

Clasificación de radiografías de tórax para la detección de neumonía mediante MobileNetV2

Sonia Guadalupe López López, Luis Fernando Aguilar Cruz

Departamento de Ingeniería Biomédica, Universidad Modelo

Mérida, Yucatán, México

15257650@modelo.edu.mx

Resumen- La neumonía es una patología respiratoria cuya detección oportuna puede favorecer el manejo clínico y disminuir el riesgo de complicaciones. La radiografía de tórax es una modalidad frecuente para su valoración inicial; sin embargo, la interpretación visual depende de la calidad de la imagen, la presencia de hallazgos sutiles y la experiencia del observador. En este trabajo se desarrolló y evaluó un clasificador binario de radiografías de tórax basado en MobileNetV2 con transferencia de aprendizaje. El proyecto fue elaborado como una aplicación académica de procesamiento de imágenes médicas por estudiantes de Ingeniería Biomédica de la Universidad Modelo. Las imágenes fueron redimensionadas a 224 x 224 píxeles, convertidas a tres canales y normalizadas. El entrenamiento se realizó durante 10 épocas con batch size de 32. La evaluación consideró exactitud, precisión, sensibilidad, especificidad, matriz de confusión, curva ROC y AUC. En el conjunto de evaluación se obtuvo una exactitud de 97.27%, precisión de 98.97%, sensibilidad de 95.52%, especificidad de 99.01% y AUC de 0.9965. Los resultados muestran un desempeño alto para distinguir radiografías normales y con neumonía, aunque el sistema debe interpretarse como herramienta de apoyo académico y no como diagnóstico clínico autónomo.

Abstract- Pneumonia is a respiratory condition whose timely detection may support clinical management and reduce the risk of complications. Chest radiography is commonly used for initial assessment; nevertheless, visual interpretation depends on image quality, subtle radiographic findings, and observer experience. This work developed and evaluated a binary chest X-ray classifier based on MobileNetV2 with transfer learning. The project was conducted as an academic application of medical image processing by Biomedical Engineering students from Universidad Modelo. Images were resized to 224 x 224 pixels, converted to three channels, and normalized. Training was performed for 10 epochs with a batch size of 32. Evaluation included accuracy, precision, sensitivity, specificity, confusion matrix, ROC curve, and AUC. The evaluation subset achieved 97.27% accuracy, 98.97% precision, 95.52% sensitivity, 99.01% specificity, and an AUC of 0.9965. These results indicate high performance for distinguishing normal and pneumonia chest X-rays; however, the system should be considered an academic support tool rather than an autonomous clinical diagnosis method.

Palabras clave- radiografía de tórax, neumonía, MobileNetV2, aprendizaje profundo, procesamiento de imágenes médicas, curva ROC, AUC.

I. INTRODUCCIÓN

La neumonía es una infección respiratoria que puede ser causada por bacterias, virus u hongos y su gravedad depende de factores como la edad del paciente, su estado general de salud y el agente causal de la infección. debido a su frecuencia y a las complicaciones que puede generar cuando no se identifica tempranamente, llegando a ser la cuarta causa de muerte a nivel nacional en México y una de las cinco

principales causas de mortalidad general, provocando más de 36,000 defunciones anuales.

Para su diagnóstico, se realiza una exploración física y solicita estudios de apoyo, entre los que destaca la radiografía de tórax. En imagenología, este estudio permite observar alteraciones como consolidaciones, infiltrados u opacidades pulmonares asociadas a procesos infecciosos. Sin embargo, a pesar de ser una herramienta accesible y ampliamente utilizada, la interpretación de las radiografías puede variar por factores técnicos, anatómicos y humanos, como la calidad de la imagen, la presencia de hallazgos sutiles y la experiencia del observador [1], [2].

En este contexto, el procesamiento de imágenes médicas y el aprendizaje profundo han abierto nuevas posibilidades para analizar información visual de manera más reproducible. En particular, las redes neuronales convolucionales han mostrado utilidad en tareas de clasificación radiográfica, ya que pueden aprender patrones directamente desde las imágenes sin depender exclusivamente de características definidas manualmente [3]-[6]. No obstante, uno de los principales desafíos ha sido identificar modelos que, además de ofrecer buen desempeño, sean eficientes en términos computacionales y viables para entornos académicos o de recursos limitados.

Entre las arquitecturas disponibles, MobileNetV2 destaca por su eficiencia en visión por computadora. Su diseño basado en convoluciones separables en profundidad, residuos invertidos y cuellos de botella lineales permite reducir el número de parámetros y el costo computacional sin perder capacidad de extracción de características relevantes [10]. Bajo este enfoque, el presente trabajo implementa MobileNetV2 para clasificar radiografías de tórax en dos categorías: normal y neumonía.

II. PLANTEAMIENTO DEL PROBLEMA Y OBJETIVOS

La interpretación de radiografías de tórax puede depender de la experiencia del lector, la calidad del estudio y la carga de trabajo. En escenarios con múltiples estudios por revisar, un sistema automatizado de clasificación puede servir como apoyo para organizar imágenes y estimar la probabilidad de hallazgos compatibles con neumonía. El problema consiste en desarrollar y evaluar un clasificador binario basado en MobileNetV2, considerando su desempeño mediante métricas objetivas y su alcance.

A. Objetivo general

Desarrollar y evaluar un sistema de clasificación de radiografías de tórax para la detección de neumonía mediante MobileNetV2 con transferencia de aprendizaje.

B. Objetivos específicos

- Seleccionar un conjunto de datos público de radiografías de tórax con las clases normal y neumonía.

- Estandarizar las imágenes mediante conversión a tres canales y normalización.
- Implementar MobileNetV2 como base de extracción de características para clasificación binaria.
- Entrenar el modelo durante 10 épocas con batch size de 32.
- Evaluar el desempeño mediante exactitud, precisión, sensibilidad, especificidad, matriz de confusión, curva ROC y AUC.
- Analizar los resultados, ventajas, limitaciones y posibles mejoras del sistema.

III. REVISIÓN DE LA LITERATURA

El uso de aprendizaje profundo en radiografías de tórax ha mostrado resultados prometedores en la detección de neumonía. Un punto de partida importante es el trabajo de Kermany et al. [1], quienes demostraron que los modelos de aprendizaje profundo pueden identificar diagnósticos tratables a partir de imágenes médicas, incluyendo radiografías de tórax. A partir de este antecedente, distintas investigaciones comenzaron a explorar el potencial de las redes neuronales convolucionales para distinguir entre estudios normales y estudios con neumonía.

En esta línea, Jaiswal et al. [3] reportaron que las arquitecturas profundas aplicadas a radiografías de tórax son capaces de aprender patrones relevantes directamente desde la imagen, sin depender exclusivamente de características diseñadas manualmente. De manera complementaria, Chouhan et al. [4] evaluaron un enfoque basado en aprendizaje por transferencia, comparando distintas arquitecturas preentrenadas para la detección de neumonía, y encontraron que este tipo de estrategias permite obtener desempeños favorables reduciendo tanto el tiempo de entrenamiento como la necesidad de construir una red desde cero.

Posteriormente, Elshennawy e Ibrahim [5] propusieron el marco Deep-Pneumonia, en el cual analizaron diversos modelos de aprendizaje profundo para clasificar radiografías de tórax. Sus hallazgos confirmaron que las arquitecturas convolucionales preentrenadas pueden alcanzar resultados altos en tareas de clasificación binaria, reforzando así su utilidad en problemas de imagenología médica. En un sentido similar, Zhang et al. [6] también abordaron la detección de neumonía mediante redes convolucionales, mostrando que este enfoque constituye una alternativa viable para apoyar la interpretación de radiografías de tórax.

En conjunto, estos trabajos evidencian que las redes neuronales convolucionales preentrenadas representan una estrategia sólida para la clasificación de radiografías de tórax, particularmente cuando se busca un equilibrio entre desempeño y costo computacional. Bajo este panorama, MobileNetV2 resulta una opción especialmente adecuada, ya que su diseño basado en convoluciones separables en profundidad, residuos invertidos y cuellos de botella lineales reduce el número de parámetros y favorece su implementación en entornos con recursos limitados [10]-[12]. Por ello, en el presente trabajo se adopta esta arquitectura como base para distinguir entre radiografías normales y radiografías con neumonía.

IV. METODOLOGÍA

A. Conjunto de datos

Se utilizó como fuente el conjunto público Chest X-Ray Images (Pneumonia), integrado por 5856 radiografías de tórax, de las cuales 1583 corresponden a la clase normal y 4273 a la clase neumonía [15]. Debido al desbalance entre ambas clases, para el desarrollo experimental del modelo no se trabajó con la distribución completa original, sino con un subconjunto balanceado. Para ello, se seleccionaron 1341 imágenes de la clase normal y 1341 imágenes de la clase neumonía, obteniendo un total de 2682 imágenes.

Posteriormente, este subconjunto balanceado se dividió de forma estratificada en 1877 imágenes para entrenamiento, 402 imágenes para validación y 403 imágenes para prueba, con el fin de conservar una distribución equilibrada de las clases y obtener una evaluación más consistente del clasificador.

B. Preprocesamiento de imágenes

Las radiografías fueron redimensionadas a 224×224 píxeles, tamaño compatible con MobileNetV2. Debido a que muchas de las imágenes se encuentran originalmente en escala de grises, se cargaron en formato de tres canales para ajustarse al formato de entrada esperado por el modelo. Posteriormente, se aplicó la función `preprocess_input` para normalizar las intensidades de los píxeles de acuerdo con la configuración requerida por MobileNetV2 [10], [16].

C. Arquitectura del modelo

MobileNetV2 se empleó como base convolucional preentrenada para la extracción de características visuales, utilizando la configuración `include_top = False` para eliminar la capa clasificadora original [10]. La salida de la red base fue procesada mediante una capa `GlobalAveragePooling2D`, con el objetivo de resumir los mapas de características en un vector compacto. Sobre esta representación se añadió una capa densa de 128 neuronas con activación ReLU, seguida de una capa Dropout de 0.5 para reducir el sobreajuste. Finalmente, se incorporó una capa de salida de una neurona con activación sigmoide, adecuada para clasificar las radiografías en dos categorías: normal y neumonía.

D. Entrenamiento

El entrenamiento se realizó con las 1877 imágenes correspondientes al conjunto de entrenamiento, mientras que el seguimiento del desempeño del modelo se llevó a cabo con las 402 imágenes del conjunto de validación. El modelo fue entrenado durante 10 épocas con un batch size de 32. Al tratarse de una tarea de clasificación binaria, se utilizó binary cross-entropy como función de pérdida y Adam como optimizador [13], [16]. Durante el entrenamiento se registraron la pérdida y la exactitud tanto en entrenamiento como en validación, con el propósito de revisar la estabilidad del aprendizaje y detectar posibles señales de sobreajuste.

E. Evaluación

El desempeño final del modelo se evaluó utilizando las 403 imágenes del conjunto de prueba mediante exactitud, precisión, sensibilidad, especificidad, matriz de confusión, curva ROC y AUC [14]. Para la interpretación de la matriz de confusión se consideró la clase neumonía como clase positiva. En ese sentido, los falsos negativos corresponden a radiografías con neumonía clasificadas como normales, error que debe analizarse con especial cuidado en una aplicación médica.

V. RESULTADOS

Las curvas de entrenamiento muestran una reducción marcada de la pérdida durante las primeras épocas, seguida de una estabilización con valores bajos. La pérdida de validación se mantuvo cercana a la de entrenamiento, con variaciones moderadas, lo que no sugiere un sobreajuste severo. En la curva de exactitud, ambos subconjuntos se ubicaron por encima de 0.97 hacia las últimas épocas, con una exactitud de validación final aproximada de 0.977.

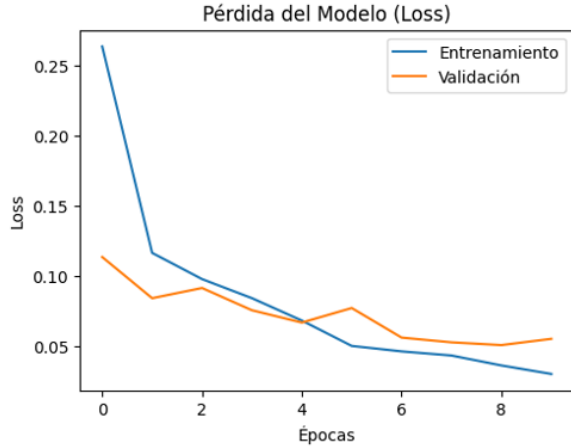


Fig. 1. Curva de pérdida durante el entrenamiento.1

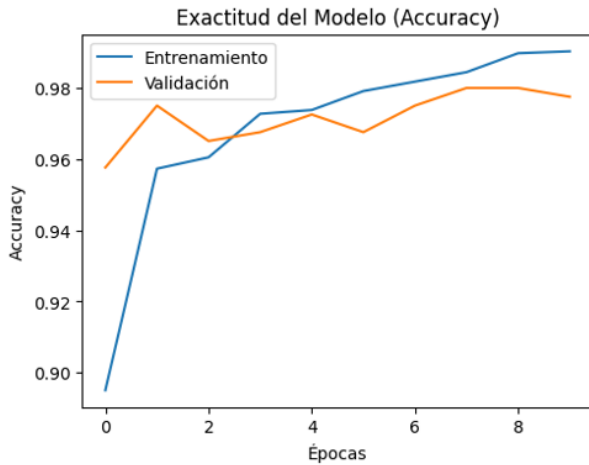


Fig. 2. Curva de exactitud durante el entrenamiento y la validación.2

La curva ROC obtuvo un AUC de 0.9965, lo que indica una alta capacidad de separación entre las clases normal y neumonía. La matriz de confusión registró 200 verdaderos negativos, 192 verdaderos positivos, 2 falsos positivos y 9 falsos negativos en un conjunto de evaluación de 403 imágenes.

VI. DISCUSIÓN

Los resultados obtenidos muestran que MobileNetV2 fue capaz de diferenciar radiografías normales y con neumonía con desempeño alto. La exactitud global de 97.27% indica una proporción elevada de aciertos, mientras que la precisión de 98.97% muestra que la mayoría de las imágenes clasificadas como neumonía correspondieron realmente a esa clase. La sensibilidad de 95.52% sugiere una detección

A partir de la matriz de confusión se calcularon 392 clasificaciones correctas de 403 imágenes evaluadas. Los 2 falsos positivos corresponden a radiografías normales clasificadas como neumonía; los 9 falsos negativos corresponden a radiografías con neumonía clasificadas como normales. Aunque el número de errores fue bajo, los falsos negativos requieren atención porque representan el tipo de error con mayor relevancia en una tarea de detección de patología.

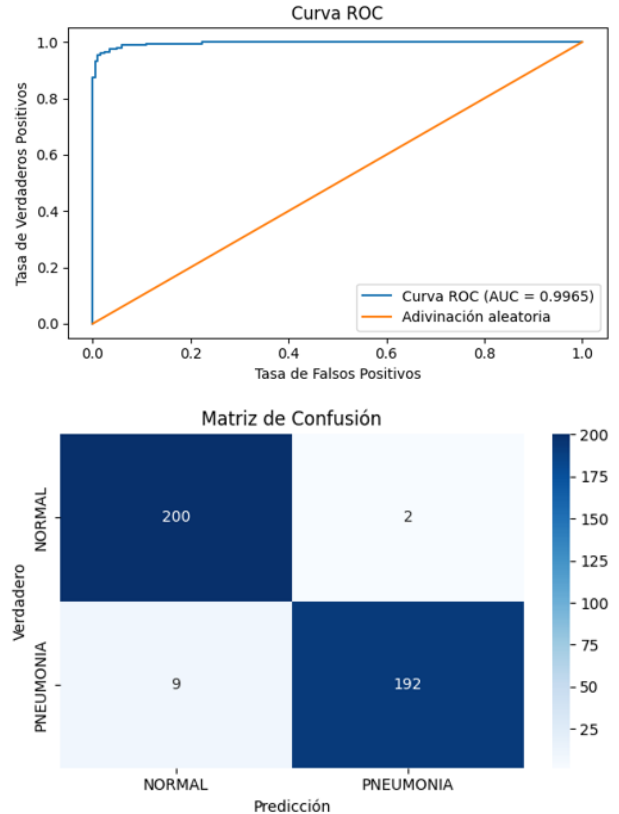


Fig. 3. Curva ROC y matriz de confusión del conjunto de prueba.

Tabla 1.

Métricas de desempeño del modelo MobileNetV2	
Métrica	Resultado
Exactitud	97.27%
Precisión	98.97%
Sensibilidad	95.52%
Especificidad	99.01%
AUC	99.65%
Clasificaciones correctas	392 / 403

adecuada de casos positivos, aunque no elimina la presencia de falsos negativos.

El AUC de 0.9965 complementa la interpretación porque resume la capacidad del modelo para separar ambas clases bajo distintos umbrales de decisión. Esta métrica resulta especialmente útil cuando el conjunto de datos presenta desbalance o cuando se busca evaluar la discriminación general del clasificador más allá de una decisión fija.

El uso de MobileNetV2 también representa una ventaja práctica. Al ser una arquitectura ligera, permite entrenar y evaluar el modelo en un entorno accesible como Google Colab. Esto favorece su uso en actividades académicas de Ingeniería Biomédica relacionadas con procesamiento de imágenes médicas, aprendizaje profundo y evaluación cuantitativa de modelos.

No obstante, el sistema tiene limitaciones. El conjunto de datos público no necesariamente reproduce todas las condiciones de adquisición, edad, variabilidad anatómica y protocolos presentes en un hospital. Además, el modelo clasifica la imagen completa y no localiza de manera explícita las regiones pulmonares responsables de la predicción. Por ello, el resultado debe entenderse como una aproximación experimental y no como una herramienta diagnóstica autónoma.

VII. CONCLUSIONES

Se desarrolló un sistema de clasificación binaria de radiografías de tórax para detección de neumonía mediante MobileNetV2 con transferencia de aprendizaje. El flujo incluyó redimensión a 224 x 224 píxeles, conversión a tres canales y normalización, entrenamiento durante 10 épocas con batch size de 32 y evaluación mediante matriz de confusión, curva ROC y AUC.

El modelo alcanzó exactitud global de 97.27%, precisión de 98.97%, sensibilidad de 95.52%, especificidad de 99.01% y AUC de 0.9965. Estos valores indican un desempeño favorable para separar radiografías normales y con neumonía dentro del conjunto evaluado. La matriz de confusión mostró pocos errores, aunque los falsos negativos continúan siendo el punto crítico para una aplicación médica.

Como trabajo futuro se propone evaluar el modelo con datos externos, aplicar validación cruzada o separación por paciente, incorporar técnicas de interpretabilidad como Grad-CAM, analizar estrategias para reducir falsos negativos y comparar MobileNetV2 con otras arquitecturas ligeras bajo las mismas condiciones experimentales.

REFERENCIAS

- [1] D. S. Kermany et al., "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell*, vol. 172, no. 5, pp. 1122-1131, 2018, doi: 10.1016/j.cell.2018.02.010.
- [2] J. P. Metlay et al., "Diagnosis and Treatment of Adults with Community-acquired Pneumonia," *American Journal of Respiratory and Critical Care Medicine*, vol. 200, no. 7, pp. e45-e67, 2019, doi: 10.1164/rccm.201908-1581ST.
- [3] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement*, vol. 145, pp. 511-518, 2019, doi: 10.1016/j.measurement.2019.05.076.
- [4] V. Chouhan et al., "A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images," *Applied Sciences*, vol. 10, no. 2, p. 559, 2020, doi: 10.3390/app10020559.
- [5] N. M. Elshennawy and D. M. Ibrahim, "Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images," *Diagnostics*, vol. 10, no. 9, p. 649, 2020, doi: 10.3390/diagnostics10090649.
- [6] D. Zhang, F. Ren, Y. Li, L. Na, and Y. Ma, "Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network," *Electronics*, vol. 10, no. 13, p. 1512, 2021, doi: 10.3390/electronics10131512.
- [7] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, "Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19," *Cognitive Computation*, 2021, doi: 10.1007/s12559-020-09787-5.
- [8] P. K. Wong et al., "Automatic Detection of Multiple Types of Pneumonia: Open Dataset and a Multi-Scale Attention Network," *Biomedical Signal Processing and Control*, vol. 73, p. 103415, 2022, doi: 10.1016/j.bspc.2021.103415.
- [9] P. Szepesi and L. Szilagyi, "Detection of Pneumonia Using Convolutional Neural Networks and Deep Learning," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 3, pp. 1012-1022, 2022, doi: 10.1016/j.bbe.2022.08.001.
- [10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 4510-4520.
- [11] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv:1704.04861*, 2017.
- [12] J. Deng et al., "ImageNet: A Large-Scale Hierarchical Image Database," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, 2009, pp. 248-255.
- [13] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv:1603.04467*, 2016.
- [14] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006, doi: 10.1016/j.patrec.2005.10.010.
- [15] P. Mooney, "Chest X-Ray Images (Pneumonia)," *Kaggle*, 2018. [Online]. Available: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>. Accessed: May 15, 2026.
- [16] F. Chollet et al., "Keras," 2015. [Online]. Available: <https://keras.io>. Accessed: May 15, 2026.