, p. 100195, 2023.

- [13] J. P. Bocanegra-Valencia, L. V. Urrelo-Huiman, and S. E. Cieza-Mostacero, “Deep learning para la detección de cáncer de pulmón: una revisión sistemática de la literatura,” *Revista Ibérica de Sistemas e Tecnologías de Informação (RISTI)*, no. E75, pp. 160–174, Nov. 2024.
- [14] S. C. Casadiego, K. Y. S. Mojica, K. Puerto, C. Niño, B. M. Delgado, and D. Guevara-Ibarra, “Efecto de los filtros morfológicos en los procesos de detección de objetos en movimiento,” *Mundo FESC*, vol. 11, no. 21, pp. 87–95, 2021.
- [15] M. E. L. López, R. O. Flores, A. A. G. Cuellar, K. G. Hidalgo, and E. J. H. Nava, “De lo bidimensional a lo tridimensional: filtros para la caracterización morfológica y reconstrucción 3D,” *Revista CiBlyT*, vol. 19, no. 55, p. 37, 2024.
- [16] J. Zheng, Y. Gao, H. Zhang, Y. Lei, and J. Zhang, “OTSU multi-threshold image segmentation based on improved particle swarm algorithm,” *Applied Sciences*, vol. 12, no. 22, p. 11514, 2022.
- [17] C. Huang, X. Li, and Y. Wen, “An otsu image segmentation based on fruit fly optimization algorithm,” *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 183–188, 2021.
- [18] B. G. Chepino, R. R. Yacoub, A. Aula, M. Saleh, and B. W. Sanjaya, “Effect of minmax normalization on ORB data for improved ANN accuracy,” *Journal of Electrical Engineering, Energy, and Information Technology (J3EIT)*, vol. 11, no. 2, p. 29, 2023.
- [19] G. Ferreira, L. Ishiuchi, L. Teixeira, E. Oliveira, and W. D. Parreira, “Aprimoramento da classificação de linfócitos e monócitos em imagens médicas: o impacto do CLAHE em redes neurais convolucionais,” in *Anais do Computer on the Beach*, vol. 16, 2025, pp. 199–212.
- [20] M. V. D. dos Reis and P. M. de Sousa, “Quality assessment of coffee beans using convolutional neural networks with wavelet and CLAHE techniques,” in *Anais do XV Workshop de Sistemas de Informação (WSIS)*, Rio Paranaíba, MG, Brazil, 2024, pp. 33–38.
- [21] K. Bhatt, T. Kumar, S. Hassan, and S. Kimothi, “2D discrete wavelet transformation (2D-DWT) for nanoscale morphological analysis,” *Prabha Materials Science Letters*, vol. 2, no. 2, pp. 140–153, 2023.
- [22] K. Ding, Z. Niu, J. Hui, X. Zhou, and F. T. S. Chan, “A weld surface defect recognition method based on improved MobileNetV2 algorithm,” *Mathematics*, vol. 10, no. 19, p. 3678, 2022.
- [23] R. Indraswari, R. Rokhana, and W. Herulambang, “Melanoma image classification based on MobileNetV2 network,” *Procedia Computer Science*, vol. 197, pp. 198–207, 2022.

- [24] X. Deng, M. Shi, B. Khan, Y. H. Choo, F. Ghaffar, and C. P. Lim, “A lightweight CNN model for UAV-based image classification,” *Soft Computing*, vol. 29, no. 4, pp. 2363–2378, 2025.
- [25] C. E. D. Vanegas, J. C. G. Mejía, F. A. V. Agudelo, and D. E. S. Durán, “A representation based on essence for the CRISP-DM methodology,” *Computación y Sistemas*, vol. 27, no. 3, pp. 675–689, 2023.
- [26] H. Alyasriy and M. AL-Huseiny, “The IQ-OTHNCCD lung cancer dataset,” Mendeley Data, 2021.
- [27] S. G. Armato III, L. Hadjiiski, G. D. Tourassi, K. Drukker, M. L. Giger, F. Li, G. Redmond, K. Farahani, J. S. Kirby, and L. P. Clarke, “SPIE-AAPM-NCI lung nodule classification challenge dataset,” The Cancer Imaging Archive, 2015.
- [28] P. P. Malla, S. Sahu, and A. I. Alutaibi, “Classification of tumor in brain MR images using deep convolutional neural network and global average pooling,” *Processes*, vol. 11, no. 3, p. 679, 2023.
- [29] A. V. Morozov, V. L. Levkivskyi, and D. D. Plechystyy, “Activation functions in neural networks: overview and comparison,” *Scientific Notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, vol. 35, no. 3, pp. 144–151, 2024.
- [30] I. Salehin and D.-K. Kang, “A review on dropout regularization approaches for deep neural networks within the scholarly domain,” *Electronics*, vol. 12, no. 14, p. 3106, 2023.
- [31] R. O. Ogundokun, S. Misra, A. O. Akinrotimi, and H. Ogul, “MobileNet-SVM: a lightweight deep transfer learning model to diagnose BCH scans for IoMT-based imaging sensors,” *Sensors*, vol. 23, no. 2, p. 656, 2023.