



21th ECORFAN® International Conference - Science, Technology and Innovation



Booklets

RENIECYT - LATINDEX - Research Gate - DULCINEA - CLASE - Sudoc - HISPANA - SHERPA UNIVERSIA - Google Scholar DOI - REDIB - Mendeley - DIALNET - ROAD - ORCID - V|LEX

Title: Detection of Lung and Colon Cancer with InceptionResNetV2: A Transfer Learning Approach

Authors: Ochoa-Ornelas, Raquel, Gudiño-Ochoa, Alberto, García-Rodríguez, Julio Alberto and Uribe-Toscano, Sofía

Instituto Tecnológico de Ciudad Guzmán S-4687-2018 0000-0003-1824-5789 668976
 Instituto Tecnológico de Ciudad Guzmán HDN-0235-2022 000-0002-2366-7452 1135216
 Centro Universitario del Sur LBI-3198-2024 0000-0003-0339-0545 367137
 Centro Universitario del Sur JRY-7988-2023 0009-0009-8325-8237

Editorial label ECORFAN: 607-8695

BECORFAN Control Number: 2024-01

BECORFAN Classification (2024): 121224-0001

RNA: 03-2010-032610115700-14

Pages: 36

CONAHCYT classification:

Area: Engineering

Field: Technological sciences

Discipline: Computer technology

Subdiscipline: Artificial intelligence

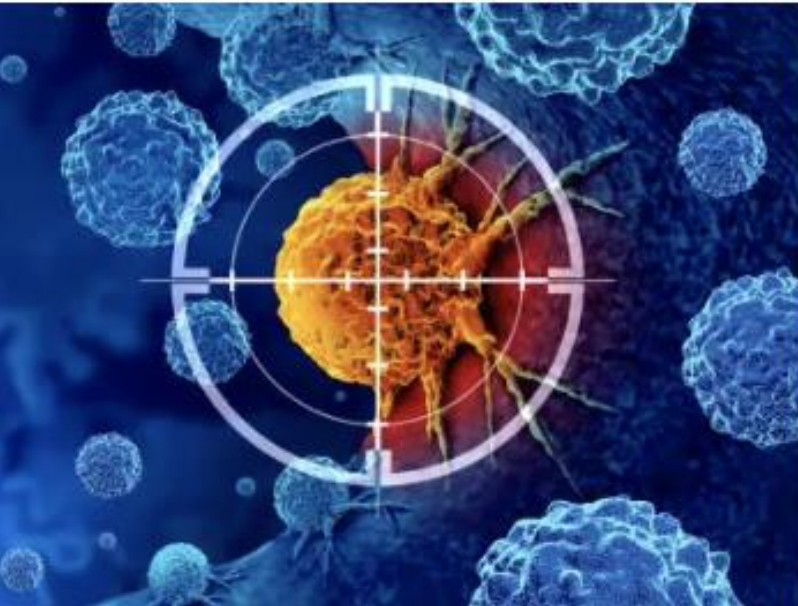
ECORFAN-México, S.C.

Park Pedregal Business. 3580,
Anillo Perif., San Jerónimo
Aculco, Álvaro Obregón,
01900 Ciudad de México, CDMX,
Phone: +52 1 55 6159 2296
Skype: MARVID-México S.C.
E-mail: contact@rinoe.org
Facebook: RINOE-México S. C.
Twitter: [@Rinoe_México](https://twitter.com/Rinoe_México)

www.marvid.org

Holdings

Mexico	Colombia	Guatemala
Bolivia	Cameroon	Democratic
Spain	El Salvador	Republic
Ecuador	Taiwan	of Congo
Peru	Paraguay	Nicaragua



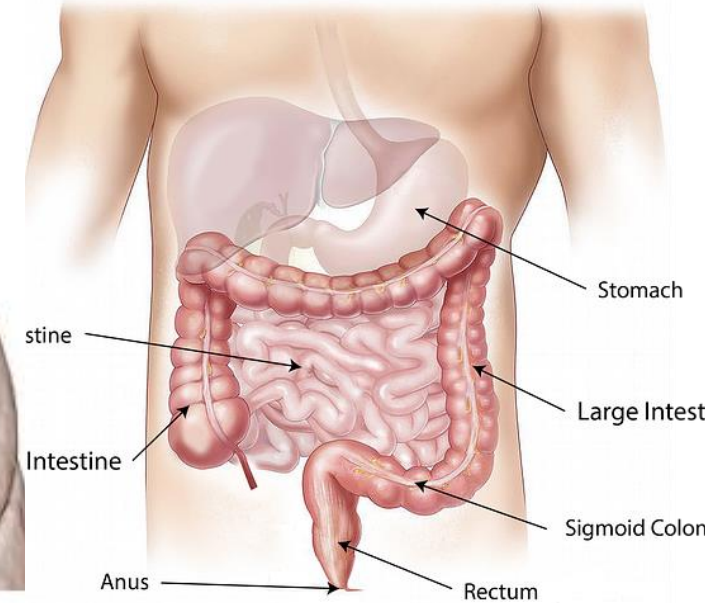
Introducción

El cáncer es la segunda causa de muerte en todo el mundo, con más de 19 millones de nuevos diagnósticos y 10 millones de muertes notificadas en 2020 (Sung et al., 2020). Los cánceres de pulmón y colon representan 4,19 millones de casos y 2,7 millones de muertes.

Importancia de la detección temprana

La detección temprana suele ser difícil sin exámenes médicos especializados, lo que resalta la importancia de los exámenes regulares para las personas con mayor riesgo (Crosby et al., 2023).

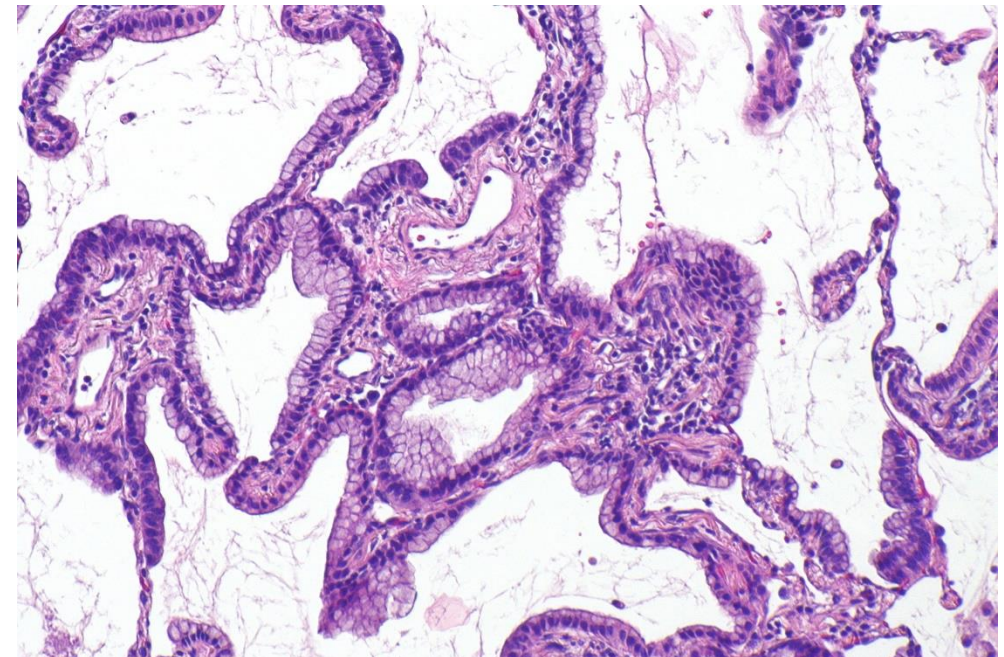
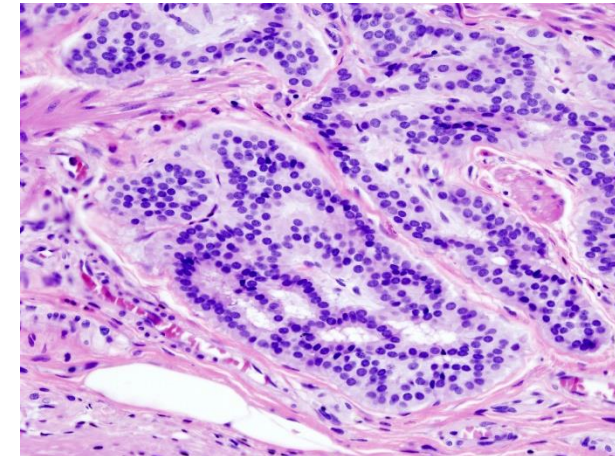
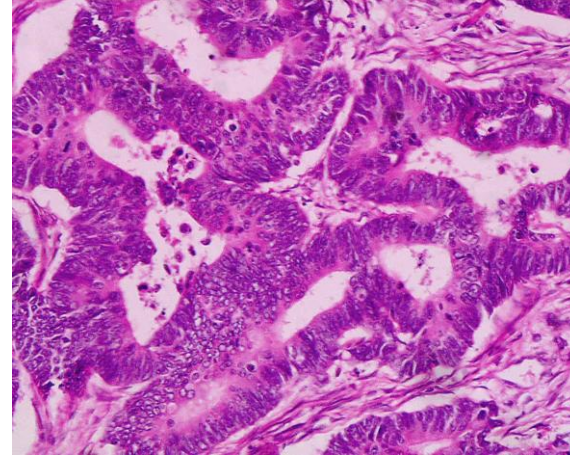
Los métodos de diagnóstico suelen incluir técnicas de imagen como radiografías, tomografías computarizadas y biopsias.



Fuente: Crosby, D., Bhatia, S., Brindle, K. M., Coussens, L. M., Dive, C., Emberton, M., ... & Balasubramanian, S. (2022). [Early detection of cancer](#). *Science*, 375(6586), eaay9040.

Rol del análisis histopatológico en el diagnóstico del cáncer

El análisis histopatológico sigue siendo el **estándar de oro** para confirmar los diagnósticos de cáncer. Sin embargo, la inspección histopatológica requiere mucho trabajo y depende de la experiencia subjetiva de los patólogos, lo que genera posibles inconsistencias en el diagnóstico (Panayides et al., 2023).



Factores de riesgo y desafíos en el diagnóstico

1

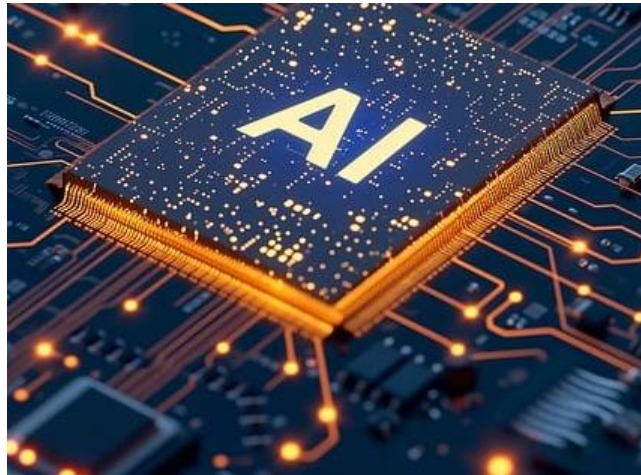
Puntos fuertes de la detección temprana

Los avances en las tecnologías de diagnóstico por imágenes y los protocolos de detección aumentan el potencial de diagnóstico temprano de cánceres de pulmón y colon, mejorando significativamente los resultados de los pacientes

2

Oportunidades para la integración de la IA

La aplicación de la inteligencia artificial y el aprendizaje automático en los procesos de diagnóstico ofrece vías prometedoras para aumentar la precisión y la eficiencia en la detección del cáncer.



3

Debilidades en el reconocimiento de síntomas

Los síntomas no específicos y las presentaciones asintomáticas a menudo conducen a un diagnóstico tardío, lo que complica la intervención oportuna y las estrategias de tratamiento para ambos tipos de cáncer.

4

Amenazas de factores ambientales

La exposición constante a carcinógenos y factores de riesgo relacionados con el estilo de vida sigue planteando desafíos importantes para reducir la incidencia de cánceres de pulmón y de colon.

Importancia de la detección temprana con IA

Pros

- Tasas de supervivencia mejoradas
- Opciones de tratamiento temprano
- Mejores resultados para los pacientes
- Reducción de la complejidad del tratamiento
- Mayor eficacia de la detección
- Beneficios de la intervención oportuna

Contras

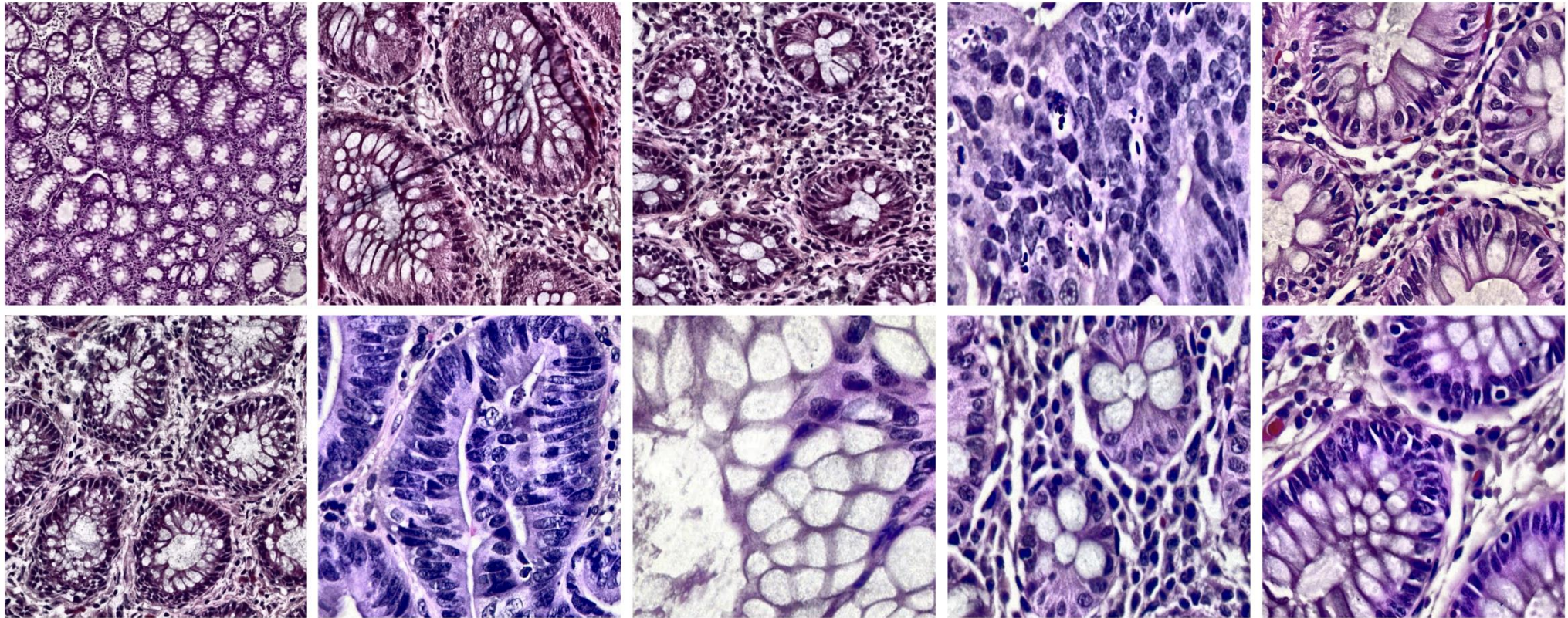
- Altos costos de detección
- Posibles falsos positivos
- Ansiedad de los pacientes por las pruebas de detección
- Acceso limitado a la tecnología
- Variabilidad en las pautas de detección
- Necesidad de educación del paciente

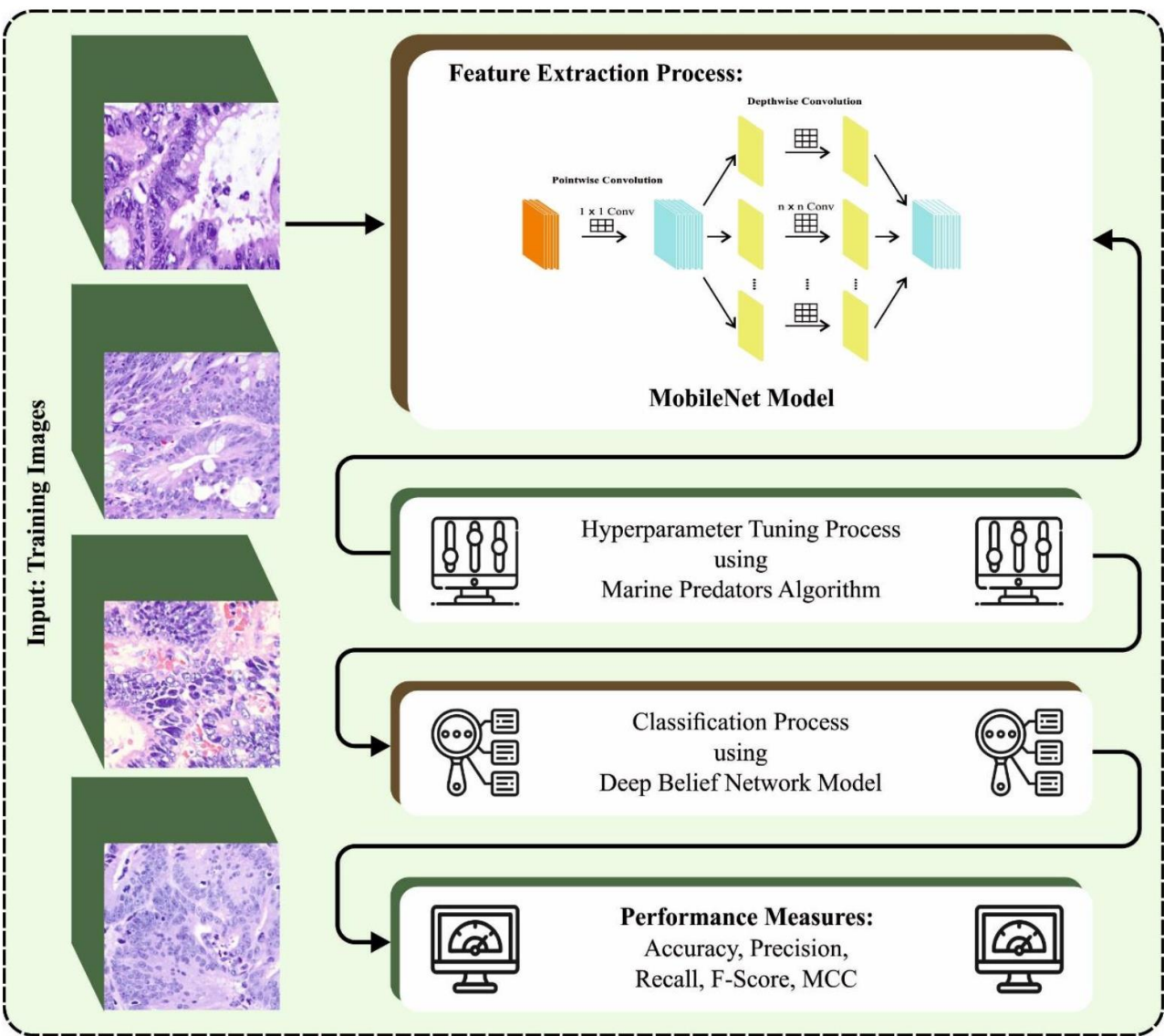


Metodología

Inception-ResNet V2

Este estudio se centra en el uso de Inception-ResNet V2 para clasificar los cánceres de pulmón y colon a partir de nuevas imágenes histopatológicas, con el objetivo de mejorar tanto la precisión diagnóstica como la interpretabilidad (XAI).

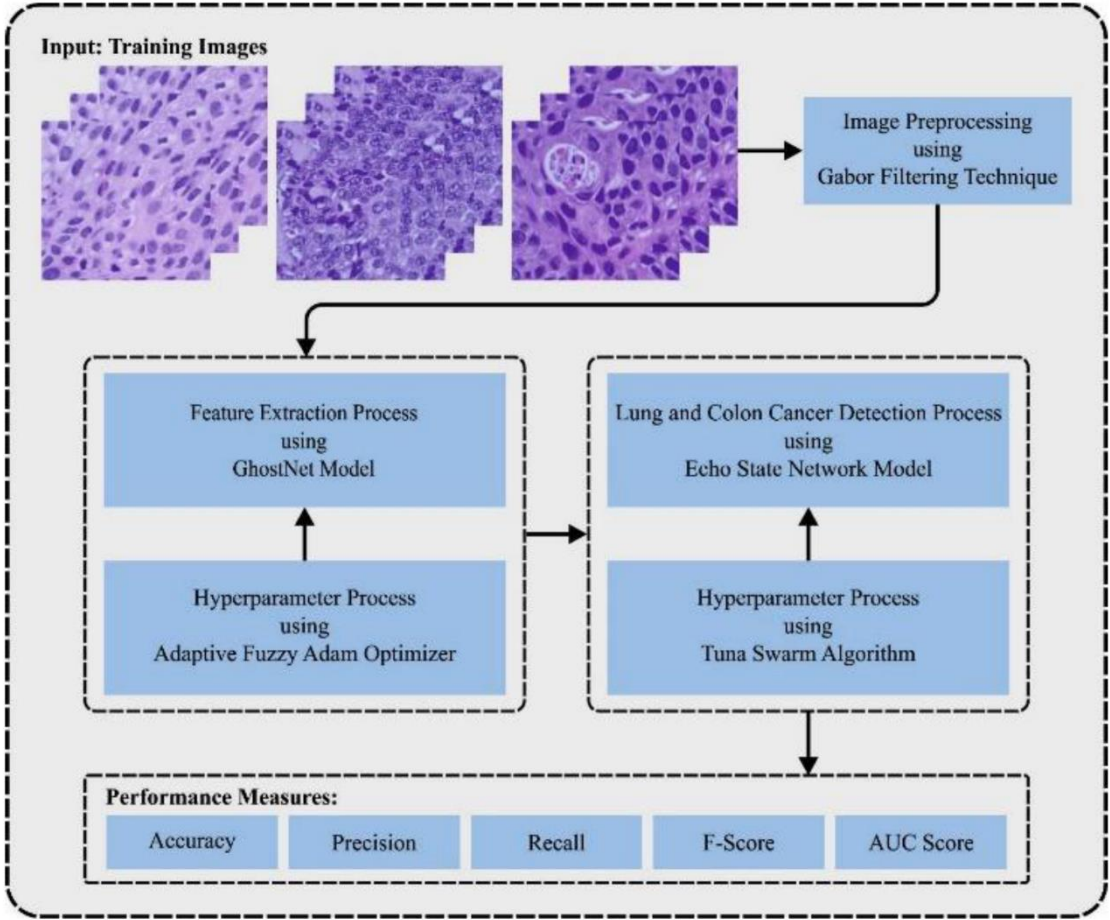




Fuente: Mengash, H. A., Alamgeer, M., Maashi, M., Othman, M., Hamza, M. A., Ibrahim, S. S., ... & Yaseen, I. (2023). Leveraging marine predators algorithm with deep learning for lung and colon cancer diagnosis. *Cancers*, 15(5), 1591.

Tabla 1. Comparación de métodos existentes

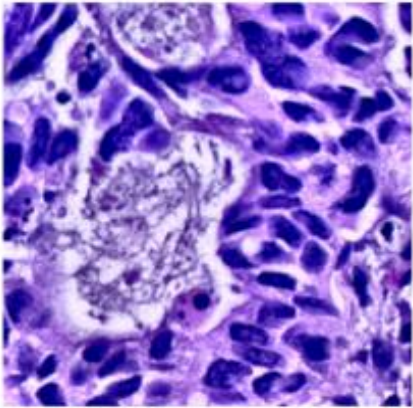
Methods	<i>Accu_y</i>	<i>Prec_n</i>	<i>Reca_l</i>	<i>F_{Score}</i>
BICLCD-TSADL	99.33	98.31	98.31	98.31
MPADL-LC3	98.81	97.92	97.87	97.91



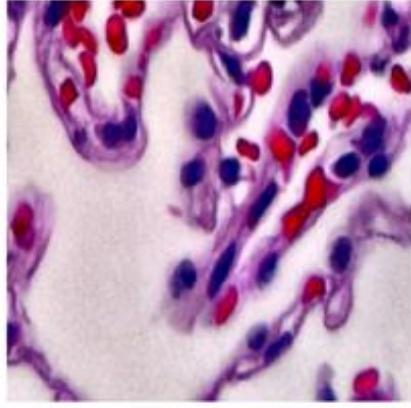
Fuente: Obayya, M., Arasi, M. A., Alruwais, N., Alsini, R., Mohamed, A., & Yaseen, I. (2023). Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model. *IEEE Access*.

Dataset

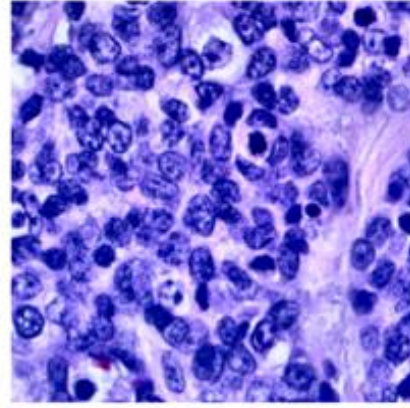
Lung Adenocarcinoma



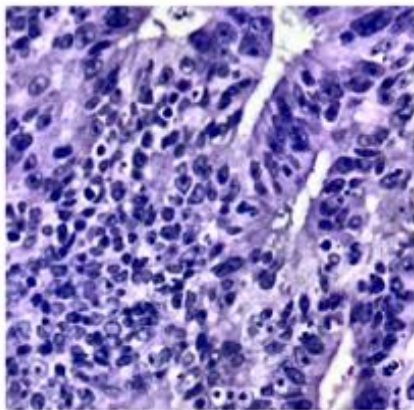
Lung Benign Tissue



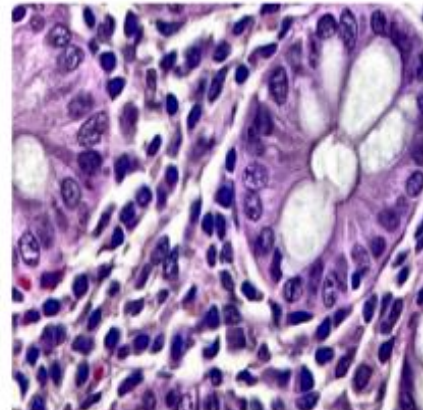
Lung Squamous Cell Carcinoma



Colon Adenocarcinoma



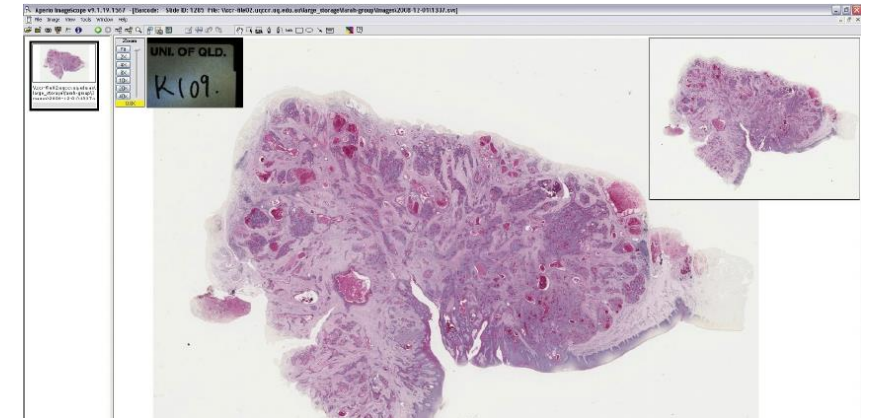
Colon Benign Tissue



Leica
BIO SYSTEMS

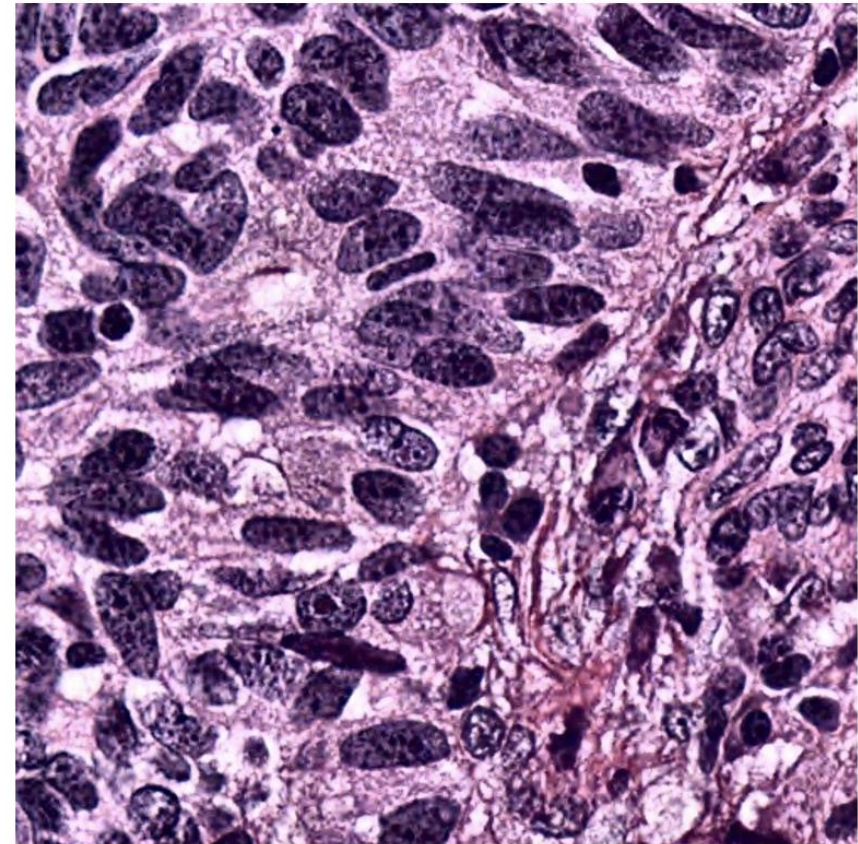
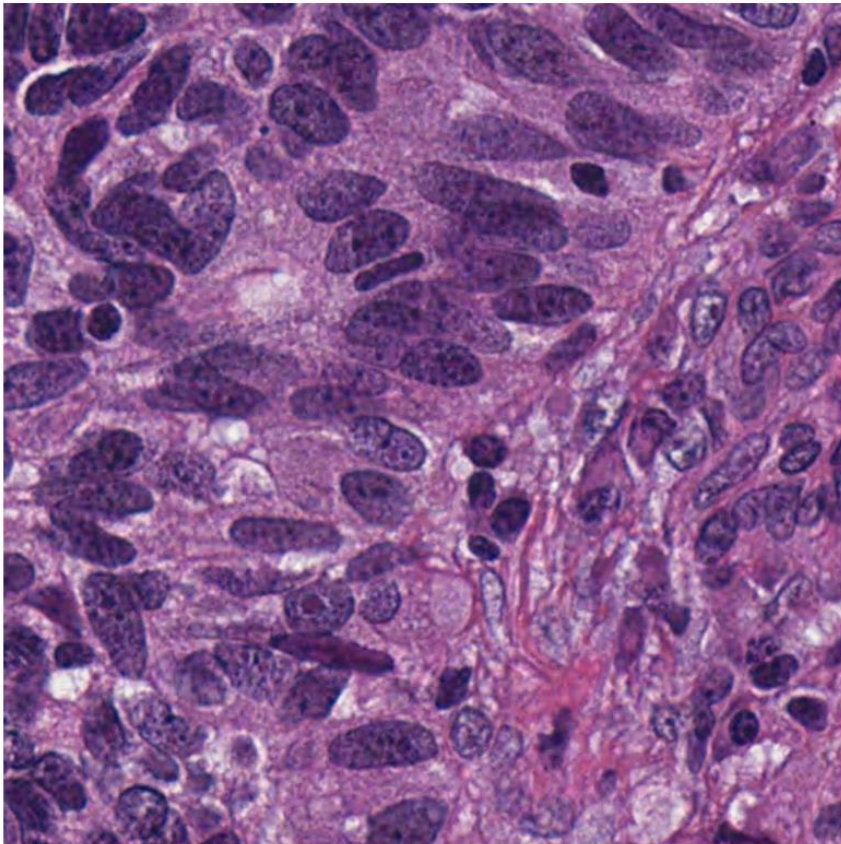


NATIONAL
CANCER
INSTITUTE



Las nuevas imágenes fueron capturadas con un Ampere Image Scope a 20x y 40x

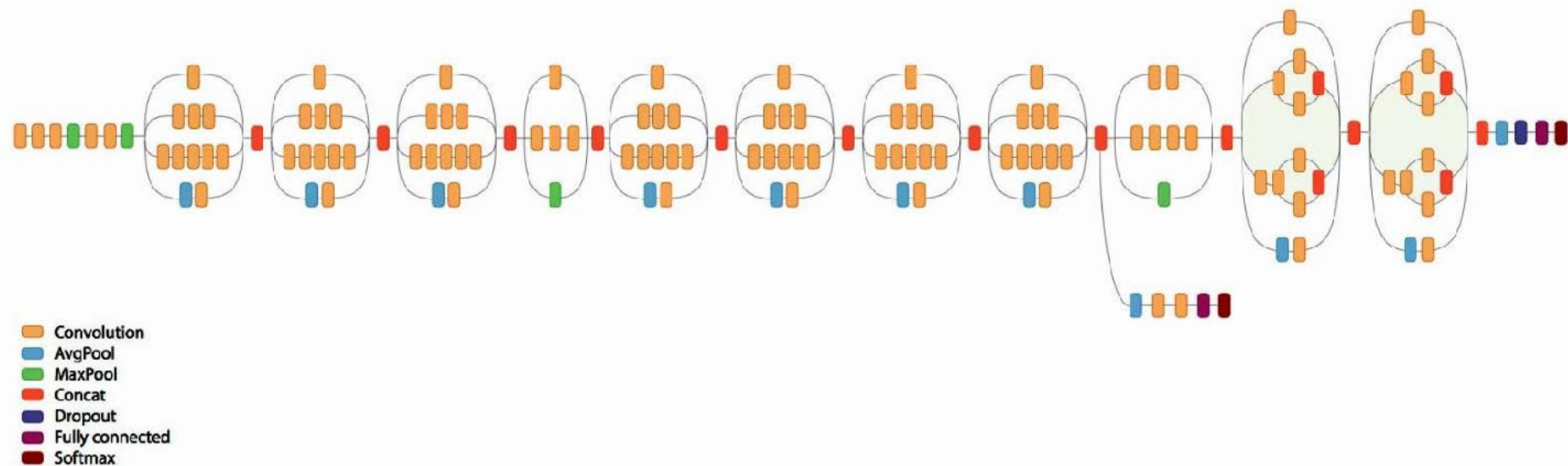
Mejora de Contraste de Imágenes con CLAHE



Tejido de cáncer de pulmón adenocarcinoma

Descripción general de la arquitectura de InceptionResNetV2

InceptionResNetV2 combina las ventajas de los módulos Inception y las conexiones residuales, mejorando la extracción de características y la eficiencia computacional, lo que lo hace particularmente eficaz para tareas complejas de clasificación de imágenes en el campo de las imágenes médicas.



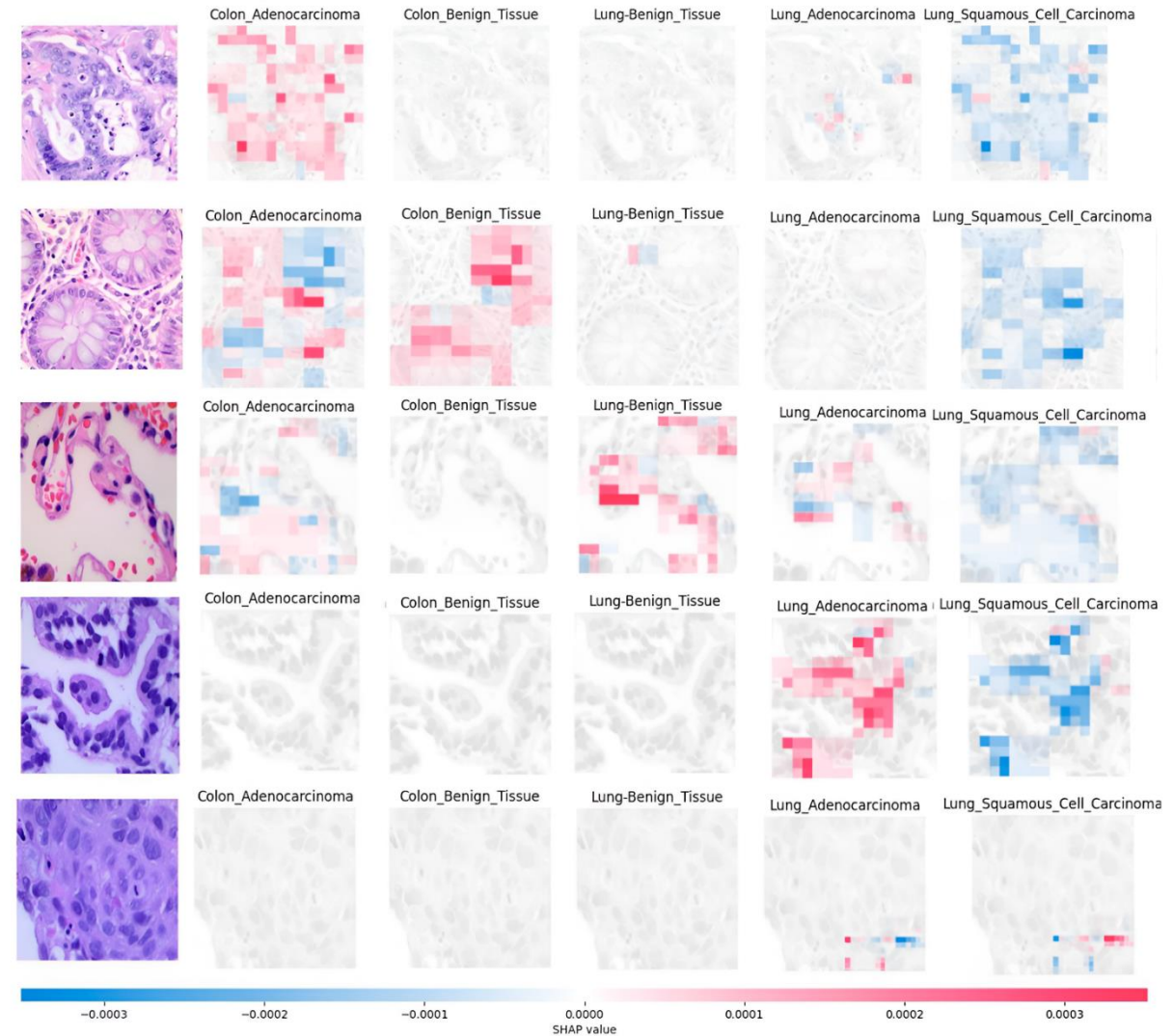
Las dimensiones de entrada
se establecieron en
224x224x3

Adamax

80:20% Train:Test

Valores SHAP para obtener información sobre características

El estudio utiliza valores SHAP para cuantificar la contribución de las características individuales en las predicciones del modelo, lo que mejora la interpretabilidad al revelar cómo las características histopatológicas específicas influyen en los resultados de diagnóstico de los cánceres de pulmón y colon.





Resultados

Evaluación de métricas de resultados

La pérdida de entrenamiento disminuyó significativamente en las primeras épocas y se estabilizó alrededor de la décima época, coincidiendo con la mayor pérdida de validación.

Se implementó una detención temprana para detener el entrenamiento después de la décima época y evitar el sobreajuste.

Box

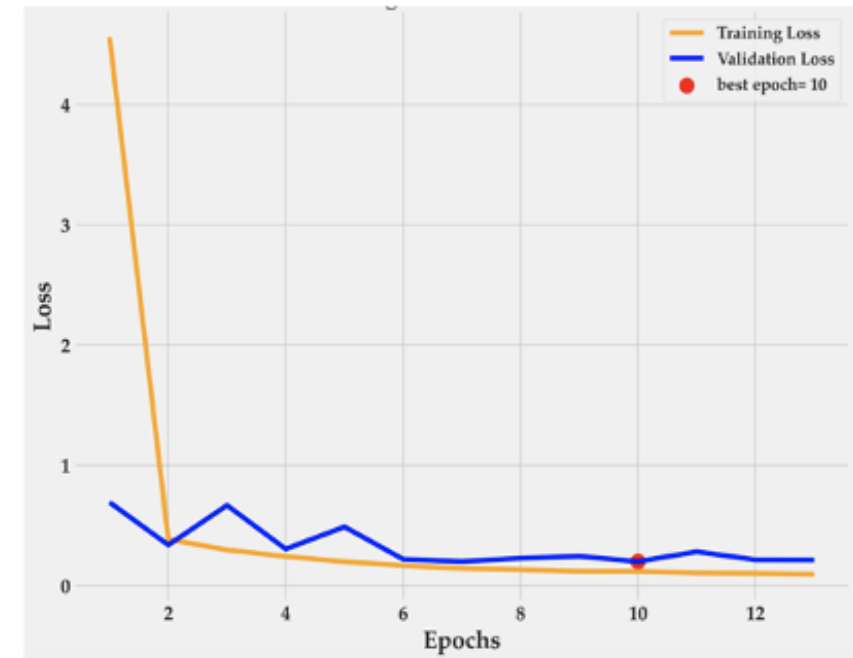


Figure 2

Training and validation loss over epochs, with the best validation loss at epoch 10 (red point).

Source: Own elaboration

Evaluación métrica

La precisión de la validación alcanzó un máximo del 96 % en la época 6, pero luego fluctuó, lo que sugiere que el modelo estaba cerca de su límite de generalización. Estas observaciones resaltan la importancia de equilibrar el entrenamiento y la validación para evitar el sobreajuste.

Box

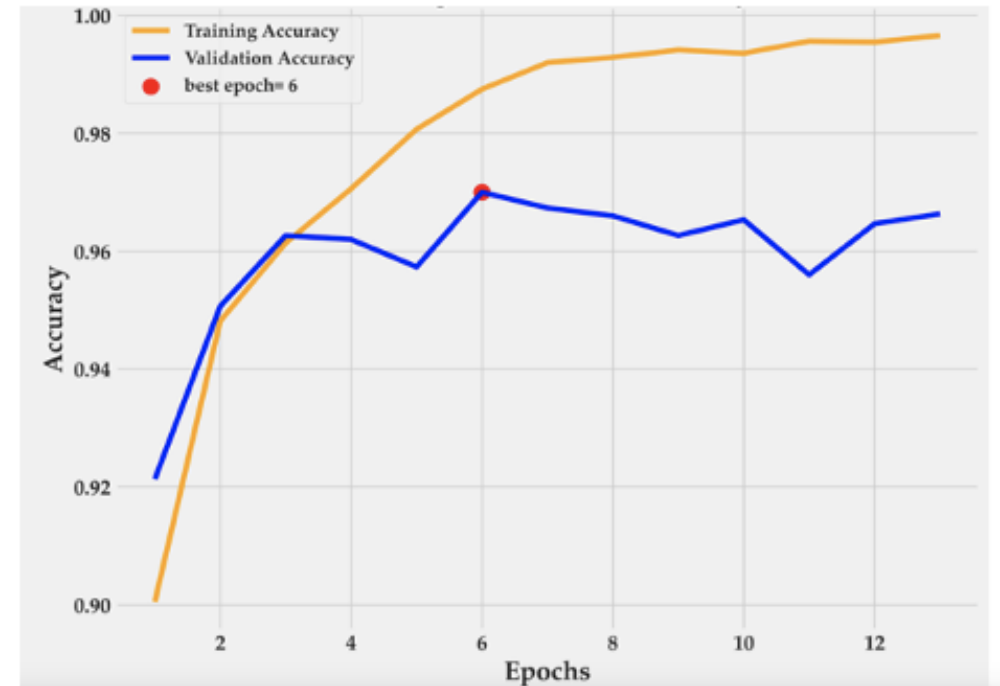


Figure 3

Training and validation accuracy over epochs, with the highest validation accuracy at epoch 6 (red point)

Source: Own elaboration

Evaluación métrica

Box

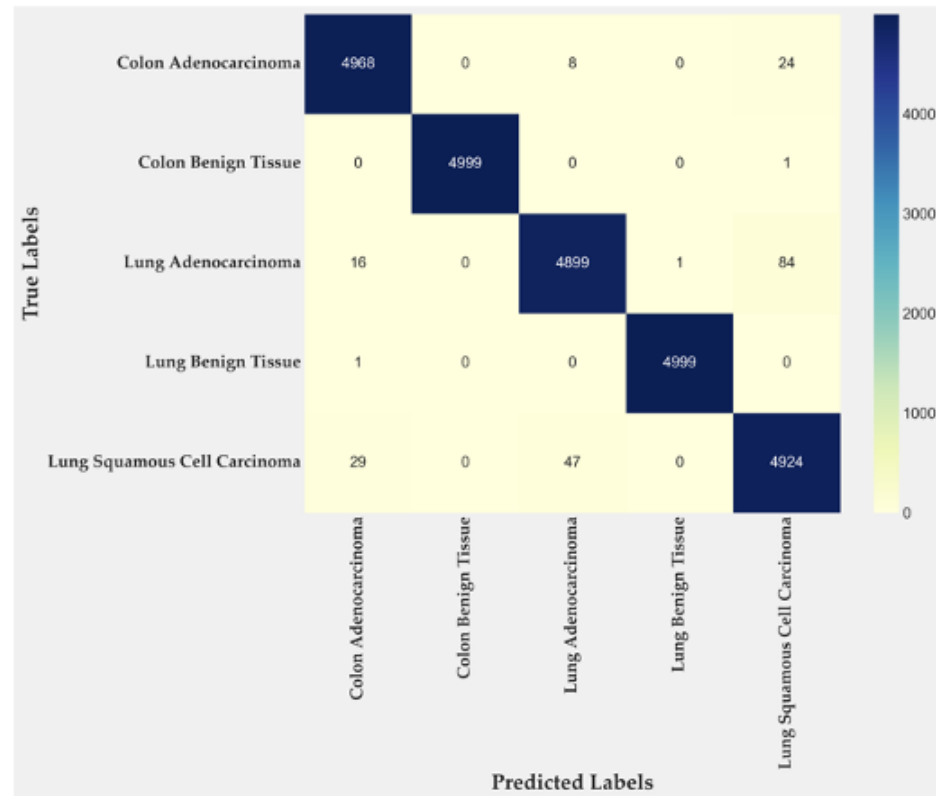


Figure 4

Confusion matrix for the classification of the entire dataset

Source: Own elaboration

Evaluación métrica

La Tabla 1 muestra una efectividad superior al 99% en tejidos benignos, mientras que en tejidos malignos oscila entre el 97% y el 99%. El valor medio es del 99,15% en precisión.

Box

Table 1

Report classification of the entire dataset.

Class	Precision (%)	Recall (%)	F1-Score (%)
Colon Adenocarcinoma	99.08	99.36	99.22
Colon Benign Tissue	100	99.98	99.99
Lung Adenocarcinoma	98.88	97.98	98.43
Lung Benign Tissue	99.98	99.98	99.98
Lung Squamous Cell Carcinoma	97.83	98.48	98.15

Source: Own elaboration

Evaluación métrica

Box

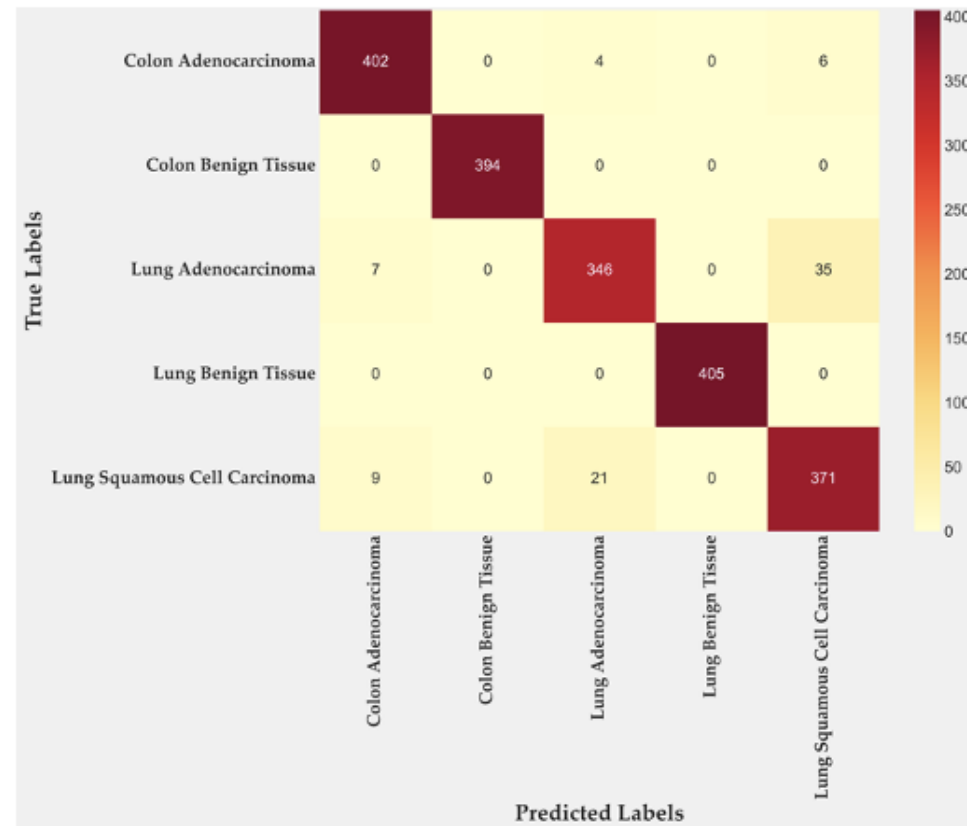


Figure 5

Confusion matrix illustrating the model's performance on the testing dataset

Source: Own elaboration

Evaluación métrica

Box

Table 2

Report classification of the testing dataset.

Class	Precision (%)	Recall (%)	F1-Score (%)
Colon Adenocarcinoma	96.17	97.57	96.86
Colon Benign Tissue	100	100	100
Lung Adenocarcinoma	93.26	89.17	91.17
Lung Benign Tissue	100	100	100
Lung Squamous Cell Carcinoma	90.04	95.90	95.89

Source: Own elaboration

Metric Evaluation

El modelo mostró un buen desempeño en las clases de adenocarcinoma de pulmón y tejido pulmonar benigno, clasificando con precisión 4004 y 3987 muestras, respectivamente.

Box

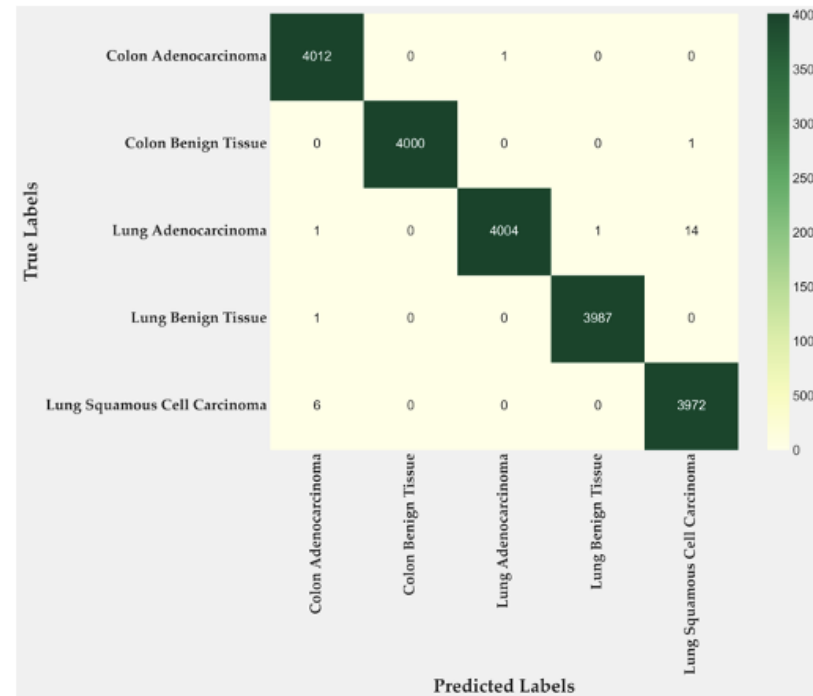


Figure 6

Confusion matrix illustrating the model's performance on the training dataset

Source: Own elaboration

Evaluación métrica

Box

Table 3

Report classification of the training dataset.

Class	Precision (%)	Recall (%)	F1-Score (%)
Colon Adenocarcinoma	99.80	99.97	99.88
Colon Benign Tissue	100	99.97	99.98
Lung Adenocarcinoma	99.97	99.60	99.78
Lung Benign Tissue	99.97	99.97	99.97
Lung Squamous Cell Carcinoma	99.62	99.84	99.73

Source: Own elaboration

Evaluación métrica integral

Análisis comparativo del modelo en diferentes conjuntos de datos, utilizando métricas clave como precisión, exactitud, recuperación, puntuación F1 y coeficiente de compensación de Matthews (MCC).

Box

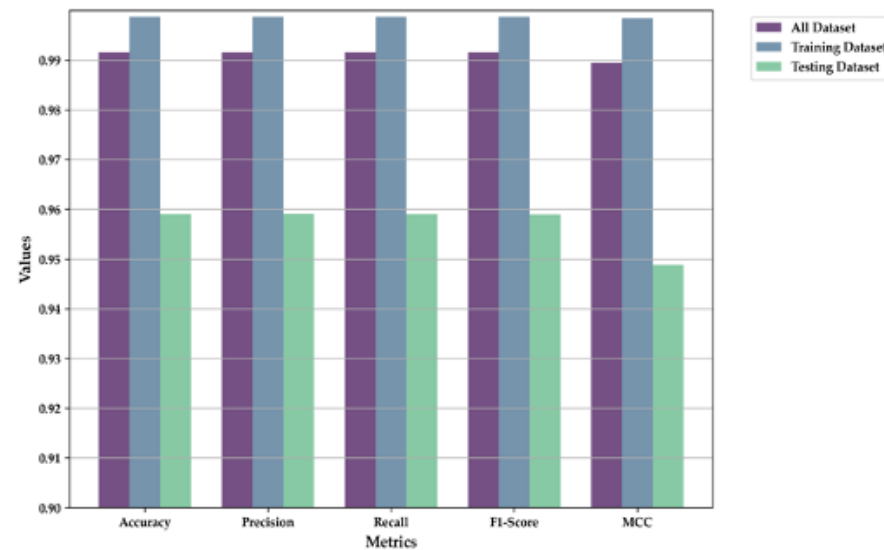


Figure 7

Comparison of average Accuracy, Precision, Recall, F1-Score, and MCC across the training, testing, and full dataset.

Source: Own elaboration

Evaluación métrica

Las curvas ROC para cinco clases histopatológicas en el conjunto de datos de prueba. Estas curvas evalúan la capacidad del modelo para diferenciar entre verdaderos positivos y falsos positivos en cada clase, y el área bajo la curva (AUC) actúa como un indicador clave de rendimiento, donde TPR representa la tasa de verdaderos positivos o el recuerdo.

Box

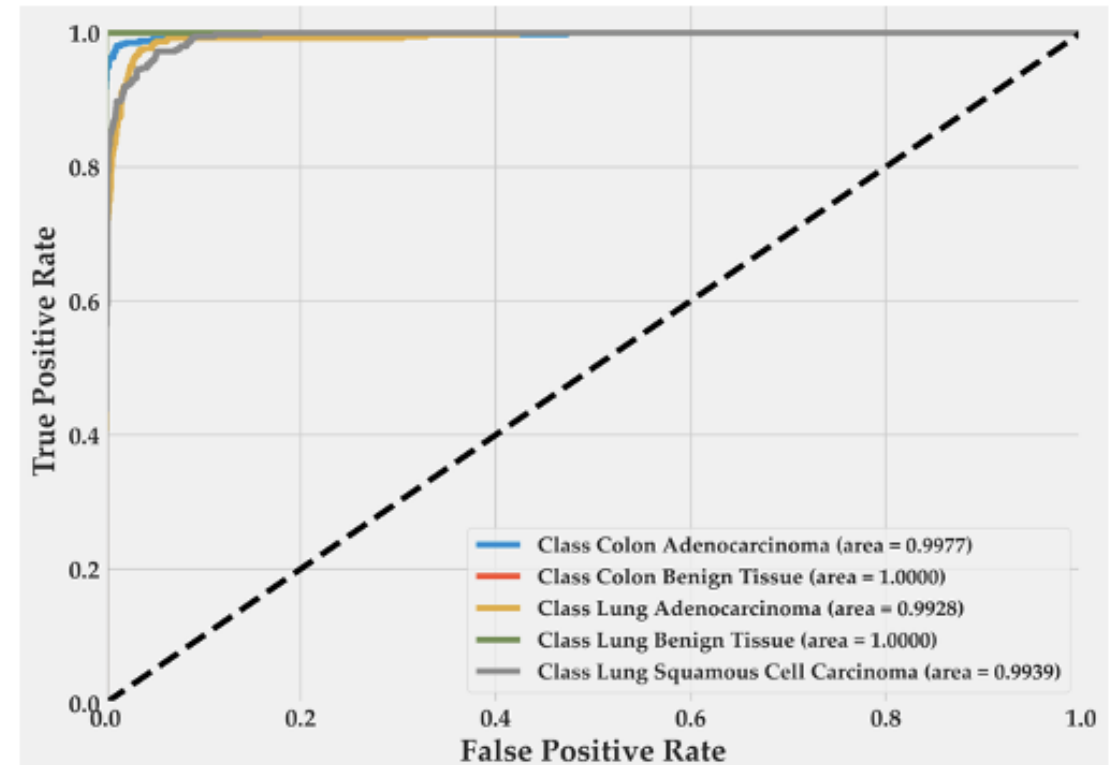


Figure 8

ROC curves for the testing dataset.

Source: Own elaboration

Evaluación métrica

Las curvas ROC para cinco clases histopatológicas en el conjunto de datos de prueba. Estas curvas evalúan la capacidad del modelo para diferenciar entre verdaderos positivos y falsos positivos en cada clase, y el área bajo la curva (AUC) actúa como un indicador clave de rendimiento, donde TPR representa la tasa de verdaderos positivos o el recuerdo.

Box

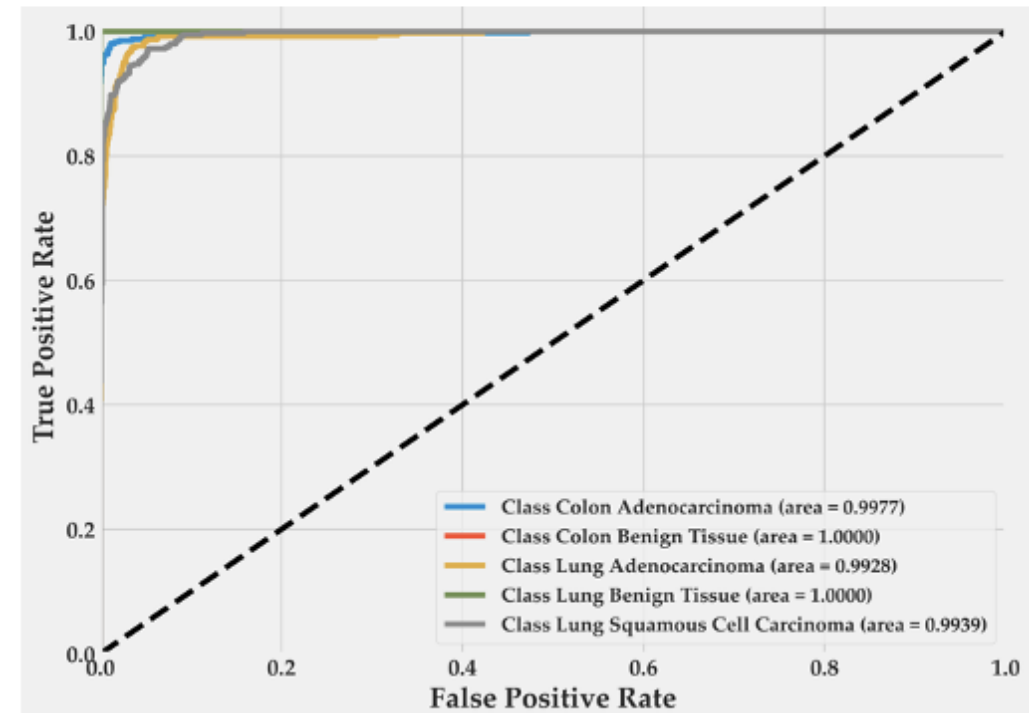


Figure 8

ROC curves for the testing dataset.

Source: Own elaboration

Evaluación métrica

Las curvas de precisión-recuperación son particularmente valiosas en conjuntos de datos desequilibrados en cuanto a clases, ya que se centran en la capacidad del modelo para predecir correctamente los casos positivos.

Box

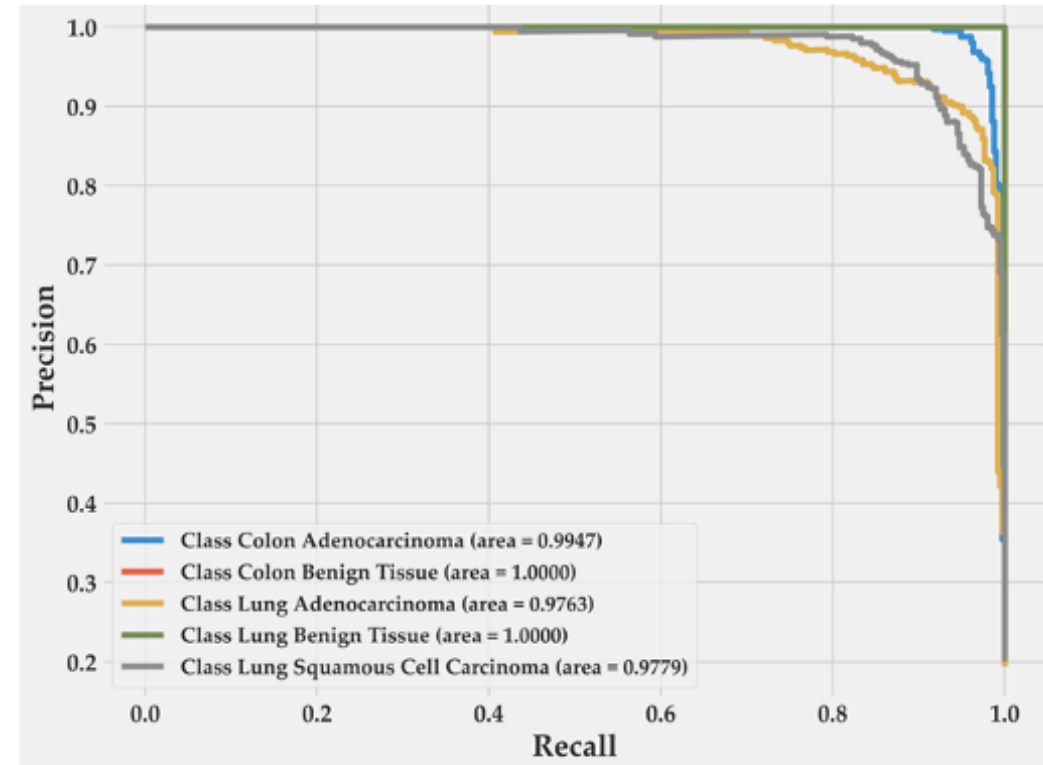


Figure 9

Precision-Recall curves for the testing dataset

Source: Own elaboration

Box

El análisis del valor SHAP para varias imágenes histopatológicas en cinco clases de cáncer, interpretando la contribución de las características individuales a las predicciones del modelo. Las áreas rojas indican características que aumentan la probabilidad de una clase específica, mientras que las áreas azules la disminuyen. La intensidad del color refleja la magnitud del valor SHAP.

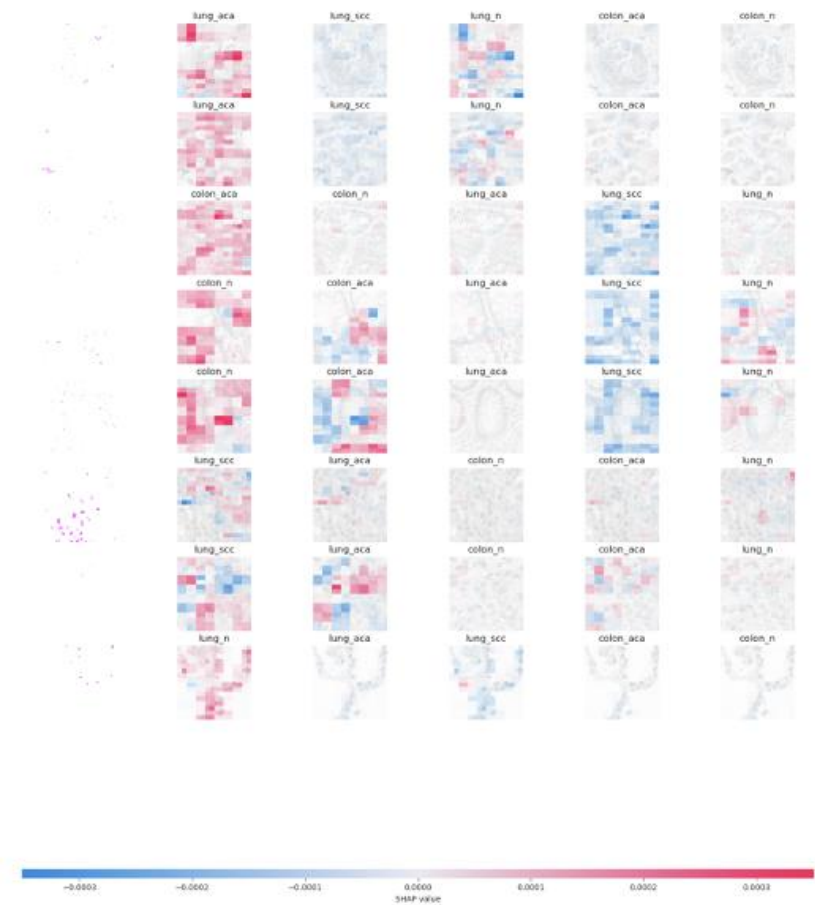
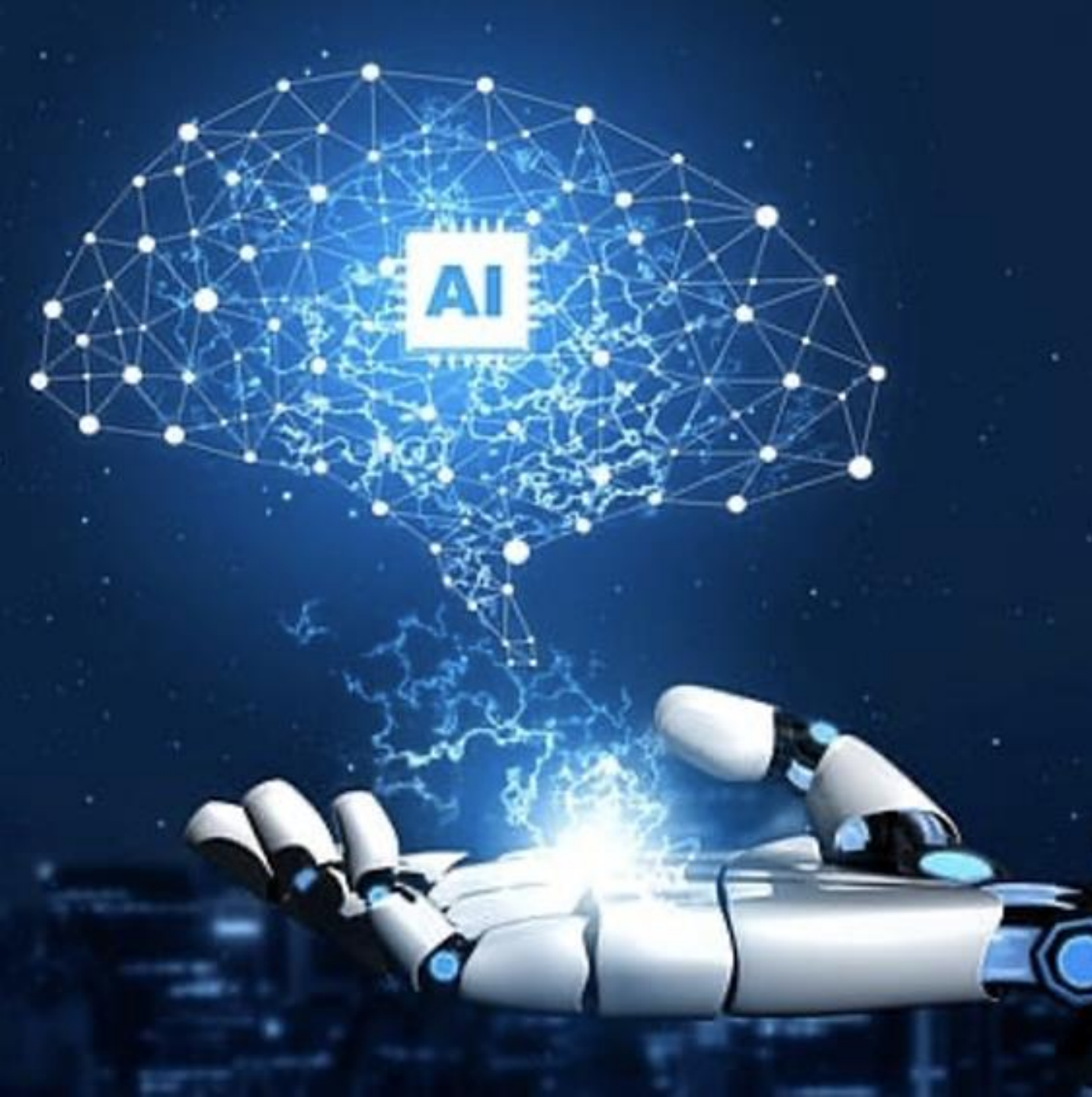


Figure 10

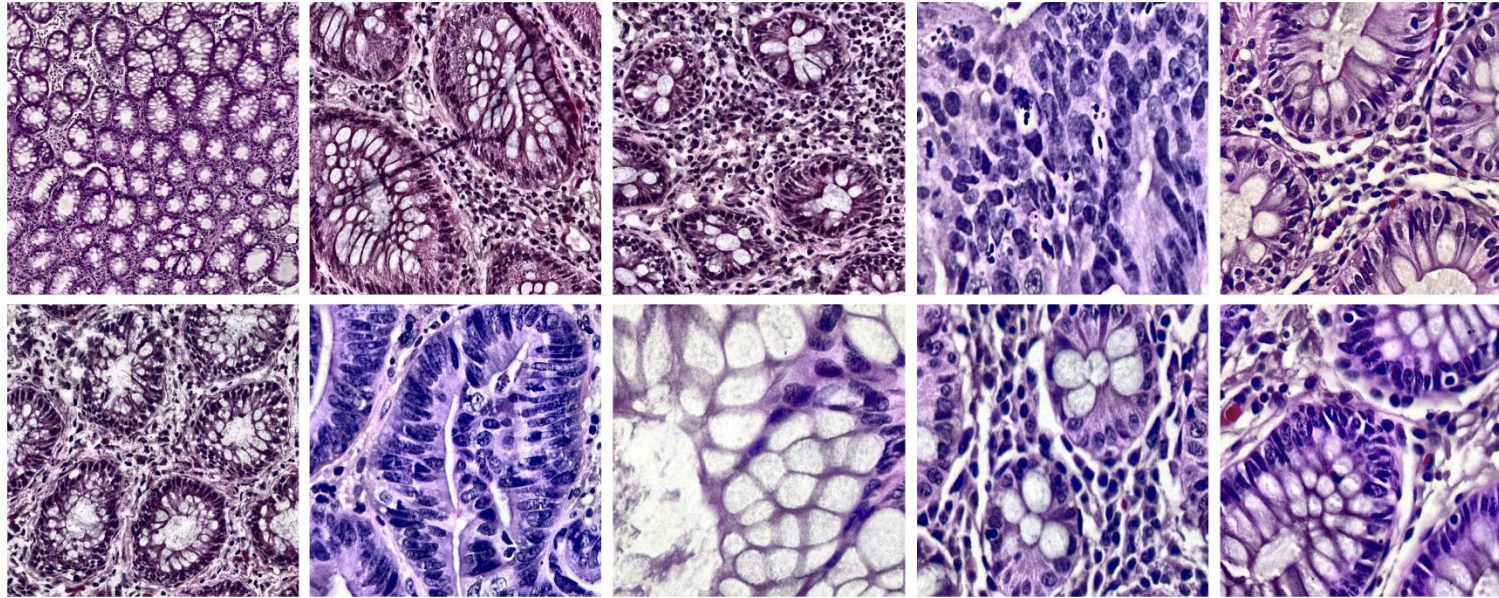
SHAP value analysis across multiple histopathological images.

Source: Own elaboration



Conclusiones

Conclusiones



Se sugiere ampliar el conjunto de datos con muestras más diversas para mejorar la generalización. Se recomienda optimizar el proceso de extracción y clasificación de características y realizar validaciones en entornos clínicos. En resumen, el método es prometedor para la detección automatizada del cáncer, con potencial para futuras mejoras que aumenten su aplicabilidad clínica.

En este estudio se utilizó la arquitectura Inception-ResNetV2 y el aprendizaje por transferencia para clasificar imágenes histopatológicas para la detección de cáncer de pulmón y colon. El modelo mostró un rendimiento excepcional, con una precisión del 95,90 % y métricas de rendimiento casi perfectas en el conjunto de datos de entrenamiento, que superaron el 99,87 %.

Agradecimientos

Nos gustaría expresar nuestro agradecimiento al Tecnológico Nacional de México (TecNM) y al Instituto Tecnológico de Ciudad Guzmán por su apoyo en esta investigación. Además, agradecemos al Dr. Andrew A. Borkowski, la Dra. Marilyn M. Bui, el Dr. L. Brannon Thomas, la MT Catherine P. Wilson, la RN Lauren A. DeLand y el Dr. Stephen M. Mastorides por proporcionar el conjunto de datos LC25000 utilizado. en este estudio.



Referencias

Referencias

Antecedentes

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). [Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries](#). *CA: a cancer journal for clinicians*, 71(3), 209-249.

Kiri, S., & Ryba, T. (2024). [Cancer, metastasis, and the epigenome](#). *Molecular Cancer*, 23(1), 154.

Li, N., Zhai, Z., Zheng, Y., Lin, S., Deng, Y., Xiang, G., ... & Wang, M. (2021). [Association of 13 occupational carcinogens in patients with cancer, individually and collectively, 1990-2017](#). *JAMA Network Open*, 4(2), e2037530-e2037530.

Referencias

Básicas

Cao, M., Li, H., Sun, D., & Chen, W. (2020). [Cancer burden of major cancers in China: a need for sustainable actions](#). *Cancer Communications*, 40(5), 205-210.

Ranasinghe, R., Mathai, M., & Zulli, A. (2022). [A synopsis of modern-day colorectal cancer: Where we stand](#). *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1877(2), 188699.

Crosby, D., Bhatia, S., Brindle, K. M., Coussens, L. M., Dive, C., Emberton, M., ... & Balasubramanian, S. (2022). [Early detection of cancer](#). *Science*, 375(6586), eaay9040.

Referencias

Soporte

Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsaftaris, S. A., Young, A., ... & Pattichis, C. S. (2020). [AI in medical imaging informatics: current challenges and future directions](#). *IEEE journal of biomedical and health informatics*, 24(7), 1837-1857.

Ahmed, A. A., Abouzid, M., & Kaczmarek, E. (2022). [Deep learning approaches in histopathology](#). *Cancers*, 14(21), 5264.

Alabdulqader, E. A., Umer, M., Alnowaiser, K., Wang, H., Alarfaj, A. A., & Ashraf, I. (2024). [Image Processing-based Resource-Efficient Transfer Learning Approach for Cancer Detection Employing Local Binary Pattern Features](#). *Mobile Networks and Applications*, 1-17.

Tummala, S., Kadry, S., Nadeem, A., Rauf, H. T., & Gul, N. (2023). [An explainable classification method based on complex scaling in histopathology images for lung and colon cancer](#). *Diagnostics*, 13(9), 1594.

Hasan, M. A., Haque, F., Sabuj, S. R., Sarker, H., Goni, M. O. F., Rahman, F., & Rashid, M. M. (2024). [An End-to-End Lightweight Multi-Scale CNN for the Classification of Lung and Colon Cancer with XAI Integration](#). *Technologies*, 12(4), 56.

Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., & Mastorides, S. M. (2019). [Lung and colon cancer histopathological image dataset \(lc25000\)](#). *arXiv preprint arXiv:1912.12142*.

Referencias

Diferencias

Al-Jabbar, M., Alshahrani, M., Senan, E. M., & Ahmed, I. A. (2023). [Histopathological Analysis for Detecting Lung and Colon Cancer Malignancies Using Hybrid Systems with Fused Features](#). *Bioengineering*, 10(3), 383.

Kharel, N., Alsadoon, A., Prasad, P. W. C., & Elchouemi, A. (2017, April). [Early diagnosis of breast cancer using contrast limited adaptive histogram equalization \(CLAHE\) and Morphology methods](#). In *2017 8th International Conference on Information and Communication Systems (ICICS)* (pp. 120-124). IEEE.

Yahia Ibrahim, N., & Talaat, A. S. (2022). [An enhancement technique to diagnose colon and lung cancer by using double CLAHE and deep learning](#). *International Journal of Advanced Computer Science and Applications*, 13(8).

Raju, M. S. N., & Rao, B. S. (2022). [Classification of Colon and Lung Cancer Through Analysis of Histopathology Images Using Deep Learning Models](#). *Ingenierie des Systemes d'Information*, 27(6).

Mengash, H. A., Alamgeer, M., Maashi, M., Othman, M., Hamza, M. A., Ibrahim, S. S., ... & Yaseen, I. (2023). [Leveraging marine predators algorithm with deep learning for lung and colon cancer diagnosis](#). *Cancers*, 15(5), 1591.

Referencias

Diferencias

Alqahtani, H., Alabdulkreem, E., Alotaibi, F., Alnfiai, M. M., Singla, C., & Salama, A. S. (2023). [Improved Water Strider Algorithm With Convolutional Autoencoder for Lung and Colon Cancer Detection on Histopathological Images](#). *IEEE Access*.

Obayya, M., Arasi, M. A., Alruwais, N., Alsini, R., Mohamed, A., & Yaseen, I. (2023). [Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model](#). *IEEE Access*.

Singh, O., Kashyap, K. L., & Singh, K. K. (2024). [Lung and Colon Cancer Classification of Histopathology Images Using Convolutional Neural Network](#). *SN Computer Science*, 5(2), 223.

Farhadipour, A. (2024). [Lung and colon cancer detection with convolutional neural networks and adaptive histogram equalization](#). *Iran Journal of Computer Science*, 7(2), 381-395.

Moin, M. B., Faria, F. T. J., Saha, S., Rafa, B. K., & Alam, M. S. (2024). [Exploring Explainable AI Techniques for Improved Interpretability in Lung and Colon Cancer Classification](#). *arXiv preprint arXiv:2405.04610*.

Referencias

Discusiones

Neshat, M., Ahmed, M., Askari, H., Thilakaratne, M., & Mirjalili, S. (2024). [Hybrid Inception Architecture with Residual Connection: Fine-tuned Inception-ResNet Deep Learning Model for Lung Inflammation Diagnosis from Chest Radiographs](#). *Procedia Computer Science*, 235, 1841-1850.

Mehmood, S., Ghazal, T. M., Khan, M. A., Zubair, M., Naseem, M. T., Faiz, T., & Ahmad, M. (2022). [Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing](#). *IEEE Access*, 10, 25657-25668.



MARVID®

© MARVID-Mexico

No part of this document covered by the Federal Copyright Law may be reproduced, transmitted or used in any form or medium, whether graphic, electronic or mechanical, including but not limited to the following: Citations in articles and comments Bibliographical, compilation of radio or electronic journalistic data. For the effects of articles 13, 162, 163 fraction I, 164 fraction I, 168, 169, 209 fraction III and other relative of the Federal Law of Copyright. Violations: Be forced to prosecute under Mexican copyright law. The use of general descriptive names, registered names, trademarks, in this publication do not imply, uniformly in the absence of a specific statement, that such names are exempt from the relevant protector in laws and regulations of Mexico and therefore free for General use of the international scientific community. BECORFAN is part of the media of MARVID-Mexico., E: 94-443.F: 008- (www.marvid.org/booklets)