



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

**Clasificación de pacientes con
COVID-19 mediante algoritmos de
machine learning fundamentados
en biomarcadores**

TESIS

Que para obtener el título de

Ingeniero en Sistemas Biomédicos

P R E S E N T A

Isael Guillermo Garcia Macedo

DIRECTOR DE TESIS

Dr. Didier Torres Guzmán



Ciudad Universitaria, Cd. Mx., 2026



**PROTESTA UNIVERSITARIA DE INTEGRIDAD Y
HONESTIDAD ACADÉMICA Y PROFESIONAL
(Titulación con trabajo escrito)**



De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado CLASIFICACION DE PACIENTES CON COVID-19 MEDIANTE ALGORITMOS DE MACHINE LEARNING FUNDAMENTADOS EN BIOMARCADORES que presenté para obtener el título de INGENIERO EN SISTEMAS BIOMÉDICOS es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi Entidad Académica, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de titulación.

ISAIEL GUILLERMO GARCIA MACEDO
Número de cuenta: 317259690

Dedicatoria

El presente trabajo va dirigido a mi abuela (Hortensia Mendoza Minigo) que sin su apoyo, sacrificio y empeño en mi cuidado y crecimiento es claro que no estaría aquí; con una vida llena de oportunidades, rodeada de gente que me quiere y aprecia. Gracias a ti, tu bondad, empatía y gran corazón no sería el hombre que soy, mis ideas y moral con las cuales dirijo mi vida serían muy diferentes sin tu compañía abuela. Ya puedes estar tranquila, que con este trabajo ya has finalizado esta tarea que te encomendaste, que fue la de cuidarme.

Agradecimiento

Una de las etapas más importantes de mi vida acaba con la presentación de este trabajo, es claro que sin el apoyo y cariño de todas las personas que me acompañaron no estaría culminando mi carrera con todo el crecimiento que conllevo.

Quiero agradecer:

A mi familia por haber sido mi principal apoyo en todos los años de mi vida, mi abuelo (Adán Timoteo García Gonzalez), a mi tío (Juan Carlos Garcia Mendoza), a mi hermana (Cinthya García Macedo), a mi tía (Patricia García Mendoza), a mi padre (Pedro García Mendoza) y claramente mi abuela; sin sus cuidados y apoyos en las diferentes etapas de mi crecimiento no podría haber llegado lejos.

A mi segunda familia que los conozco desde que era niño y que sin su guía no estaría en la UNAM, aquí se encuentra mi mejor amigo (Marco Armando Gonzalez Morales), Ricardo Antonio Gonzalez Morales, Claudia Morales Ramírez y Marco Antonio Gonzalez Villa; sin su intervención en diferentes momentos de mi vida es claro que no tendría la paz con la que gozo hoy.

A Mary (Mariana Samantha Zurita Perez), mi compañera, quien en estos casi seis años de relación me ha brindado un acompañamiento que me ha permitido cuestionar y cambiar aspectos de mi que no me gustaban, además de siempre creer en mí y en los momentos más difíciles de la carrera ayudarme a levantarme; también agradezco a sus padres (Olga y Ángel) por siempre darme un espacio en su hogar y hacerme parte de su familia.

Al Grupo de Acción Revolucionaria por darme un espacio de organización y lucha; por demostrarme que con el trabajo colectivo podemos cambiar nuestras condiciones materiales, sin sus aprendizajes y trabajo conjunto no hubiera podido ayudar a los diferentes estudiantes que se acercaron al CEAFL en su momento. Aquí se encuentran Nikté, Natalia, Diego, Octavio, Rosh, Olga, Dulce, Osmany, Alejandro, Aideé, Elisa,

Clarisa, Alejandra, Daniel, Esteban, Arturo y demás personas que no puedo mencionar pero que igualmente tuvieron un gran impacto.

A mi asesor por darme una guía en este año de desarrollo de la tesis para poder culminar esta etapa de mi vida.

Es claro que me hacen falta muchas personas más por agradecer; que ayudaron y aportaron en mi camino y quiero que sepan que los llevo conmigo.

Por la revolución proletaria y una palestina libre

Índice

Dedicatoria.....	2
Agradecimiento.....	3
Capítulo 1. Aspectos Metodológicos.....	8
1.1. Introducción.....	8
1.2. Planteamiento del problema.....	10
1.3. Hipótesis.....	10
1.3.1. Objetivo.....	11
1.4. Antecedentes.....	12
1.4.1. Evolución de la segmentación pulmonar.....	12
1.4.2. Clasificación de patologías mediante IA.....	13
1.5. Propuesta de solución.....	14
Capítulo 2. Marco teórico.....	17
2.1. Segmentación y procesamiento de imágenes médicas.....	17
2.1.1. Métodos geométricos y tradicionales.....	18
2.1.2. Aprendizaje profundo en segmentación.....	19
2.1.3. Retos técnicos en patologías severas.....	19
2.2. Biomarcadores.....	20
2.2.1. Biomarcadores morfológicos.....	20
2.2.2. Cuantificación del daño parenquimatoso.....	21
2.2.3. Análisis de textura del parénquima.....	22
2.3. Algoritmos de machine learning.....	24
2.3.1. Regresión logística.....	25
2.3.2. K-vecinos más cercanos.....	25
2.3.3. Máquina de vectores de soporte.....	26
2.3.4. Bosque aleatorio (Random Forest).....	26
2.3.5. Extreme gradient boosting (XGBoost).....	26

2.3.6. Redes neuronales artificiales: Perceptrón Multicapa (MLP).....	27
Capítulo 3. Metodología.....	29
3.1. Materiales y cohortes de estudio.....	29
3.1.1. Origen de los datos.....	29
3.1.2. Criterios de inclusión y exclusión.....	30
3.2. Técnicas de procesamiento.....	31
3.2.1. Adquisición y selección interactiva de la región de interés.....	31
3.2.2. Mejora de contraste y normalización.....	31
3.2.3. Reducción de ruido mediante filtrado espacial.....	32
3.2.4. Umbralización dual y limitaciones del método de Otsu.....	32
3.2.5. Suavizado de bordes con filtro Savitzky-Golay.....	33
3.2.6. Algoritmo de corrección de bordes para daño pulmonar severo.....	34
3.3. Validación del algoritmo.....	35
3.3.1. Procedimiento con phantom 3D y referencia manual.....	35
3.3.2. Métricas de desempeño volumétrico.....	37
3.4. Obtención de los biomarcadores.....	38
3.4.1. Cuantificación y caracterización de lesiones (ROI).....	38
3.4.2. Análisis de textura mediante GLCM.....	39
3.4.3. Biomarcadores morfológicos estructurales.....	40
3.5. Entrenamiento y validación de modelos de machine learning.....	41
3.5.1. Preparación y partición del conjunto de datos.....	42
3.5.2. Optimización de hiperparámetros mediante Grid Search.....	42
3.5.3. Métricas de evaluación del desempeño.....	43
3.5.4. Análisis de importancia de características e interpretabilidad.....	45
Capítulo 4. Resultados y discusión.....	45
4.1. Análisis del resultado de segmentación.....	45
4.2. Análisis exploratorio de datos.....	48
4.2.1. Balance de la variable objetivo.....	48

4.2.2. Poder discriminante de los biomarcadores.....	49
4.2.3. Mapa de calor de correlación.....	50
4.3. Entrenamiento y optimización de los modelos.....	52
4.3.1. Fundamentos de la validación cruzada (k-fold CV).....	53
4.3.2. Proceso de optimización por algoritmo.....	53
A. Regresión Logística (Logistic Regression).....	53
B. Bosques Aleatorios (Random Forest).....	54
C. Máquinas de Soporte Vectorial (SVM - RBF).....	55
D. K-Vecinos Más Cercanos (KNN).....	56
E. Red Neuronal Artificial (MLP).....	56
F. XGBoost (Extreme Gradient Boosting).....	57
4.4. Evaluación de los modelos sobre el conjunto de prueba.....	58
4.4.1. Análisis de la evaluación.....	58
4.4.2. Análisis de modelos de alto desempeño: XGBoost, Random Forest y MLP.....	59
4.4.3. Evaluación de modelos de estabilidad robusta: Regresión logística, KNN y SVM..	61
4.5. Análisis de importancia de variables e interpretabilidad.....	63
4.5.1. Interpretabilidad en el modelo líder: XGBoost.....	63
4.5.2. Interpretabilidad en la red neuronal: MLP.....	65
4.5.3. Interpretabilidad en el modelo KNN.....	68
4.5.4. Interpretabilidad en el modelo SVM.....	70
4.5.5. Interpretabilidad en el modelo Random Forest.....	71
4.5.6. Interpretabilidad en el modelo Regresión Logística.....	74
4.6. Discusión.....	76
Conclusiones.....	78

Capítulo 1. Aspectos Metodológicos

1.1. Introducción

La tomografía computarizada (TC) de tórax se ha consolidado como una herramienta diagnóstica fundamental en neumología, permitiendo la visualización detallada del parénquima pulmonar, las vías respiratorias y la vasculatura (Yan et al., 2021; Akbari et al., 2021). Su evolución tecnológica, desde la TC helicoidal hasta la TC de alta resolución y los escáneres multidetectores, ha proporcionado imágenes tridimensionales de alta resolución que superan ampliamente las capacidades de la radiografía tradicional (Turk & Kılıçaslan, 2025; Poola et al., 2023). Esta riqueza de datos anatómicos es crucial para la evaluación de enfermedades como el cáncer de pulmón, las enfermedades infecciosas como la COVID-19, y las patologías crónicas como la EPOC y las enfermedades intersticiales (Yan et al., 2021; Akbari et al., 2021).

Sin embargo, existe una carga de trabajo considerable para los radiólogos, quienes deben evaluar, caracterizar y detectar características pulmonares en un gran volumen de imágenes; un proceso que a menudo es lento y susceptible a la variabilidad interobservador (Poola et al., 2023; Yan et al., 2021). Ante esto, los avances tecnológicos, particularmente en inteligencia artificial, ofrecen herramientas prometedoras para automatizar estas tareas, aligerar dicha carga y mejorar la precisión del diagnóstico (Zeng et al., 2024; Akbari et al., 2021).

Recientemente, el aprendizaje profundo (*deep learning*, DL) ha demostrado un rendimiento excepcional en el análisis de imágenes de TC, con redes neuronales convolucionales (CNN) que alcanzan precisiones superiores al 90% en tareas diagnósticas (Akbari et al., 2021; Turk & Kılıçaslan, 2025; Yan et al., 2021). No obstante, estos modelos presentan limitaciones críticas: requieren enormes recursos computacionales para el entrenamiento, son "ávidos de datos" y operan como "cajas negras", con una falta de transparencia que dificulta la confianza clínica (Li, 2023; Yan et

al., 2021). Además, su rendimiento puede degradarse notablemente ante las alteraciones patológicas severas que provocan la fusión pulmonar, ya que no fueron diseñados originalmente para manejar la pérdida de límites anatómicos convencionales (Yan et al., 2021; Li, 2023).

Una alternativa prometedora es el aprendizaje de máquina (*machine learning*, ML) interpretable basado en biomarcadores diseñados, donde características significativas y predefinidas —derivadas del conocimiento clínico— son utilizadas para entrenar modelos más simples y transparentes (Li, 2023). Este paradigma produce modelos cuyas decisiones pueden rastrearse hasta características específicas, alineándose con el razonamiento clínico y los principios de la Inteligencia Artificial Explicable (XAI) (Li, 2023). En particular, los biomarcadores morfológicos de estructuras pulmonares como el árbol vascular y aéreo pulmonar representan una fuente viable de características interpretables que capturan cambios globales en la arquitectura pulmonar ante enfermedades como la COVID-19 (Torres et al., 2024; Yan et al., 2021).

La presente tesis se centra en este enfoque orientado a biomarcadores. Se propone una comparación rigurosa de diversos algoritmos de IA para la clasificación de patologías pulmonares ocasionadas por COVID-19, utilizando biomarcadores morfológicos como la compacidad discreta, el volumen y la tortuosidad. Al integrar métodos avanzados de segmentación —capaces de manejar incluso casos severos de daño pulmonar— con modelos de clasificación supervisada, este trabajo busca demostrar que es posible alcanzar un alto rendimiento diagnóstico manteniendo la eficiencia computacional y la interpretabilidad necesaria para el apoyo al diagnóstico clínico real.

1.2. Planteamiento del problema

La interpretación manual de TC es un proceso lento y propenso a la variabilidad interobservador, lo que ha impulsado el desarrollo de herramientas de segmentación automatizada para la extracción de biomarcadores diagnósticos. Sin embargo, surge una paradoja técnica: la fiabilidad de algoritmos estándar, como la U-Net, disminuye drásticamente en casos de COVID-19 grave debido a la alteración de la morfología del

tejido y a eventos de fusión pulmonar que borran los límites anatómicos convencionales. Esta limitación técnica impide una delimitación precisa de las estructuras, comprometiendo la calidad de los biomarcadores extraídos y reduciendo la confianza en los sistemas de diagnóstico asistido precisamente en los escenarios clínicos donde el análisis cuantitativo es más crítico.

Por otro lado, aunque el DL intenta resolver estas fallas, su implementación se ve obstaculizada por su naturaleza de "caja negra", su alta demanda computacional y la necesidad de conjuntos masivos de datos anotados. Existe una brecha significativa entre los resultados de investigación y la práctica clínica debido a la falta de transparencia y validación externa de estos modelos. Por lo tanto, surge la necesidad de desarrollar enfoques alternativos basados en el ML interpretable. Este paradigma utiliza biomarcadores morfológicos específicos para entrenar modelos más simples y transparentes, cuyas decisiones pueden rastrearse hasta características alineadas con el razonamiento médico y con un uso eficiente de recursos computacionales.

1.3. Hipótesis

La implementación de modelos de ML fundamentados en biomarcadores morfológicos —específicamente compacidad discreta, volumen y tortuosidad— permite alcanzar una precisión diagnóstica comparable a los modelos de DL en la clasificación de patologías pulmonares por COVID-19. A diferencia de los enfoques de "caja negra" característicos del DL, este paradigma orientado a biomarcadores ofrece una interpretabilidad inherente alineada con el razonamiento clínico y los principios de la XAI, optimizando además el uso de recursos computacionales necesarios para el entrenamiento y validación de los algoritmos.

1.3.1. Objetivo

Objetivo General

- Clasificar pacientes con lesiones pulmonares mediante la utilización de algoritmos de inteligencia artificial, integrando técnicas de segmentación robustas y la cuantificación de biomarcadores morfológicos para el apoyo al diagnóstico clínico de patologías ocasionadas por COVID-19.

Objetivos Específicos

- **Optimizar un método de segmentación automatizada:** Desarrollar un algoritmo que delimite con precisión los contornos pulmonares e incluya las lesiones en pacientes con cuadros severos de COVID-19, implementando una lógica de respaldo diseñada para abordar eventos de fusión pulmonar.
- **Extraer y cuantificar biomarcadores estructurales:** Obtener un vector de características que integre biomarcadores morfológicos (compacidad discreta, volumen y tortuosidad), estructuras pulmonares (árboles vascular y de las vías respiratorias), así como métricas de textura para caracterizar de forma sistémica la arquitectura pulmonar.
- **Entrenar y optimizar modelos de clasificación:** Implementar y comparar el rendimiento de seis algoritmos de ML supervisado mediante búsqueda en cuadrícula para maximizar la exactitud diagnóstica.
- **Evaluar el rendimiento diagnóstico:** Validar los modelos optimizados utilizando métricas de exactitud, precisión, sensibilidad, F1-score y el área bajo la curva (AUC-ROC).
- **Validar la interpretabilidad de los modelos:** Aplicar técnicas de importancia por permutación y valores Shapley Additive Explanations (SHAP) para explicar las decisiones de los modelos, garantizando que el diagnóstico pueda rastrearse hasta biomarcadores específicos alineados con la práctica clínica.

1.4. Antecedentes

El desarrollo de herramientas computacionales para el análisis de enfermedades pulmonares ha evolucionado rápidamente, impulsado principalmente por la necesidad de procesar grandes volúmenes de datos de TC de manera precisa y eficiente. Esta evolución ha permitido que el procesamiento automatizado se convierta en una pieza clave para mitigar la carga de trabajo de los radiólogos y mejorar la precisión del diagnóstico asistido por computadora.

1.4.1. Evolución de la segmentación pulmonar

La segmentación es el proceso fundacional que permite delimitar estructuras de interés para la extracción de biomarcadores cuantitativos cruciales para el diagnóstico y seguimiento clínico (Yan et al., 2021). Tradicionalmente, se han empleado métodos geométricos como el algoritmo de Marcha Adaptativa de Bordes (ABM), el cual recorre el contorno y genera "puentes" sobre las concavidades causadas por los nódulos para evitar la subsegmentación (Pu et al., 2008). Otros enfoques basados en la curvatura local permiten identificar puntos dominantes de inflexión en el contorno para unir y refinar los bordes de manera iterativa hasta alcanzar la convergencia (Singadkar et al., 2021).

En la última década, la arquitectura de red neuronal U-Net se ha establecido como el estándar para la segmentación biomédica debido a su diseño de codificador-decodificador que combina características de alto nivel con información espacial de alta resolución (Akbari et al., 2021; Turk & Kılıçaslan, 2025; Yan et al., 2021). No obstante, se ha reportado que el rendimiento de estos modelos de DL se degrada notablemente ante alteraciones patológicas severas, como la fusión pulmonar, ya que no fueron diseñados originalmente para manejar la pérdida total de límites anatómicos (Yan et al., 2021; Li, 2023). Esta limitación justifica el desarrollo de lógicas de respaldo y métodos de refinamiento de bordes más robustos que incluyan con precisión las lesiones en pacientes con cuadros graves.

1.4.2. Clasificación de patologías mediante IA

La literatura reciente muestra una clara transición de los métodos tradicionales de ML hacia modelos de DL para el análisis de imágenes médicas.

- **Modelos de DL:** Diversos estudios han empleado CNN como ResNet50, Xception y EfficientNet para la clasificación de COVID-19, logrando exactitudes superiores al 95% (Guhan et al., 2022; Hou & Navarro-Cia, 2022; Ifty et al., 2024). Algunos enfoques avanzados utilizan CNN 3D para analizar anomalías en un contexto espacio-temporal, permitiendo distinguir entre COVID-19 y neumonía adquirida en la comunidad con valores de AUC de hasta 0.92 (Harmon et al., 2020). Sin embargo, se ha criticado que estos modelos suelen carecer de validación externa en conjuntos de datos no vistos, lo que puede llevar a una sobreestimación de su rendimiento real y comprometer su generalización (Lee et al., 2023; Roberts et al., 2021).
- **Aprendizaje de máquina e interpretabilidad:** Frente a la naturaleza de "caja negra" del DL, han resurgido propuestas de ML interpretable que utilizan vectores de características radiómicas o biomarcadores morfológicos (Brocki & Chung, 2023; Rahaghi et al., 2021). Estudios comparativos han demostrado que algoritmos como Naïve Bayes o soporte vectorial de máquina (SVM) pueden alcanzar exactitudes del 97% al 99% cuando se combinan con una extracción de características adecuada mediante técnicas como las matrices de co-ocurrencia de niveles de gris (GLCM) (Albataineh et al., 2024; Guhan et al., 2022). Estos modelos ofrecen, además, una complejidad computacional significativamente menor, lo que facilita su implementación en entornos clínicos con recursos limitados (Sarv Ahrabi et al., 2021; Lee et al., 2023).
- **Biomarcadores morfológicos y XAI:** La integración de biomarcadores, como la compacidad discreta y la tortuosidad, permiten alinear las decisiones del modelo con el razonamiento clínico (Torres Guzmán et al., 2024; Brocki & Chung, 2023). A pesar de que algunos modelos interpretables pueden presentar un rendimiento

ligeramente inferior a las arquitecturas de DL más complejas, su capacidad para ser validados por expertos mediante técnicas de XAI los posiciona como una alternativa crítica para la implementación en entornos clínicos reales (Brocki & Chung, 2023; Lee et al., 2023).

En resumen, mientras que el estado del arte ha priorizado la exactitud diagnóstica mediante modelos complejos, persiste una brecha significativa en cuanto a la generalización en casos graves y la transparencia de los algoritmos (Lee et al., 2023; Ahmad et al., 2025). Esta tesis se sitúa en esa intersección, buscando robustez en la segmentación e interpretabilidad en la clasificación mediante un enfoque orientado a biomarcadores.

1.5. Propuesta de solución

La propuesta de esta tesis consiste en el desarrollo de un sistema integral de apoyo al diagnóstico clínico que combina el procesamiento digital de imágenes con algoritmos de ML supervisado. A diferencia de los modelos de "caja negra" desde la entrada y salida, este enfoque se estructura en etapas modulares que permiten supervisar y validar cada fase del proceso diagnóstico, garantizando tanto la precisión como la transparencia.

El flujo de trabajo propuesto se divide en las siguientes fases fundamentales:

1. **Segmentación pulmonar optimizada:** Para abordar las fallas de los métodos tradicionales en casos de patología severa, se implementa una técnica de segmentación que incluye el preprocesamiento de la imagen mediante filtros Gaussianos y de mediana para reducir el ruido. La metodología introduce una lógica de respaldo diseñada específicamente para manejar eventos de fusión pulmonar, asegurando que los contornos se delimiten correctamente incluso cuando los límites anatómicos se pierden. Además, se aplica un suavizado de contornos mediante un filtro Savitzky-Golay para mejorar la detección de características geométricas esenciales.

2. **Extracción de biomarcadores cuantitativos:** A partir de las máscaras segmentadas, se extrae un vector de características compuesto por 14 métricas diseñadas para capturar la integridad del sistema pulmonar de forma sistémica. Estas incluyen:
 - 2.1. **Biomarcadores morfológicos:** Compacidad discreta, volumen y tortuosidad, extraídos tanto del árbol vascular como del árbol de las vías respiratorias para capturar alteraciones estructurales a gran escala.
 - 2.2. **Estadísticas de daño parenquimatoso:** Promedio, desviación estándar, curtosis y asimetría de un índice de lesión (ROI) que cuantifica la extensión de las anomalías de alta densidad.
 - 2.3. **Características de textura:** Contraste, correlación, energía y homogeneidad derivadas de la Matriz de Co-ocurrencia de Niveles de Gris para caracterizar la heterogeneidad del tejido.
3. **Modelado y clasificación supervisada:** Se realiza una comparación exhaustiva de seis algoritmos de ML: Regresión Logística (RL), K-Vecinos más Cercanos (KNN), SVM, Bosque Aleatorio (Random Forest), Redes Neuronales de tipo Perceptrón Multicapa (MLP) y XGBoost. Estos modelos se someten a una optimización de hiperparámetros mediante búsqueda en cuadrícula (Grid Search) con validación cruzada estratificada para asegurar su robustez.
4. **Evaluación e interpretabilidad clínica:** El rendimiento se mide rigurosamente en conjuntos de prueba independientes utilizando métricas como exactitud, sensibilidad y el AUC-ROC. Crucialmente, la propuesta integra un análisis de interpretabilidad mediante importancia por permutación y valores SHAP, lo que permite identificar qué biomarcadores específicos (por ejemplo, el volumen del árbol vascular) determinaron la clasificación de un paciente, alineando el resultado del modelo con el razonamiento médico.

Esta solución integral busca demostrar que el uso de biomarcadores morfológicos no solo ofrece una alternativa eficiente en términos computacionales frente al DL, sino que proporciona la transparencia necesaria para su adopción segura en la práctica clínica real.

Para delimitar con claridad el alcance de esta investigación, es fundamental precisar que los datos de los biomarcadores morfológicos (volumen, tortuosidad y compacidad discreta) del árbol vascular y la vía aérea se adoptaron de la línea de base establecida por Torres Guzmán et al. (2024). Por consiguiente, las contribuciones originales y directas de este trabajo se centran en el desarrollo del algoritmo de segmentación primaria en MATLAB, el diseño e implementación de la lógica de respaldo para mitigar eventos de fusión pulmonar, la extracción autónoma de los descriptores estadísticos y texturales, y la arquitectura, optimización e interpretación con SHAP del pipeline completo de Machine Learning.

Capítulo 2. Marco teórico

La TC de alta resolución constituye el estándar técnico para la visualización tridimensional del parénquima y la vasculatura pulmonar (Raju et al., 2017; Campos & Diaz, 2018). El análisis cuantitativo moderno se fundamenta en la segmentación anatómica, un proceso que permite la transición de la inspección visual hacia la extracción de biomarcadores morfológicos (Bodduluri et al., 2021; Pu et al., 2020). Estos indicadores transforman la densidad de los píxeles en parámetros con significado físico, permitiendo una caracterización sistémica de la arquitectura vascular y aérea que es crucial para evaluar el impacto de patologías infecciosas (Torres Guzmán et al., 2024).

La integración de estos biomarcadores en modelos de inteligencia artificial define el paradigma de clasificación propuesto en esta investigación. Frente a la naturaleza de "caja negra" y la alta demanda computacional de los modelos de DL convencional (Lee et al., 2023), el ML basado en características diseñadas ofrece una alternativa eficiente y transparente (Brocki & Chung, 2023). Este paradigma permite que las decisiones del modelo puedan rastrearse hasta rasgos específicos, cumpliendo con los principios de la XAI y fortaleciendo la confianza clínica necesaria para la adopción segura de estas tecnologías en la práctica médica real (Ahmad et al., 2025).

2.1. Segmentación y procesamiento de imágenes médicas

La segmentación de imágenes médicas se define como el proceso de delimitación precisa de estructuras anatómicas de interés, tales como los pulmones y las lesiones parenquimatosas. Este paso es considerado el pilar fundamental de los sistemas de diagnóstico asistido por computadora, ya que permite la extracción de biomarcadores cuantitativos esenciales para el diagnóstico, la estadificación y el seguimiento clínico de diversas patologías (Yan et al., 2021). A pesar de su importancia, la interpretación manual de los extensos volúmenes de datos generados por la TC sigue siendo un proceso lento y susceptible a una alta variabilidad interobservador (Poola et al., 2023). La necesidad de automatizar esta tarea se volvió crítica durante la pandemia de

COVID-19, donde fue imperativo cuantificar el daño pulmonar de forma rápida y objetiva para optimizar la gestión de recursos hospitalarios (Akbari et al., 2021; Yan et al., 2021; Zeng et al., 2024).

2.1.1. Métodos geométricos y tradicionales

Dentro de las técnicas clásicas de procesamiento de imágenes, se han desarrollado enfoques geométricos avanzados para enfrentar las irregularidades de los contornos pulmonares. El algoritmo de Marcha Adaptativa de Bordes (Adaptive Border Marching, ABM) es una técnica destacada que recorre el contorno pulmonar y genera "puentes" matemáticos sobre las concavidades que pueden ser causadas por la presencia de nódulos yuxtapleurales (Pu et al., 2008). Su principal innovación es la implementación de un paso de marcha que se adapta localmente a la morfología, lo que previene la sobresegmentación en regiones cóncavas naturales como el mediastino (Pu et al., 2008).

Complementariamente, existen métodos fundamentados en el análisis de la curvatura local. Estos algoritmos se encargan de identificar puntos de alta inflexión en el contorno, conocidos como picos y valles de curvatura, los cuales suelen señalar la ubicación de vasos sanguíneos o lesiones periféricas (Singadkar et al., 2021). Una vez que se detectan estos puntos dominantes, se conectan mediante algoritmos de marcha que refinan el borde de forma iterativa hasta alcanzar la convergencia del contorno final (Singadkar et al., 2021). No obstante, la eficacia de estos métodos geométricos depende intrínsecamente de que la extracción inicial del contorno esté libre de ruido y artefactos (Singadkar et al., 2021).

2.1.2. Aprendizaje profundo en segmentación

En la última década, la arquitectura de red neuronal U-Net se ha consolidado como el estándar de facto para la segmentación de imágenes biomédicas debido a su excepcional capacidad de localización (Akbari et al., 2021; Turk & Kılıçaslan, 2025). Su estructura consiste en un diseño de codificador-decodificador simétrico con conexiones

saltadas (skip connections), lo que le permite integrar características semánticas de alto nivel con información espacial de alta resolución para lograr una segmentación precisa a nivel de píxel (Yan et al., 2021).

A pesar de sus ventajas, los modelos de DL enfrentan limitaciones significativas en el entorno clínico. Requieren conjuntos de datos masivos y meticulosamente anotados por expertos para su entrenamiento (Sarsembayeva et al., 2025). Además, su rendimiento tiende a degradarse ante alteraciones patológicas severas que causan la pérdida de los límites anatómicos convencionales, un fenómeno técnico conocido como fusión pulmonar (Li, 2023; Yan et al., 2021).

2.1.3. Retos técnicos en patologías severas

La segmentación automatizada en el contexto de patologías pulmonares graves, como la COVID-19 en estado crítico, representa un desafío técnico mayor, dado que la fiabilidad de los algoritmos suele disminuir a medida que aumenta la complejidad del daño tisular (Mansoor et al., 2014). Esto crea una paradoja diagnóstica donde las herramientas son más propensas a fallar precisamente cuando el análisis cuantitativo es más urgente (Mansoor et al., 2014).

Para contrarrestar estas deficiencias, se requiere un preprocesamiento robusto que incluya normalización de intensidad y filtrado de ruido mediante filtros de mediana y Gaussianos (Sarsembayeva et al., 2025). Asimismo, el empleo de técnicas de suavizado como el filtro Savitzky-Golay aplicado a las coordenadas de los contornos permite preservar las características geométricas fundamentales, facilitando la corrección de bordes y mejorando la precisión global de la máscara final en casos de alta complejidad.

2.2. Biomarcadores

La caracterización del sistema respiratorio mediante biomarcadores cuantitativos representa una transición fundamental de la observación cualitativa hacia la medicina de precisión. En particular, los biomarcadores morfológicos derivados de los árboles

vascular y aéreo pulmonar constituyen una fuente de características interpretables que permiten capturar alteraciones estructurales profundas en la arquitectura del órgano. A diferencia de los enfoques convencionales que se limitan al análisis de texturas locales, estos biomarcadores evalúan el impacto sistémico de la enfermedad en toda la estructura pulmonar, lo cual resulta crítico para el estudio de patologías como la COVID-19, la EPOC y la hipertensión pulmonar (Bodduluri et al., 2021; Pu et al., 2020). Estas afecciones provocan remodelaciones estructurales significativas, tales como la pérdida de vasos pequeños (pruning) y cambios geométricos en las vías respiratorias, los cuales pueden ser cuantificados objetivamente (Bodduluri et al., 2021; Pu et al., 2020).

2.2.1. Biomarcadores morfológicos

Dentro de esta investigación, se analizan tres biomarcadores morfológicos propuestos: la compacidad discreta, el volumen y la tortuosidad (Torres Guzmán et al., 2024). Estos parámetros son extraídos tanto para el árbol vascular como para el árbol de las vías respiratorias mediante el procesamiento tridimensional de las segmentaciones:

- **Volumen:** Representa la cuantificación tridimensional del espacio ocupado por las estructuras anatómicas segmentadas. Este biomarcador permite identificar alteraciones sistémicas en la arquitectura pulmonar (Torres Guzmán et al., 2024).

$$V = \sum_{i=1}^N d_{voxel} * n$$

$V = volumen$; $n = número de voxels$; $d_{voxel} = dimensiones del voxel$

- **Tortuosidad:** Evalúa el grado de sinuosidad o curvatura de los conductos aéreos y vasos sanguíneos, siendo un indicador altamente sensible a la remodelación vascular provocada por el daño crónico o infeccioso (Torres Guzmán et al., 2024).

$$TR = \frac{A - A_{min}}{A_{max} - A_{min}}$$

TR = tortuosidad; A = área de la superficie; A_{min} = área mínima de la superficie;
 A_{max} = área máxima de la superficie

- **Compacidad Discreta:** Evalúa la eficiencia del empaquetamiento de la estructura al relacionar su volumen con el área superficial; cambios en este índice sugieren alteraciones en la complejidad geométrica del sistema pulmonar debido a la enfermedad (Torres Guzmán et al., 2024).

$$DC = \frac{n - \frac{A}{6}}{n - (\sqrt[3]{n})^2}$$

DC = compacidad discreta; A = área de la superficie; n = número de voxels

Para la conformación del vector de características, se utilizaron los datos de volumen, tortuosidad y compacidad discreta derivados del trabajo de Torres Guzmán et al. (2024), los cuales fueron generados mediante el entorno 3D Slicer.

2.2.2. Cuantificación del daño parenquimatoso

Además de la morfología, es imperativo cuantificar la extensión de las lesiones en el parénquima pulmonar. Para este fin, se implementa un índice de lesión, el cual se calcula mediante la segmentación de áreas con anomalías de alta densidad, dentro de la región pulmonar (Yan et al., 2021). Matemáticamente, este índice representa la proporción del área de lesión respecto al área total del pulmón en cada corte axial (Albataineh et al., 2024).

A partir de este índice, se derivan descriptores estadísticos para resumir la distribución del daño a lo largo de todo el volumen del paciente:

- **Promedio y Desviación Estándar:** Proporcionan una medida de la severidad media y la dispersión del daño tisular.

$$Radio\ de\ infección = \frac{Área\ de\ lesión}{Área\ del\ pulmón}$$

$$Promedio = \frac{Radio\ de\ infección}{Número\ de\ cortes\ de\ la\ TC}$$

$$\text{Desviación estándar} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 *$$

- **Curtosis y Asimetría:** Describen la forma de la distribución del daño, permitiendo identificar si la afectación es homogénea o si existen picos de daño localizados en regiones específicas (Yan et al., 2021).

$$\text{Curtosis} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2\right)^2} *$$

$$\text{Asimetría} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}\right)^3} *$$

*N = número total de cortes procesados ; x_i = valor del radio de lesión en el corte i;

$\bar{x}$ = promedio del radio de lesión

2.2.3. Análisis de textura del parénquima

El análisis de textura constituye una herramienta esencial en el procesamiento de imágenes médicas, ya que permite cuantificar la heterogeneidad del tejido pulmonar mediante el estudio de la relación espacial de las intensidades de los píxeles. A diferencia de los biomarcadores morfológicos que evalúan la estructura macroscópica, el análisis textural se enfoca en las propiedades locales del parénquima, permitiendo identificar patrones que no son fácilmente detectables mediante la inspección visual. Esta técnica es fundamental para caracterizar el estado del tejido y distinguir entre parénquima sano y diversas manifestaciones patológicas (Guhan et al., 2022).

La metodología principal para este análisis se fundamenta en la GLCM, un método estadístico de segundo orden que describe la frecuencia con la que diferentes combinaciones de niveles de gris ocurren de forma adyacente en una imagen. Para

caracterizar la arquitectura del parénquima, se extraen cuatro descriptores cuantitativos promediados a lo largo de todos los cortes axiales de la TC:

- **Contraste:** Mide las variaciones locales de intensidad entre píxeles vecinos. Valores elevados de contraste indican una textura con cambios bruscos de intensidad, característicos de zonas donde coexisten tejido sano y lesiones de alta densidad.

$$Contraste = \sum_{i,j=0}^{N_g-1} p(i,j)(i-j)^2*$$

- **Correlación:** Cuantifica la dependencia lineal de los niveles de gris en la imagen. Una correlación alta sugiere que el tejido presenta patrones repetitivos o estructuras organizadas, mientras que una baja correlación indica una textura más aleatoria o degradada por la enfermedad.

$$Correlación = \sum_{i,j=0}^{N_g-1} p(i,j) \left[\frac{(i-u_i)(j-u_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \right] *$$

- **Energía (Uniformidad):** Representa la suma de los cuadrados de los elementos de la GLCM. En tejidos uniformes y homogéneos, el valor de la energía es mayor, lo que permite diferenciar parénquima estable de regiones con infiltrados heterogéneos.

$$Energía = \sum_{i,j=0}^{N_g-1} p(i,j)^2*$$

- **Homogeneidad:** Indica qué tan cerca se encuentra la distribución de los elementos de la GLCM respecto a la diagonal de la matriz. Esta métrica es inversamente proporcional al contraste y ayuda a identificar la suavidad del tejido pulmonar.

$$Homogeneidad = \sum_{i,j=0}^{N_g-1} \frac{p(i,j)}{1+(i-j)^2}^*$$

**i = nivel de gris en la fila de la matriz ; j = nivel de gris en la fila de la matriz;*

p(i, j) = probabilidad conjunta de encontrar un par de píxeles con intensidades i y j;

N_g = número total de niveles de gris en la imagen ; u = promedio de las intensidades;

σ = desviación estándar de las intensidades;

La combinación de estos parámetros texturales con biomarcadores morfológicos y estadísticos permite conformar un vector de características altamente discriminatorio. Este enfoque integrado busca que los modelos de ML alcancen niveles de exactitud diagnóstica altas con el objetivo de ofrecer una alternativa transparente y computacionalmente eficiente a los modelos de DL convencionales.

2.3. Algoritmos de machine learning

La literatura científica reciente documenta una transición significativa desde los métodos tradicionales de ML hacia modelos de DL para la clasificación de imágenes de TC. El DL, particularmente a través de CNN, ha demostrado un rendimiento excepcional con precisiones y sensibilidades que superan frecuentemente el 90% en la detección de patologías como la COVID-19 o el cáncer de pulmón (Guhan et al., 2022; Harmon et al., 2020; Sarv Ahrabi et al., 2021). No obstante, estos modelos enfrentan desafíos críticos relacionados con su alta demanda de recursos computacionales para el entrenamiento y su naturaleza de "caja negra" (Bodduluri et al., 2021; Lee et al., 2023; Zhang et al., 2023). Esta falta de transparencia limita su aceptación en entornos clínicos, ya que dificulta la interpretación de sus decisiones y la confianza por parte de los especialistas (Brocki & Chung, 2023; Dubois et al., 2024).

Como respuesta a estas limitaciones, ha resurgido el interés por el ML interpretable fundamentado en vectores de características diseñadas o biomarcadores derivados del conocimiento del dominio clínico (Brocki & Chung, 2023). Este paradigma utiliza algoritmos supervisados —como las SVM o KNN— que han reportado exactitudes

competitivas entre el 97% y el 99% en diversas tareas de clasificación (Guhan et al., 2022; Albataineh et al., 2024). A diferencia del DL, este enfoque permite que las decisiones del sistema se rastreen hasta rasgos físicos y estructurales específicos, alineándose con el razonamiento médico y permitiendo un entrenamiento eficiente incluso con tamaños de muestra moderados (Li et al., 2023; Pu et al., 2020).

Finalmente, la integración de técnicas de XAI es fundamental para cerrar la brecha entre la investigación técnica y la práctica médica (Lee et al., 2023). Herramientas como la segmentación previa de las áreas anormales del pulmón permiten a los clínicos verificar visualmente las regiones de interés que determinaron el diagnóstico, aumentando la transparencia y la confiabilidad del sistema (Harmon et al., 2020). La validación robusta en conjuntos de datos diversos y la capacidad de explicar los resultados son requisitos indispensables para superar los obstáculos de interpretabilidad y asegurar una integración efectiva y segura de estas tecnologías en la práctica clínica real (Ahmad et al., 2025; Lee et al., 2023).

2.3.1. Regresión logística

La RL es un modelo estadístico fundamental para la clasificación binaria que predice la probabilidad de pertenencia a una clase mediante una función sigmoide. Es valorada en el ámbito médico por su alta interpretabilidad, permitiendo asociar directamente los coeficientes del modelo con la relevancia de biomarcadores específicos (Brocki & Chung, 2023). En tareas de clasificación de nódulos y patologías pulmonares, variantes como la RL con regularización Lasso han demostrado ser eficaces para reducir la dimensionalidad y mejorar la generalización del modelo (Brocki & Chung, 2023).

2.3.2. K-vecinos más cercanos

El algoritmo KNN es un método de aprendizaje no paramétrico que clasifica cada instancia basándose en la mayoría de votos de sus vecinos más cercanos en un espacio de características multidimensional (Albataineh et al., 2024). Este enfoque es robusto para identificar patrones de similitud en biomarcadores de textura y densidad sin asumir una distribución lineal de los datos. Su precisión depende de la métrica de distancia y la

optimización del hiperparámetro K, logrando en estudios clínicos exactitudes superiores al 98% en la detección de severidad de infecciones virales (Albataineh et al., 2024).

2.3.3. Máquina de vectores de soporte

Las SVM son clasificadores que buscan el hiperplano óptimo que maximiza el margen entre clases en un espacio de alta dimensión (Albataineh et al., 2024). Mediante el uso de funciones de núcleo (kernels), este algoritmo puede capturar relaciones no lineales complejas entre biomarcadores radiómicos y morfológicos (Brocki & Chung, 2023). Es uno de los modelos de mejor rendimiento reportados, alcanzando precisiones de hasta el 99.12% en la clasificación de pacientes sanos frente a patológicos (Albataineh et al., 2024; Guhan et al., 2022).

2.3.4. Bosque aleatorio (Random Forest)

El Random Forest es un algoritmo de aprendizaje por ensamble que combina múltiples árboles de decisión para mejorar la estabilidad y precisión de la clasificación (Brocki & Chung, 2023). Su arquitectura permite evaluar la "importancia de las características", lo que facilita identificar qué biomarcadores, como el volumen vascular o la tortuosidad, son más determinantes para el diagnóstico. Es una herramienta resiliente al sobreajuste y ha sido validada para la clasificación de malignidad en nódulos pulmonares con altos índices de recall (Brocki & Chung, 2023).

2.3.5. Extreme gradient boosting (XGBoost)

XGBoost es una implementación avanzada del algoritmo de aumento de gradiente (gradient boosting) diseñada para ser altamente eficiente y flexible. A diferencia de los bosques aleatorios tradicionales, XGBoost entrena árboles de forma secuencial donde cada nuevo árbol corrige los errores de los anteriores, utilizando una función de pérdida regularizada para prevenir el sobreajuste. Aunque se basa en estructuras de árboles de decisión (Albataineh et al., 2024), su capacidad para manejar datos dispersos y su velocidad computacional lo posicionan como un modelo de vanguardia para procesar vectores de características heterogéneos en medicina.

2.3.6. Redes neuronales artificiales: Perceptrón Multicapa (MLP)

El Perceptrón Multicapa es una arquitectura de red neuronal de alimentación directa que aprende relaciones no lineales complejas mediante capas de neuronas interconectadas. A diferencia de las arquitecturas de DL que procesan imágenes completas (Sarv Ahrabi et al., 2021), el MLP se aplica sobre el vector de biomarcadores ya extraídos, lo que permite aprovechar la potencia de aprendizaje de las redes neuronales con una carga computacional significativamente menor. Esto facilita la optimización de hiperparámetros para maximizar el F1-score en diagnósticos complejos (Sarv Ahrabi et al., 2021).

En conclusión, el marco teórico presentado subraya que, si bien el DL ha alcanzado niveles de rendimiento excepcionales con precisiones superiores al 90% en la detección de patologías pulmonares (Guhan et al., 2022; Harmon et al., 2020), su adopción clínica masiva se ve frenada por la opacidad de sus procesos y la alta demanda de infraestructura computacional (Lee et al., 2023; Bodduluri et al., 2021). Frente a este escenario, la utilización de un vector de características diseñado —que integra biomarcadores morfológicos, estadísticos y texturales— surge como una solución robusta que prioriza la transparencia diagnóstica y la eficiencia (Brocki & Chung, 2023). Este enfoque no solo optimiza el uso de recursos, sino que permite que cada decisión del modelo esté fundamentada en rasgos físicos y estructurales alineados con el conocimiento médico, garantizando la interpretabilidad inherente necesaria para su validación en entornos clínicos reales (Li et al., 2023; Pu et al., 2020).

Por lo tanto, la base conceptual establecida permite justificar la transición de la teoría a la implementación técnica. La integración de técnicas de XAI junto con algoritmos supervisados de alto rendimiento como SVM, Random Forest y XGBoost, asegura un equilibrio necesario entre la exactitud diagnóstica y la confianza clínica (Lee et al., 2023; Ahmad et al., 2025). Habiendo definido la relevancia de una segmentación robusta, la validez biológica de los biomarcadores seleccionados y la capacidad operativa de los modelos de ML (Torres Guzmán et al., 2024; Zhang et al., 2023), se sientan las bases para detallar el flujo metodológico en el siguiente capítulo (Dubois et al., 2024). En consecuencia, el Capítulo 3 describirá los materiales, las cohortes de pacientes y los

procedimientos experimentales utilizados para transformar estos fundamentos teóricos en un sistema funcional de apoyo al diagnóstico neumológico (Soyyigit, 2020; Clark et al., 2013).

Capítulo 3. Metodología

El presente capítulo detalla el flujo de trabajo experimental diseñado para la segmentación y clasificación automatizada de patologías pulmonares. La metodología se estructura en cinco etapas principales: 1) adquisición y depuración de datos, 2) preprocesamiento de imágenes, 3) segmentación de estructuras anatómicas mediante una lógica híbrida, 4) extracción de biomarcadores morfológicos y texturales, y 5) entrenamiento y validación de los modelos de ML. Este enfoque busca garantizar la reproducibilidad de los resultados y la robustez del sistema ante casos de alta complejidad clínica.

3.1. Materiales y cohortes de estudio

La validez de un sistema de diagnóstico asistido por computadora depende fundamentalmente de la calidad y diversidad de los datos utilizados para su entrenamiento y validación (Roberts et al., 2021). Para esta investigación, se conformó una base de datos heterogénea proveniente de repositorios internacionales de acceso abierto, lo que permite asegurar la representatividad de diferentes protocolos de adquisición y diversas manifestaciones patológicas.

3.1.1. Origen de los datos

Se utilizaron dos fuentes principales de datos para construir las cohortes de pacientes:

1. **Conjunto de Datos COVID-19 (Harvard Dataverse):** Se integraron tomografías de tórax de pacientes con diagnóstico confirmado de COVID-19 en diversos estadios de severidad, permitiendo el análisis de lesiones parenquimatosas y eventos de fusión pulmonar (Soyyigit, 2020).
2. **Conjunto de Datos Control:** El grupo control (n=78) fue recopilado de forma retrospectiva a partir de dos colecciones indexadas en el repositorio público *The Cancer Imaging Archive (TCIA)*. Esta cohorte de contraste no representa sujetos

completamente sanos o libres de patología, sino pacientes evaluados por anomalías torácicas independientes de la infección viral por SARS-CoV-2. El desglose exacto de las fuentes del grupo de control se detalla en la Tabla 3.1. (Clark et al., 2013).

Tabla 3.1. Composición del cohorte control extraído de TCIA.

Colección de Origen	Número de Estudios	Descripción del Estudio	DOI de la Colección
CT Lymph Nodes	61	Evaluación de ganglios mediastinales	10.7937/K9/TCIA.2015.AQIIDCNM
LIDC-IDRI	17	Tamizaje pulmonar y caracterización de nódulos	10.7937/K9/TCIA.2015.LO9QL9SX

3.1.2. Criterios de inclusión y exclusión

Para garantizar la integridad del análisis, los estudios de TC debieron cumplir con los siguientes criterios: ser imágenes de tórax en formato DICOM o NIfTI, poseer una resolución espacial suficiente para la reconstrucción 3D y contar con la etiqueta diagnóstica correspondiente. Se excluyeron aquellos estudios con artefactos de movimiento severos o truncamiento de la anatomía pulmonar que pudieran impedir una segmentación completa (Roberts et al., 2021; Ahmad et al., 2025).

3.2. Técnicas de procesamiento

La metodología de procesamiento de imágenes, desarrollada íntegramente en el entorno de MATLAB, constituye una etapa crítica para garantizar que la información contenida en los vóxeles de la TC sea interpretada de manera consistente por los algoritmos de clasificación. Debido a la alta variabilidad técnica entre protocolos de adquisición, el flujo de trabajo estandariza los datos para reducir el ruido y mejorar el contraste en las regiones de interés, estructurándose en la secuencia lógica presentada en la Figura 3.1.



Figura 3.1.- Diagrama de bloques del procesamiento de imágenes en MATLAB

3.2.1. Adquisición y selección interactiva de la región de interés

El flujo de procesamiento comienza con la carga del volumen 3D en formato NIfTI mediante la función **niftiread(...)**. Para optimizar la eficiencia computacional y eliminar estructuras anatómicas irrelevantes (como tejidos blandos periféricos y artefactos del soporte del paciente), se implementó una selección interactiva de la ROI.

Este proceso utiliza una Proyección de Máxima Intensidad (MIP) generada con la función **max(...)**, la cual permite visualizar de forma resumida la extensión de los pulmones en un plano bidimensional. Mediante la herramienta **imrect(...)**, el usuario define manualmente el área de recorte, asegurando que el análisis se limite estrictamente a la cavidad torácica. La función **wait(...)** pausa el script hasta que la selección es confirmada, permitiendo un encuadre preciso de ambos pulmones independientemente del tamaño del paciente.

3.2.2. Mejora de contraste y normalización

Una vez extraída la ROI, las imágenes se procesan para estandarizar el rango dinámico. Se utiliza la función **imadjust(...)** para realizar un mapeo de las intensidades originales a un rango que resalte el parénquima pulmonar. Posteriormente, mediante **mat2gray(...)**, se normalizan los valores de intensidad a un rango de [0,1]. Esta normalización es fundamental para la estabilidad de los algoritmos de segmentación y para que las GLCM sean comparables entre diferentes sujetos de estudio.

3.2.3. Reducción de ruido mediante filtrado espacial

Para mitigar la presencia de artefactos de adquisición y ruido cuántico, se implementó un filtrado secuencial:

- **Filtro de Mediana:** Se aplica mediante `medfilt2(...)` con una ventana de [2,2] para eliminar el ruido impulsivo ("sal y pimienta") preservando los bordes de las lesiones y vasos sanguíneos.
- **Filtro Gaussiano:** Se ejecuta con `imgaussfilt(...)` ($\sigma = 0.2$) para homogeneizar la textura interna del parénquima, facilitando la identificación de regiones de baja densidad de forma continua.

3.2.4. Umbralización dual y limitaciones del método de Otsu

La segmentación inicial de los pulmones se fundamenta en una lógica de umbralización. En la literatura de procesamiento de imágenes, el método de Otsu es el estándar para encontrar umbrales óptimos de forma automática basándose en la maximización de la varianza entre clases de un histograma bimodal.

Sin embargo, en esta investigación se determinó que el método de Otsu no es adecuado para el análisis de pulmones con patología severa (específicamente COVID-19 y neumonías graves). La razón técnica es que el método de Otsu asume que el histograma es bimodal (dos picos claros: uno para el aire y otro para el tejido). En pacientes enfermos, la presencia de opacidades genera una gran cantidad de píxeles con densidades intermedias, lo que provoca que el histograma se vuelva unimodal o multimodal. En estos casos, el algoritmo de Otsu suele calcular un umbral erróneo que omite gran parte del pulmón enfermo o lo fusiona con estructuras externas.

Para superar este desafío, se implementó una umbralización dual fija, cuya lógica matemática se define como:

$$OCT = (IG < 0.4) \cap (IG > 0)$$

En donde IG es la imagen normalizada, a la cual se le aplicaron filtros de suavizado para mitigar el ruido. A partir de esta, se obtuvo una máscara binaria resultante de una doble umbralización, siendo la variable OCT. Esta máscara permite aislar los vóxeles con densidades radiológicas compatibles con el parénquima pulmonar afectado, sirviendo como la base para los procesos posteriores de refinamiento morfológico y extracción de biomarcadores

La selección de estos límites específicos se determinó de manera estrictamente experimental mediante un análisis empírico iterativo. Debido a que las tomografías de pacientes con afecciones pulmonares severas no presentan un histograma de intensidades bimodal estable sino distribuciones de carácter unimodal o multimodal con transiciones continuas de niveles de gris provocadas por los patrones difusos, la aplicación de criterios analíticos automatizados convencionales resultaba inviable. Por lo tanto, el establecimiento de estos umbrales fijos fue el resultado de pruebas experimentales orientadas a maximizar la inclusión del tejido pulmonar real dañado y mitigar la introducción de falsos positivos en las regiones colindantes de la cavidad torácica.

3.2.5. Suavizado de bordes con filtro Savitzky-Golay

Tras la obtención de los contornos mediante **bwboundaries(...)**, se observó que la discretización de los píxeles generaba bordes irregulares que podrían afectar la precisión de las métricas. Para corregir esto, se implementó el filtro de Savitzky-Golay mediante la función **sgolayfilt(...)**.

Este filtro aplica un ajuste polinomial de tercer orden sobre una ventana de 33 puntos en las coordenadas (x,y) del contorno. A diferencia de los filtros de promedio móvil, el método de Savitzky-Golay logra suavizar el borde sin suavizar excesivamente las características geométricas reales del pulmón, permitiendo que el algoritmo iterativo de corrección de bordes trabaje sobre una base anatómica más fiel a la realidad clínica.

3.2.6. Algoritmo de corrección de bordes para daño pulmonar severo

La segmentación precisa en pacientes con daño pulmonar severo representa uno de los mayores desafíos técnicos en el análisis de imágenes médicas, ya que las lesiones a menudo oscurecen o eliminan los límites anatómicos naturales entre el parénquima y los tejidos circundantes. Para superar esta limitación, tras el suavizado inicial de los contornos, se implementó un algoritmo de corrección de bordes inspirado en los modelos geométricos de Pu et al. (2008) y Singadkar et al. (2021) .

Este algoritmo opera de manera iterativa, ajustando dinámicamente el contorno del pulmón mediante el análisis de su curvatura local:

- **Detección de puntos dominantes:** Utilizando la señal de curvatura derivada de las coordenadas suavizadas, el algoritmo emplea la función **findpeaks(...)** para localizar picos y valles de curvatura que corresponden a puntos de inflexión críticos en la morfología pulmonar.
- **Simplificación geométrica iterativa:** Basándose en una lógica de simplificación de polígonos, el sistema conecta estos puntos dominantes y evalúa la colinealidad mediante un parámetro de umbral ($\lambda = 0.9$) para refinar el borde.
- **Criterio de convergencia:** El refinamiento se ejecuta dentro de un bucle **while(...)** condicionado por una tolerancia de perímetro y un límite máximo de iteraciones (`config.iterativo.maxIteraciones = 100`). La corrección se detiene cuando la diferencia de perímetro entre la nueva máscara generada por **poly2mask(...)** y la versión anterior es de 1 mm, asegurando que el contorno final se ajuste con precisión a la periferia pulmonar real.

Es importante destacar que este algoritmo de corrección se aplica de forma independiente a cada pulmón tras la separación de contornos. En aquellos casos críticos donde la patología provoca una fusión pulmonar (pérdida del límite anatómico entre ambos pulmones), el sistema activa automáticamente la lógica de respaldo. Esta lógica detecta la presencia de un único objeto fusionado mediante **bwboundaries(...)** y opta

por omitir el refinamiento iterativo para ese corte específico, utilizando la máscara inicial fusionada para evitar errores topológicos que los algoritmos de corrección suelen cometer en escenarios de fusión total.

3.3. Validación del algoritmo

Para verificar la efectividad y precisión del método de segmentación propuesto, especialmente en su capacidad para incluir lesiones en pacientes con cuadros graves, se llevó a cabo un proceso de validación técnica. Este procedimiento permitió cuantificar la fiabilidad de las contribuciones metodológicas, como la lógica de respaldo para la fusión pulmonar y el suavizado de contornos.

3.3.1. Procedimiento con phantom 3D y referencia manual

La validación cuantitativa se fundamentó en el uso de un phantom 3D sintético en el cual tiene dos cilindros con diferentes valores en la escala de grises, diseñados para simular tejidos con distintas densidades. Como estándar de referencia, se realizó una segmentación manual del phantom utilizando el software 3D Slicer. Posteriormente, se aplicó el algoritmo desarrollado a las imágenes de TC del phantom para comparar sus resultados con dicha referencia manual.

Los detalles metodológicos complementarios de esta prueba, así como el desglose de los resultados y métricas de desempeño volumétrico obtenidos, se presentarán de forma detallada en el Capítulo 4 de este trabajo.

Es imperativo señalar una limitación metodológica fundamental dentro de este proceso de validación: debido a la ausencia de un médico radiólogo especialista disponible para realizar la segmentación manual de referencia sobre la cohorte clínica real, no fue posible generar un estándar de comparación cuantitativa para las tomografías de los pacientes. En consecuencia, el alcance de la validación matemática mediante métricas de superposición quedó acotado exclusivamente al entorno controlado del fantoma sintético 3D.

No obstante, la justificación técnica de la lógica de respaldo implementada en escenarios clínicos reales se fundamenta en su comportamiento cualitativo de protección topográfica. En cortes tomográficos donde la gravedad de la patología borra los límites anatómicos o la estructura de la parénquima tenga límites difusos, intentar aplicar un refinamiento geométrico basado en la unión iterativa de puntos periféricos induciría a errores severos y deformaciones artificiales de la estructura. Al detectar la fusión, la lógica de respaldo interviene deteniendo el algoritmo de corrección de bordes y optando por conservar la segmentación primaria del corte. Esta decisión de diseño evita que se falseen las mallas geométricas, asegurando que el parénquima pulmonar mantenga su integridad matemática antes de pasar a la etapa de extracción de características.

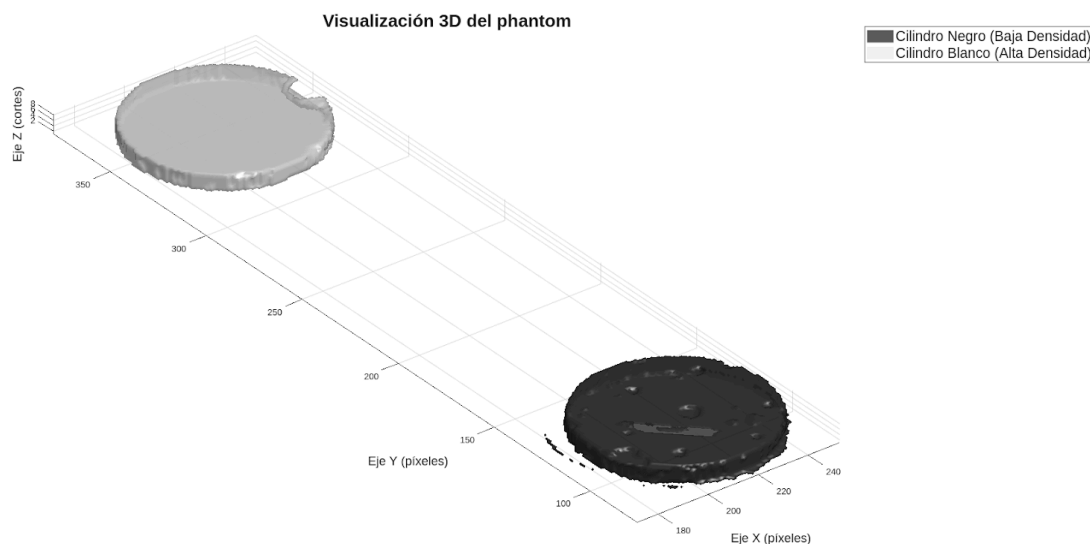


Figura 3.2.- Gráfica donde se visualiza la segmentación manual de ambos cilindros del phantom

3.3.2. Métricas de desempeño volumétrico

La precisión del algoritmo fue evaluada mediante un conjunto de métricas estándar que permiten analizar la concordancia y los errores de segmentación:

- **Fracción de superposición volumétrica (VOF):** En este trabajo se utilizó el Coeficiente de Dice para medir el grado de coincidencia entre los volúmenes.

- **Índice de jaccard (IoU):** Cuantifica la similitud mediante la intersección sobre la unión de los volúmenes automático y manual.
- **Volumetric overlap error ratio (VOER):** Mide la disparidad volumétrica general entre ambos métodos.
- **Tasa de sobresegmentación (OSR):** Representa los falsos positivos, cuantificando el volumen segmentado erróneamente fuera de la referencia.
- **Tasa de subsegmentación (USR):** Representa los falsos negativos, midiendo el volumen de referencia que el algoritmo no logró capturar.

3.4. Obtención de los biomarcadores

Posteriormente, se procedió a la obtención de los ocho biomarcadores detallados en los epígrafes 2.2.2 y 2.2.3 del Capítulo 2, los cuales comprenden la cuantificación de la lesión pulmonar y los descriptores referentes a la textura e intensidad de los cortes de TC de cada paciente.

3.4.1. Cuantificación y caracterización de lesiones (ROI)

La cuantificación del daño parenquimatoso se realiza mediante la aplicación secuencial de máscaras binarias en MATLAB, lo que permite aislar el tejido patológico del parénquima sano con alta precisión. Este proceso inicia utilizando la máscara del pulmón ya refinada como un filtro espacial sobre la imagen de intensidad y mediante una operación de enmascaramiento se eliminan todos los píxeles externos a los pulmones, asegurando que el análisis de textura y densidad se limite exclusivamente a la cavidad pulmonar y no se vea sesgado por tejidos extrapulmonares.

Para identificar la sección de daño, se aplica una técnica de umbralización dual basada en las intensidades características de las lesiones por COVID-19. En el script desarrollado, se utiliza un umbral de intensidad inferior de 0.3 y un límite superior de

0.59; esta ventana permite capturar tanto las opacidades como las zonas de consolidación más densas. La máscara resultante, identifica los píxeles candidatos a lesión, los cuales son posteriormente filtrados mediante criterios de área y excentricidad para descartar ruidos o estructuras lineales que no corresponden a infiltrados patológicos.

La medición del daño se fundamenta en un cálculo de proporción aritmética para cada corte axial procesado. Matemáticamente, la severidad en un corte específico se obtiene mediante la división del área dañada entre el área total del pulmón. Esta normalización es crítica para obtener un ratio de lesión independiente del tamaño anatómico del paciente o del nivel de profundidad del corte.

Finalmente, para consolidar un diagnóstico sistémico, el sistema recopila los ratios de todos los cortes individuales en un vector de resultados (V). A partir de este vector, se calculan cuatro variables estadísticas fundamentales que caracterizan la distribución del daño en todo el volumen pulmonar:

- **Promedio:** Se obtiene mediante la función **mean(V_filtrado)** y representa el índice global de severidad o carga lesional media del paciente.
- **Desviación Estándar:** Calculada con **std(V_filtrado)**, esta métrica cuantifica la dispersión del daño, indicando qué tan heterogénea es la distribución de la enfermedad a lo largo de los diferentes lóbulos pulmonares.
- **Asimetría:** Utilizando la función **skewness(V_filtrado)**, se mide el grado de asimetría de la distribución de las lesiones. Un valor de asimetría permite identificar si el daño se concentra mayoritariamente en ciertas secciones del volumen pulmonar (por ejemplo, predominancia basal).
- **Curtosis:** Se extrae mediante **kurtosis(V_filtrado)** para analizar la concentración del daño. Esta variable indica si el paciente presenta niveles de infección similares en todos los cortes (curtosis baja) o si existen picos de afectación muy severos en cortes específicos (curtosis alta).

La integración de estas cuatro métricas estadísticas permite transformar la información de cientos de imágenes en un perfil numérico compacto y altamente discriminatorio para los modelos de clasificación.

3.4.2. Análisis de textura mediante GLCM

Para complementar la información morfológica, se implementó un análisis de segundo orden basado en la GLCM. Este método permite cuantificar la heterogeneidad del parénquima pulmonar mediante el estudio de la relación espacial entre las intensidades de los píxeles, capturando patrones de textura que no son detectables mediante la simple inspección visual.

El procesamiento en MATLAB se realiza siguiendo un flujo de trabajo diseñado para estabilizar las métricas frente al ruido técnico de las tomografías:

- **Cuantización de intensidades:** La imagen original se reduce a 8 niveles de gris mediante la función **grayscale(...)**. Esta reducción es fundamental para disminuir la sensibilidad del algoritmo al ruido cuántico de la TC y asegurar que la matriz de co-ocurrencia no sea excesivamente dispersa, lo que permite obtener descriptores estadísticos más estables y representativos del tejido.
- **Enmascaramiento de precisión:** Para garantizar que los cálculos de textura se refieran estrictamente al tejido pulmonar, se utiliza la máscara Pulmon_corregido para asignar valores NaN (Not a Number) a todos los píxeles externos al parénquima. Al realizar este enmascaramiento antes de invocar a **graycomatrix(...)**, se evita que el fondo negro o las estructuras de la pared torácica sesguen las métricas de contraste y homogeneidad.
- **Configuración de desplazamientos (Offsets):** Se definieron cuatro desplazamientos angulares para el análisis: $[0 \ 1]$ (0°), $[-1 \ 1]$ (45°), $[-1 \ 0]$ (90°) y $[-1 \ -1]$ (135°). Esta configuración multidireccional asegura que el análisis sea invariante a la rotación, permitiendo capturar la arquitectura del parénquima independientemente de la orientación de los lóbulos en el corte axial.

A partir de la matriz generada, se extraen los cuatro biomarcadores de GLCM mediante la función **graycoprops(...)**, obteniendo el promedio del contraste, correlación, energía y homogeneidad de todos los cortes de TC de cada paciente.

3.4.3. Biomarcadores morfológicos estructurales

Esta sección de la investigación representa una continuación directa y extensión del trabajo desarrollado por Torres Guzmán et al. (2024), donde se estableció la relevancia clínica de los descriptores geométricos para evaluar las secuelas en pacientes diagnosticados con COVID-19. Debido a que este estudio se fundamenta en dicha base metodológica, los biomarcadores estructurales del árbol traqueobronquial y del sistema vascular pulmonar ya se encontraban calculados y validados, permitiendo su integración inmediata como parte del vector de 14 características esenciales para el entrenamiento de los modelos de inteligencia artificial.

Este enfoque multidimensional asegura que el sistema de clasificación no solo detecte la presencia de daño, sino que también comprenda las alteraciones estructurales profundas que afectan la función respiratoria del paciente.

3.5. Entrenamiento y validación de modelos de machine learning

El desarrollo de los modelos de clasificación se fundamenta en el uso de los biomarcadores cuantitativos extraídos previamente, los cuales sirven como entrada para sistemas de diagnóstico asistido por computadora. Esta fase es crucial para convertir los datos morfológicos y texturales en información accionable que apoye la toma de decisiones médicas.

Es fundamental precisar que el pipeline tecnológico desarrollado en esta investigación opera como un sistema híbrido que conecta de manera secuencial dos entornos de desarrollo computacionales distintos. Las etapas de preprocesamiento espacial, segmentación anatómica con corrección de bordes y la extracción inicial de los descriptores matemáticos se ejecutaron de forma centralizada en MATLAB. Una vez

cuantificados los rasgos de cada paciente, el vector de 14 características se exportó de forma automatizada hacia un archivo estructurado de valores separados por comas. Posteriormente, este archivo sirvió como canal de conexión al ser importado dentro de un entorno de Python a través de un Jupyter Notebook, donde se programaron la normalización de variables, la optimización por búsqueda en cuadrícula, el entrenamiento de los seis modelos supervisados y el análisis XAI con SHAP.

Para garantizar la reproducibilidad e integridad completa de este pipeline de software, se detalla a continuación la especificación técnica del entorno utilizado:

- **Procesamiento de Imágenes y Biomarcadores:** MATLAB (Versión R2024a]) haciendo uso especializado de los módulos *Image Processing Toolbox* y *Signal Processing Toolbox*.
- **Entorno Executable de Machine Learning:** Python (Versión 3.11]), ejecutado en la infraestructura interactiva de *Jupyter Notebook*.
- **Librerías de Modelado y Explicabilidad:** *Scikit-learn* (Versión 1.3.2]) para la manipulación estadística, partición y optimización de hiperparámetros; *XGBoost* (Versión 2.0.2]) para el algoritmo de aumento de gradiente; y *SHAP* (Versión 0.42.1]) para la extracción de valores de atribución.

3.5.1. Preparación y partición del conjunto de datos

El vector de características, compuesto por 14 variables (tortuosidad del árbol vascular, volumen del árbol vascular, compacidad discreta del árbol vascular, tortuosidad de las vías respiratorias, volumen de las vías respiratorias, compacidad de las vías respiratorias, promedio, desviación estándar, curtosis, asimetría, contraste, correlación, homogeneidad y energía), se somete a un proceso de normalización para asegurar que todas las métricas contribuyan equitativamente al aprendizaje del modelo.

Para garantizar la capacidad de generalización del sistema y evitar el sobreajuste (*overfitting*), los datos se dividen en dos conjuntos independientes:

- **Conjunto de entrenamiento (80%):** Utilizado para el ajuste de los pesos y parámetros de los algoritmos.
- **Conjunto de prueba (20%):** Reservado exclusivamente para la evaluación final, simulando el desempeño del modelo ante pacientes nuevos en un entorno clínico real.

3.5.2. Optimización de hiperparámetros mediante Grid Search

Para maximizar el rendimiento de cada algoritmo (SVM, Random Forest, XGBoost, KNN, Regresión Logística y MLP), se implementó una estrategia de Grid Search combinada con Validación Cruzada de k-pliegues (k-fold Cross-Validation). Este procedimiento permite explorar sistemáticamente diversas combinaciones de hiperparámetros, tales como:

- **SVM:** Selección del tipo de núcleo (lineal, RBF) y el parámetro de regularización C.
- **Random Forest / XGBoost:** Ajuste del número de árboles, la profundidad máxima y la tasa de aprendizaje.
- **KNN:** Optimización del número de vecinos (K) y la métrica de distancia.

La validación cruzada asegura que la selección de parámetros sea robusta y no dependa de una partición específica de los datos, mejorando la fiabilidad de los resultados consolidados.

3.5.3. Métricas de evaluación del desempeño

La eficacia de los modelos de ML se cuantifica mediante métricas estadísticas derivadas de la matriz de confusión, la cual permite desglosar el rendimiento del clasificador en cuatro categorías fundamentales: Verdaderos Positivos (TP), Verdaderos Negativos (TN), Falsos Positivos (FP) y Falsos Negativos (FN). La correcta identificación de estas

categorías es indispensable para garantizar que el sistema de diagnóstico asistido por computadora sea fiable en escenarios clínicos desafiantes.

- **Exactitud (*Accuracy*):** Representa la proporción global de predicciones correctas realizadas por el modelo respecto al total de casos evaluados. Se define matemáticamente como:

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Sensibilidad (*Recall*):** Mide la capacidad del algoritmo para identificar correctamente a los pacientes que presentan la patología. En medicina, esta métrica es crítica para minimizar los falsos negativos (FN), asegurando que ningún paciente enfermo sea omitido por el sistema. Su fórmula es:

$$Sensibilidad = \frac{TP}{TP + FN}$$

- **Especificidad:** Evalúa la capacidad del modelo para identificar correctamente a los sujetos sanos o sin daño significativo. Es fundamental para reducir la tasa de sobresegmentación o falsos positivos (FP), evitando intervenciones médicas innecesarias. Se calcula como:

$$Especificidad = \frac{TN}{TN + FP}$$

- **F1-score:** Se define como el promedio armónico entre la precisión y la sensibilidad. Esta métrica es especialmente robusta cuando existe un desequilibrio entre las clases de severidad, proporcionando una visión más equilibrada del rendimiento del clasificador que la exactitud simple.

$$F1 = 2 * \frac{Precisión * Sensibilidad}{Precisión + Sensibilidad}$$

$$Donde: Precisión = \frac{TP}{TP + FP}$$

- **AUC-ROC:** La curva ROC (*Receiver Operating Characteristic*) representa la tasa de verdaderos positivos frente a la tasa de falsos positivos a diversos umbrales de decisión. El valor AUC resultante indica la probabilidad de que el modelo clasifique una instancia positiva elegida al azar por encima de una negativa, siendo un indicador de la capacidad discriminatoria global del sistema.

La aplicación de este conjunto de métricas asegura que el algoritmo validado no sólo sea preciso en condiciones controladas, sino que posea la robustez necesaria para apoyar la toma de decisiones clínicas en enfermedades complejas como la COVID-19.

3.5.4. Análisis de importancia de características e interpretabilidad

A diferencia de los modelos de DL que actúan como "cajas negras", esta metodología prioriza la interpretabilidad. Se utilizaron técnicas de XAI, como la importancia por permutación, para determinar qué biomarcadores específicos (por ejemplo, la tortuosidad vascular o el promedio de ROI) tienen mayor peso en la predicción final. Este paso es fundamental para alinear los resultados del algoritmo con la fisiopatología observada por los especialistas y garantizar herramientas de diagnóstico más fiables.

Capítulo 4. Resultados y discusión

A continuación, se presenta la integración de los resultados obtenidos en esta investigación. Esta sección expone inicialmente la validación del algoritmo de segmentación, paso previo y esencial para asegurar la fidelidad de los datos extraídos. Posteriormente, el análisis se centra en los resultados derivados del entrenamiento y la aplicación de los modelos de ML.

4.1. Análisis del resultado de segmentación

Para validar la precisión y robustez del algoritmo de segmentación propuesto, este se aplicó al phantom 3D. La Figura 4.1 ilustra el proceso completo del algoritmo aplicado a un corte representativo del phantom, mostrando visualmente cada etapa, desde el preprocesamiento hasta la generación de las métricas de calidad por plano que nutren los resultados consolidados. Se calcularon métricas volumétricas globales para evaluar el rendimiento general del método en la segmentación de los dos cilindros, los cuales simulaban tejidos con diferentes densidades.

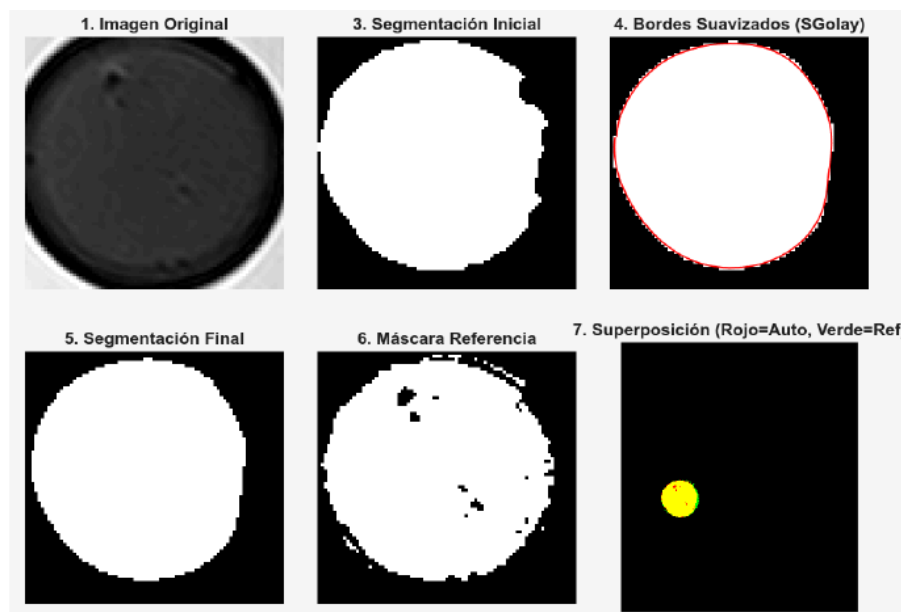


Figura 4.1.- Procesamiento del corte 8 del phantom

En la evaluación del primer cilindro (con un histograma de mayor intensidad), el algoritmo demostró un alto grado de exactitud. Se alcanzó un Coeficiente de Dice del 95.49% y un IoU del 91.38%, lo que indica una excelente concordancia entre el volumen automático y el de referencia. El volumen total segmentado fue de 13301.34 mm³, muy próximo al volumen manual de 13605.07 mm³. Las tasas de error fueron mínimas, registrando una USR del 5.57% y una OSR del 3.34%. Para el segundo cilindro (de menor densidad), el rendimiento fue igualmente sobresaliente. Se obtuvo un Coeficiente de Dice del 97.63% y un IoU del 95.37%. El volumen automático calculado fue de 16539.67 mm³ frente al volumen de referencia de 16301.73 mm³. Las tasas de error se mantuvieron bajas, con una USR del 1.66% y una OSR del 3.12%. Un análisis detallado del rendimiento en cada corte individual se presenta en la Tabla 4.1 y la Tabla 4.2.

Tabla 4.1. Resultados de la segmentación del phantom del cilindro blanco

Planos	Segmentación manual (mm ³)	Método propuesto (mm ³)	VOER	OSR %	USR %	VOF
1	0	0	0			
2	2305	2166.7	0.1	0.0021	0.081	0.9
3	2296.4	2133.5	0.102	0.016	0.087	0.898
4	2257.7	2279.2	0.037	0.024	0.014	0.963
5	2247.8	2121.8	0.116	0.032	0.088	0.884
6	2233.1	2200.5	0.08	0.034	0.048	0.92
7	2265.1	2303.8	0.045	0.032	0.015	0.955
8	0	0	0			
9	0	0	0			

Tabla 4.2. Resultados de la segmentación del phantom del cilindro negro

Planos	Segmentación manual (mm ³)	Método propuesto (mm ³)	VOER	OSR %	USR %	VOF
1	0	0	0			
2	2339.4	2294.6	0.066	0.024	0.044	0.934
3	2341.3	2388.6	0.044	0.033	0.013	0.956
4	2332.7	2377	0.034	0.027	0.008	0.966
5	2340.1	2375.7	0.028	0.022	0.007	0.972
6	2339.4	2377.6	0.027	0.022	0.006	0.973
7	2351.7	2378.2	0.033	0.023	0.012	0.967
8	2194.3	2266.9	0.128	0.086	0.053	0.872
9	62.7	81.2	0.419	0.451	0.157	0.581

Tras haber garantizado la precisión del algoritmo mediante la validación con el phantom 3D es posible transitar con confianza hacia el análisis estadístico del conjunto de datos. La hilación lógica aquí presentada permite transformar las máscaras binarias y los contornos suavizados en un vector de características cuantitativas, moviendo el enfoque desde el procesamiento de imágenes hacia la interpretación diagnóstica de los modelos de ML.

El sustento de este análisis reside en una base de datos de 150 pacientes, la cual integra y expande el conocimiento clínico establecido en investigaciones previas. Parte fundamental de este conjunto de datos proviene de la evaluación de biomarcadores en la tráquea y el árbol vascular pulmonar realizada por Torres Guzmán et al. (2024), asegurando que la arquitectura de las variables morfológicas cuente con una validación clínica antecedente. La base de datos se encuentra etiquetada mediante una variable objetivo dicotómica denominada Clasificación, donde el valor 0 identifica a sujetos control y el valor 1 representa a pacientes con diagnóstico confirmado de COVID-19, permitiendo así un entrenamiento supervisado orientado a la detección de la patología.

4.2. Análisis exploratorio de datos

El análisis exploratorio de datos (EDA) constituye un puente fundamental entre la extracción técnica de biomarcadores y la inteligencia predictiva de los modelos, actuando como un filtro de validación riguroso que garantiza la integridad y coherencia clínica de las 14 variables procesadas. Su ejecución es vital para verificar la calidad de la muestra de 150 pacientes, permitiendo identificar patrones estadísticos, detectar posibles valores atípicos y confirmar el poder discriminatorio de los rasgos morfológicos y texturales antes de ser sometidos a algoritmos supervisados. Al evaluar sistemáticamente la distribución de las clases y la interdependencia entre métricas —como la relación entre el volumen vascular y la severidad de la lesión—, el EDA otorga la transparencia necesaria para alinear el diagnóstico automatizado con la realidad fisiopatológica de enfermedades complejas como la COVID-19, asegurando que el sistema final sea estadísticamente robusto y clínicamente interpretable.

4.2.1. Balance de la variable objetivo

Como se observa en la Figura 4.2, la distribución de la muestra garantiza la robustez del aprendizaje supervisado. Con un 52.0% (78 pacientes) en la clase sin COVID-19 y un 48.0% (72 pacientes) en la clase COVID-19, el conjunto de datos se encuentra equilibrado, lo que previene sesgos en las métricas de sensibilidad y especificidad que se discutirán posteriormente.

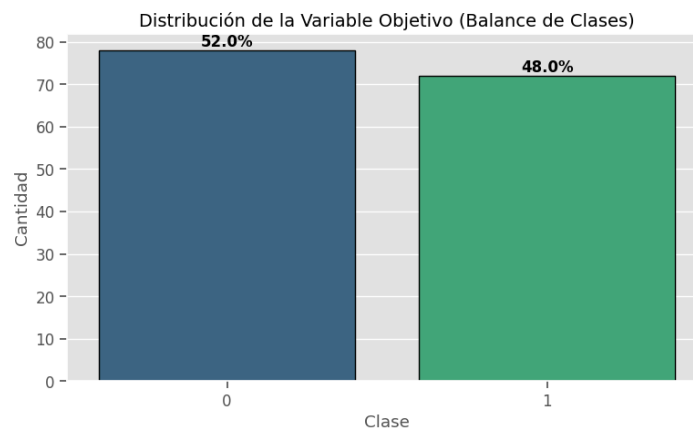


Figura 4.2.- Gráfica de barras de la distribución entre la clase 0 (pacientes sin COVID-19) y 1 (pacientes con COVID-19).

4.2.2. Poder discriminante de los biomarcadores

El análisis visual de la capacidad predictiva de los rasgos extraídos se fundamenta en la observación de los diagramas de caja para cada una de las 14 variables independientes del conjunto de datos. Como se evidencia en la Figura 4.3, la totalidad de los biomarcadores analizados exhiben una separación clara y estadísticamente significativa entre el grupo de control (Clase 0) y los pacientes diagnosticados con COVID-19 (Clase 1), lo que valida la sensibilidad de la metodología para capturar alteraciones patológicas complejas.

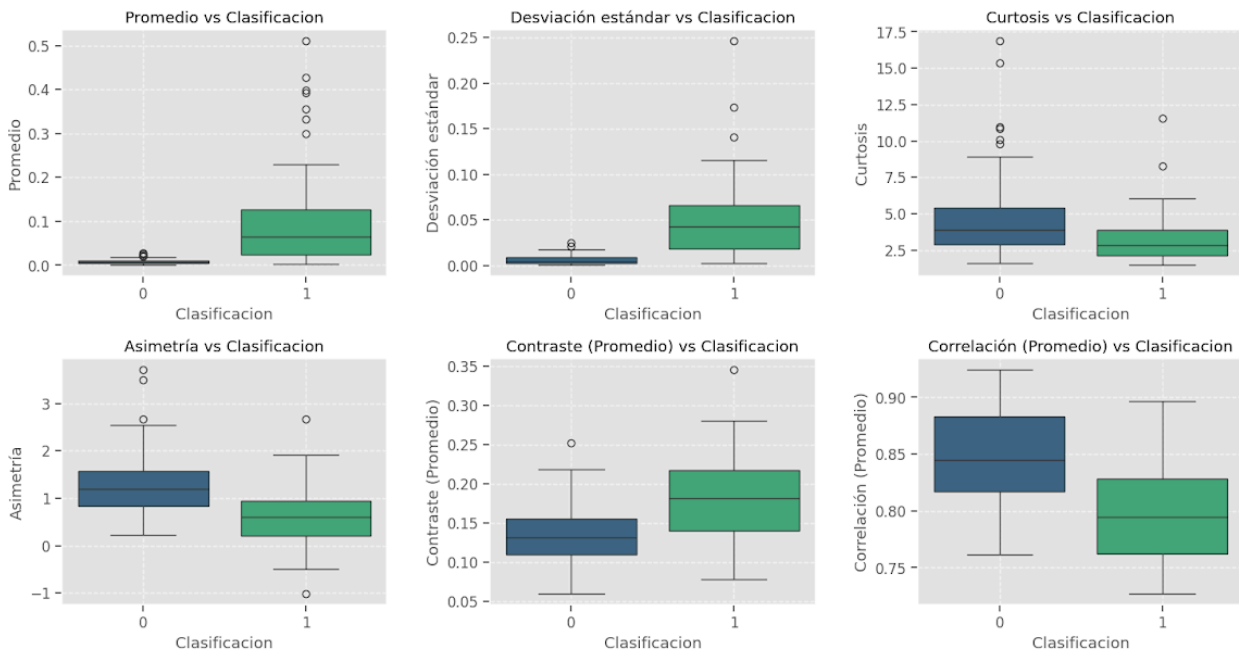


Figura 4.3.- Boxplot de los valores estadísticos (promedio, desviación estándar, curtosis y asimetría) del daño pulmonar y la textura de las imágenes (contraste y correlación) de los 150 pacientes.

La distribución de los datos permite realizar un desglose detallado de esta diferenciación según la naturaleza de las variables:

- Biomarcadores de Lesión Pulmonar:** Las variables Promedio y Desviación estándar muestran la distinción más pronunciada en el análisis; mientras que en los sujetos sanos los valores se agrupan cerca del cero debido a la ausencia de infiltrados, en los pacientes con COVID-19 la mediana se desplaza significativamente hacia valores superiores con un rango de dispersión mayor. Este comportamiento confirma que el sistema de umbralización dual propuesto es capaz de cuantificar con éxito la severidad del daño parenquimatoso.
- Biomarcadores de Textura (GLCM):** Se observa una tendencia inversa y clara en las métricas de textura. La Homogeneidad y la Energía presentan una caída drástica en la Clase 1, indicando la pérdida de uniformidad del tejido pulmonar sano. Por el contrario, el Contraste promedio se eleva notablemente en pacientes infectados, reflejando el incremento de la rugosidad y heterogeneidad espacial causada por las opacidades en vidrio deslustrado y las consolidaciones.

- **Biomarcadores Morfológicos Estructurales:** Existe una separación evidente en la arquitectura de los árboles vascular y aéreo. El Volumen [mm3]_arbol disminuye significativamente en la población enferma, situando su mediana incluso por debajo del primer cuartil del grupo de control. Asimismo, la Tortuosidad_arbol y la Compacidad discreta_airway muestran desplazamientos que confirman una remodelación geométrica profunda, sugiriendo que la infección altera no solo la densidad del tejido sino también la integridad de sus estructuras de conducción.

La ausencia de solapamiento significativo en los rangos intercuartílicos de la mayoría de estas variables demuestra que el vector de características no es solo una representación numérica, sino un conjunto de biomarcadores con un alto valor diagnóstico. Esta diferenciación visual inequívoca constituye la base estadística que permitirá a los modelos de ML alcanzar altos niveles de exactitud en la etapa de clasificación.

4.2.3. Mapa de calor de correlación

El mapa de calor permite identificar cómo interactúan los biomarcadores de distinta naturaleza (morfológicos, texturales y estadísticos), proporcionando una visión integral de la arquitectura de datos antes de la fase de clasificación. (Figura 4.4)

A partir de la observación de la Figura 4.4, se destacan los siguientes hallazgos técnicos:

- **Variables con Alta Colinealidad:** Se identifican grupos de variables con una correlación positiva extremadamente fuerte, lo que sugiere una redundancia en la información capturada. Destacan la relación entre la Homogeneidad (Promedio) y la Energía (Promedio) con un valor de 0.95, así como la dependencia entre el Promedio de la lesión y su Desviación estándar (0.93). Asimismo, la Asimetría y la Curtosis presentan una correlación de 0.87. Esta información es valiosa para algoritmos como Random Forest o XGBoost, los cuales son capaces de gestionar variables redundantes mediante la selección intrínseca de características.

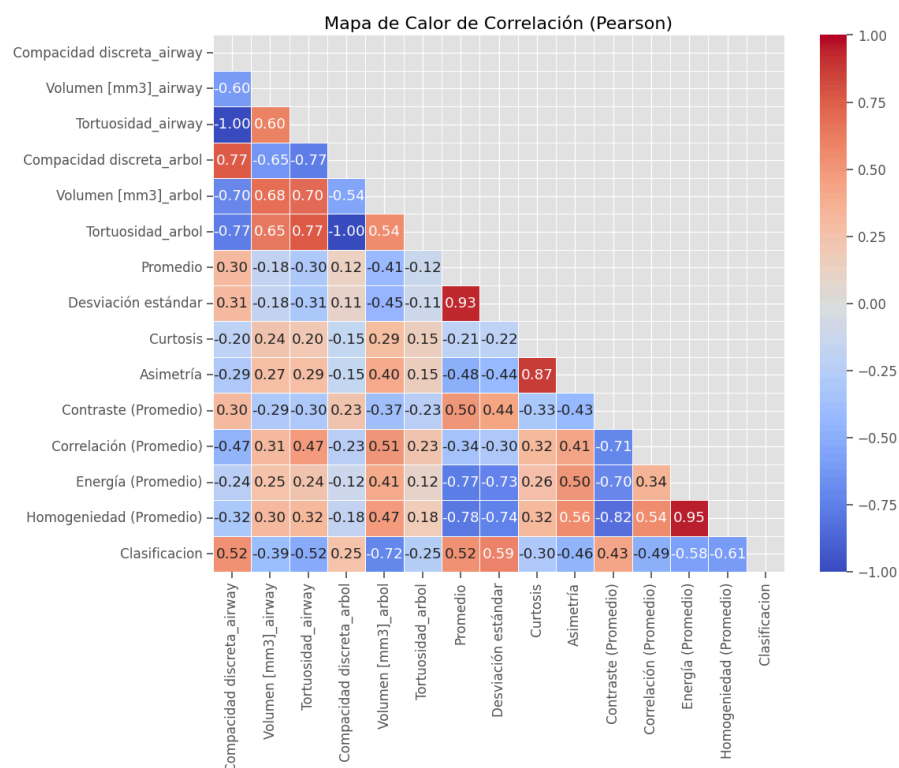


Figura 4.4.- Mapa de correlación de las 14 variables que serán utilizadas como nuestro vector de características.

- Correlación con la Variable Objetivo (Clasificación):** El análisis revela que varios biomarcadores poseen una asociación significativa con la presencia de la enfermedad (Clase 1). El Volumen [mm3]_arbol presenta la correlación negativa más fuerte (-0.72), indicando que la reducción del volumen vascular es un predictor altamente determinante para identificar la patología. Por otro lado, la Desviación estándar (0.59), el Promedio de lesión (0.52) y la Compacidad discreta_airway (0.52) muestran correlaciones positivas moderadas, confirmando que el incremento en la variabilidad del daño y la alteración de la vía aérea son rasgos característicos de la infección.
- Independencia de las Métricas de Textura:** A pesar de su alta correlación interna, las métricas GLCM como el Contraste (0.43) y la Correlación (-0.49) muestran asociaciones moderadas e independientes respecto a las variables morfológicas. Esto sugiere que el análisis de textura de segundo orden aporta

información complementaria sobre la heterogeneidad del parénquima que no puede ser capturada únicamente por la geometría del árbol vascular o aéreo.

La presencia de estas correlaciones significativas, tanto positivas como negativas, con la variable Clasificación, refuerza los hallazgos obtenidos en los diagramas de caja y confirma que el vector de características seleccionado posee la robustez estadística necesaria para proceder con el entrenamiento de modelos supervisados altamente precisos.

4.3. Entrenamiento y optimización de los modelos

El entrenamiento de los modelos se fundamenta en la búsqueda de un equilibrio óptimo entre el sesgo y la varianza, asegurando que el sistema no solo memorice los datos de los 150 pacientes, sino que aprenda patrones fisiopatológicos subyacentes. Para ello, se empleó un flujo de trabajo basado en la Búsqueda en Cuadrícula (Grid Search) asistida por una validación interna.

El dataset completo de 150 pacientes fue dividido inicialmente en un conjunto de entrenamiento (80%, n=120) y un conjunto de prueba independiente (20%, n=30) utilizando la función **train_test_split(...)**. Esta división fue configurada con estratificación activa para asegurar que tanto la clase COVID-19 como la clase No-COVID preservarán su proporción original de impacto en ambos subconjuntos.

Durante la etapa de optimización de hiperparámetros mediante Grid Search, el ajuste del conjunto de entrenamiento se validó utilizando una estrategia de Validación Cruzada Estratificada y Repetida configurada con 5 pliegues y 2 repeticiones. Absolutamente todo el proceso computacional fue fijado de manera transversal con una semilla aleatoria de control **random_state = 42(...)**.

4.3.1. Fundamentos de la validación cruzada (k-fold CV)

Para mitigar el riesgo de sobreajuste y proporcionar una estimación insesgada del rendimiento, se utilizó la Validación Cruzada de k-pliegues. En este procedimiento, el conjunto de entrenamiento se divide en k subconjuntos aleatorios; el modelo se entrena en k-1 pliegues y se evalúa en el pliegue restante, repitiendo el ciclo de forma iterativa.

La métrica de Exactitud CV reportada en esta sección representa la media aritmética de los resultados obtenidos en cada iteración:

$$Exactitud\ CV = \frac{1}{k} \sum_{i=1}^k Acc_i$$

4.3.2. Proceso de optimización por algoritmo

A continuación, se describen los resultados de la optimización y los hiperparámetros que arrojaron el mejor desempeño para cada arquitectura evaluada:

A. Regresión Logística (Logistic Regression)

Se optimizó el parámetro de regularización C, el cual determina la fuerza de la penalización aplicada a los coeficientes del modelo.

- **Mejor Accuracy CV:** 0.8889.
- **Mejor Configuración:** {'C':0.1}.
- **Análisis:** Un valor bajo de C sugiere que el modelo se beneficia de una regularización fuerte, simplificando la frontera de decisión para evitar el ruido en los datos textuales.

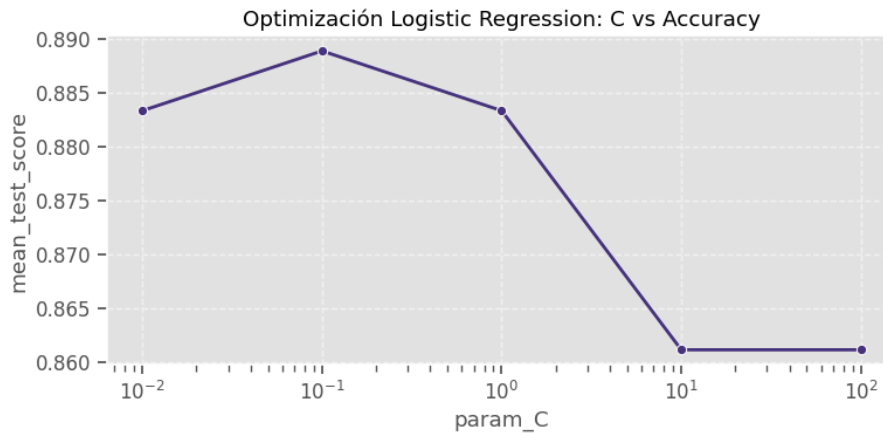


Figura 4.5.- Gráfica de optimización del modelo de regresión logística

B. Bosques Aleatorios (Random Forest)

Se evaluó el impacto del número de árboles en el ensamble para estabilizar la varianza del error.

- **Mejor Accuracy CV:** 0.8944.
- **Mejor Configuración:** {'n_estimators':50}.
- **Análisis:** La convergencia con 50 estimadores demuestra que la combinación de biomarcadores morfológicos y de lesión ofrece una señal lo suficientemente potente para que el ensamble no requiera una complejidad estructural excesiva.

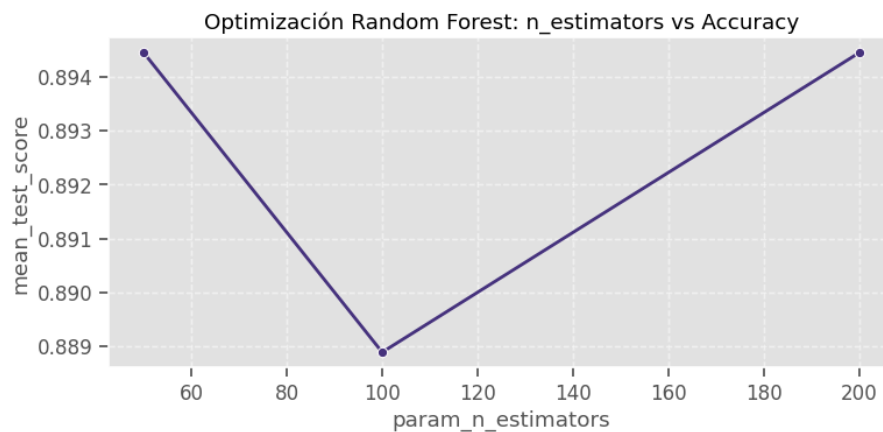


Figura 4.6.- Gráfica de optimización del modelo de random forest

C. Máquinas de Soporte Vectorial (SVM - RBF)

Se utilizó un núcleo de Función de Base Radial (RBF) para manejar la posible no linealidad entre las métricas de textura y la clase patológica.

- **Mejor Accuracy CV:** 0.8778.
- **Mejor Configuración:** {'C':0.1,'gamma':'scale'}.
- **Análisis:** El uso de un margen suave (C=0.1) permite al modelo ignorar pequeñas inconsistencias en los bordes de segmentación, logrando una clasificación más robusta.

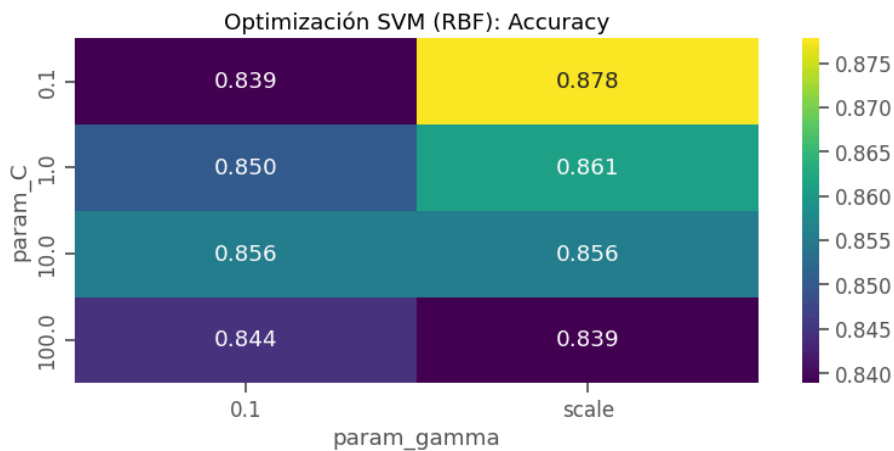


Figura 4.7.- Gráfica de optimización del modelo de SVM

D. K-Vecinos Más Cercanos (KNN)

Se exploró la vecindad óptima para la clasificación basada en la similitud de los vectores de características.

- **Mejor Accuracy CV:** 0.8833.
- **Mejor Configuración:** {'n_neighbors':11}.

- **Análisis:** Una vecindad de 11 individuos permite suavizar la clasificación, capturando la tendencia general del daño pulmonar sin verse afectado por variaciones individuales extremas.

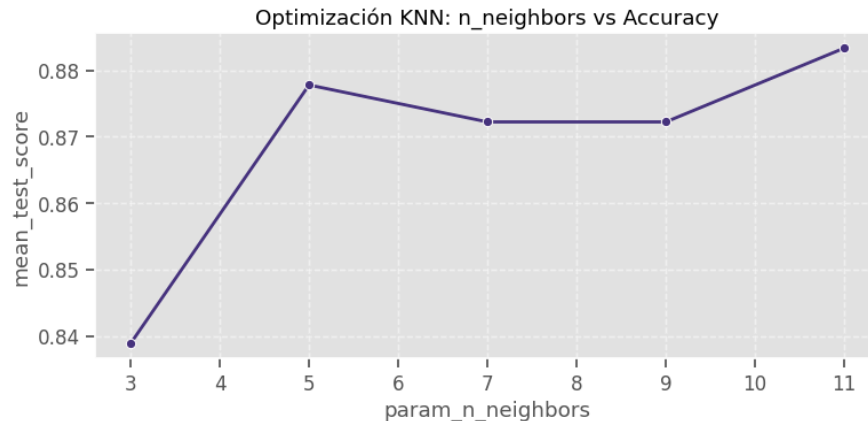


Figura 4.8.- Gráfica de optimización del modelo de KNN

E. Red Neuronal Artificial (MLP)

Se ajustó el parámetro de regularización L2 (alpha) para controlar la magnitud de los pesos en las capas ocultas.

- **Mejor Accuracy CV:** 0.8778.
- **Mejor Configuración:** {'alpha':0.01}.
- **Análisis:** La estabilidad en el entrenamiento indica que la arquitectura de la red neuronal es capaz de procesar las 14 variables de entrada sin incurrir en divergencias de optimización.

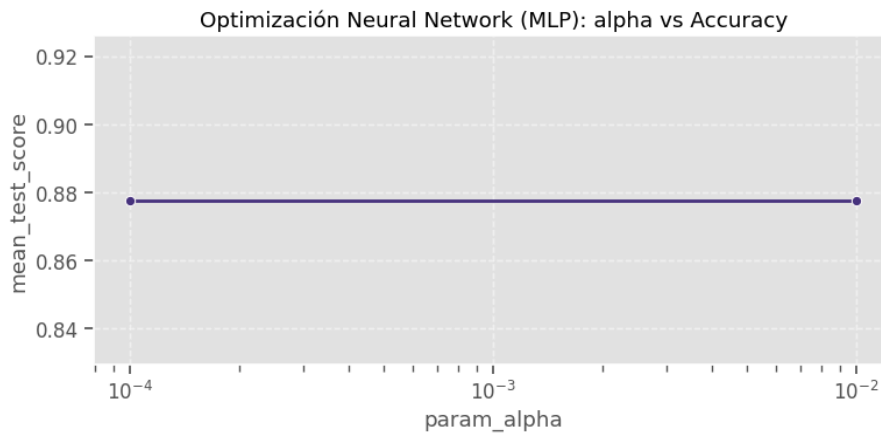


Figura 4.9.- Gráfica de optimización del modelo de MLP

F. XGBoost (Extreme Gradient Boosting)

Este algoritmo de potenciación del gradiente presentó el rendimiento más elevado de toda la experimentación.

- **Mejor Accuracy CV: 0.9056.**
- **Mejor Configuración:** {'learning_rate':0.01,'max_depth':3,'n_estimators':200}.
- **Análisis:** XGBoost aprovecha de manera excepcional la información combinada de los biomarcadores. La profundidad máxima limitada (3) previene la captura de ruido, mientras que la tasa de aprendizaje baja (0.01) permite que el modelo refine su precisión de forma incremental a lo largo de 200 iteraciones.

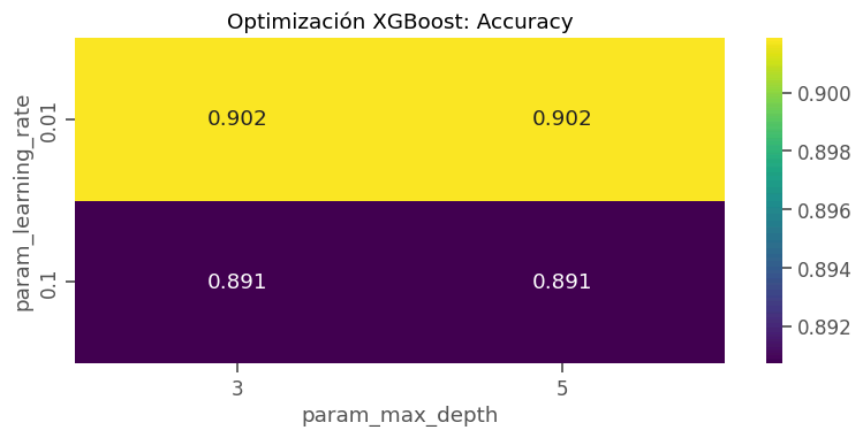


Figura 4.10.- Gráfica de optimización del modelo de XGBoost

4.4. Evaluación de los modelos sobre el conjunto de prueba

Tras concluir la fase de optimización, se realizó la validación definitiva de los modelos utilizando el conjunto de prueba (test set). Este conjunto, que representa el 20% de la muestra total (30 pacientes), se mantuvo estrictamente aislado durante todo el proceso de búsqueda de hiperparámetros y entrenamiento inicial. Evaluar los algoritmos frente a datos que nunca han procesado es el estándar de oro para determinar la capacidad de generalización del sistema en un entorno clínico real.

Tabla 4.3. Resultados Comparativos de Desempeño (Test Set)

Modelo	Exactitud	Precisión	Recall	F1-Score	AUC
XGBoost	0.9667	0.9333	1	0.9655	0.9844
Random Forest	0.9333	0.9286	0.9286	0.9286	0.9866
Red Neuronal	0.9333	0.9286	0.9286	0.9286	0.9866
Regresión Logística	0.9	0.8667	0.9286	0.8966	0.9643
KNN	0.9	0.9231	0.8571	0.8889	0.9777
SVM	0.9	0.9231	0.8571	0.8889	0.9554

4.4.1. Análisis de la evaluación

Un hallazgo técnico relevante es el incremento en la exactitud observado al pasar de la Validación Cruzada (Sección 4.2) a la evaluación final en el conjunto de prueba. Mientras que el modelo líder, XGBoost, reportó un Accuracy CV de 0.9056 durante la optimización, su desempeño alcanzó un 0.9667 en la validación definitiva.

Este comportamiento no es indicativo de un sobreajuste (overfitting); por el contrario, el overfitting se caracteriza por una exactitud alta en el entrenamiento pero drásticamente baja en la prueba. En este estudio, el incremento responde a dos factores fundamentales:

- **Uso íntegro del conjunto de entrenamiento:** Durante la validación cruzada, el modelo se entrena solo con una fracción de los datos para evaluarse en el pliegue restante. Sin embargo, para la prueba final, el modelo se reentrena utilizando el

100% de los datos de entrenamiento (el 80% de la muestra total). Este volumen adicional de ejemplos permite que el algoritmo ajuste sus fronteras de decisión de forma más robusta y precisa.

- **Poder predictivo de los biomarcadores:** La clara separación observada en el análisis exploratorio (Sección 4.1.2) confirma que los biomarcadores morfológicos y texturales poseen una firma patológica muy marcada. Al dotar a un modelo ya optimizado con descriptores tan discriminantes, la probabilidad de clasificación correcta aumenta significativamente, incluso ante datos nuevos.

4.4.2. Análisis de modelos de alto desempeño: XGBoost, Random Forest y MLP

En esta categoría se agrupan los algoritmos que lograron una exactitud superior al 93%, demostrando una capacidad excepcional para procesar la complejidad multidimensional de los 14 biomarcadores.

- **XGBoost (Líder del Estudio):** Con una Exactitud de 0.9667, este modelo de potenciación de gradiente se consolidó como la herramienta más fiable para el diagnóstico. Su matriz de confusión (Figura 4.11) revela un desempeño clínico ideal: alcanzó un Recall de 1.0000, lo que significa que identificó correctamente a los 14 pacientes positivos en el conjunto de prueba sin generar un solo falso negativo (FN=0). Este resultado es crítico para evitar que pacientes infectados sean dados de alta erróneamente.

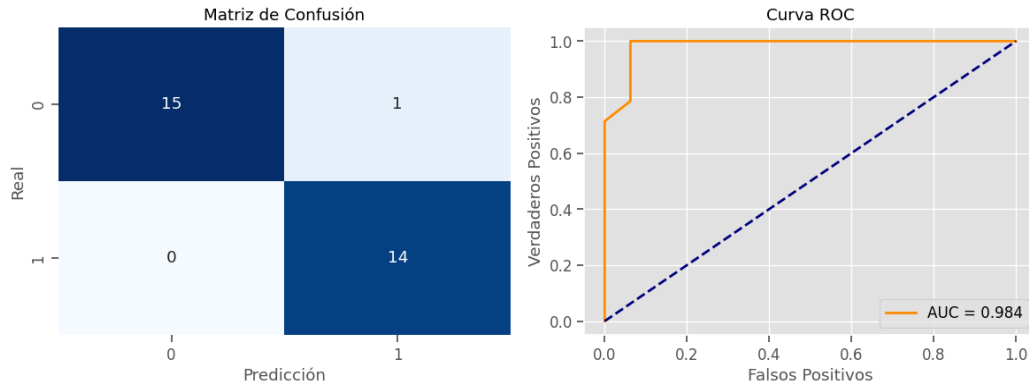


Figura 4.11.- Rendimiento del modelo XGBoost mostrando la matrix de confusión y curva ROC

- Random Forest y MLP:** Ambos modelos alcanzaron una exactitud de 0.9333 y el valor de AUC más alto del estudio (0.9866). Como se observa en sus respectivas evaluaciones (Figuras 4.12), ambos clasificadores mostraron un equilibrio notable, cometiendo solo un error de clasificación por clase (1 FP y 1 FN). Esto sugiere que tanto el ensamble de árboles como las capas neuronales capturaron con éxito la variabilidad de la textura y morfología pulmonar.

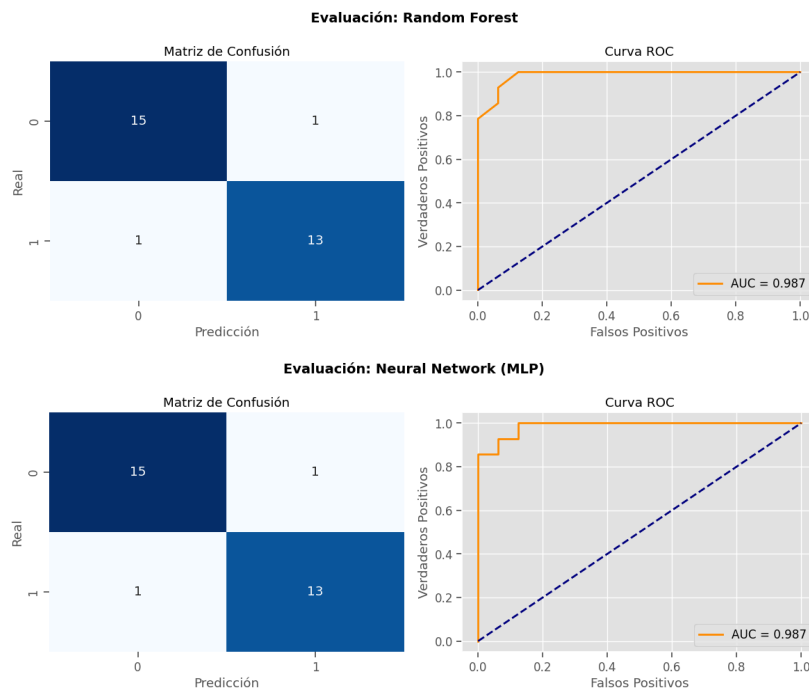


Figura 4.12.- Rendimiento del modelo RF y MLP mostrando sus matrix de confusión y curva ROC

4.4.3. Evaluación de modelos de estabilidad robusta: Regresión logística, KNN y SVM

Los tres modelos restantes alcanzaron una exactitud del 0.9000, un nivel de rendimiento que sigue siendo altamente competitivo para herramientas de diagnóstico asistido por computadora, superando incluso las estimaciones iniciales obtenidas en la fase de entrenamiento.

- **Regresión logística:** A pesar de su simplicidad lineal, este modelo demostró una inclinación favorable hacia la sensibilidad (Recall: 0.9286). Su matriz de confusión (Figura 4.13) muestra que solo omitió un caso positivo, aunque presentó una tasa ligeramente superior de falsos positivos (FP=2). Es un modelo valioso por su alta interpretabilidad en la relación entre el volumen vascular y la clase patológica.

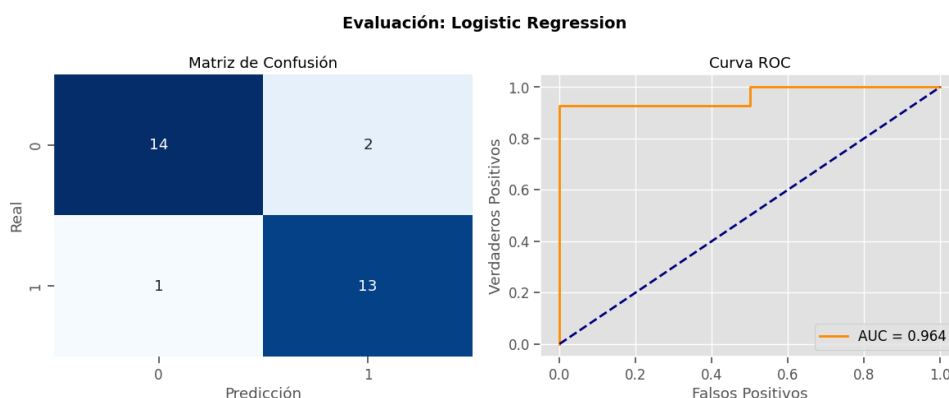


Figura 4.13.- Rendimiento del modelo de regresión logística mostrando la matrix de confusión y curva ROC

- **KNN y SVM:** Ambos algoritmos presentaron métricas idénticas de exactitud (0.9000) y F1-Score (0.8889). Curiosamente, estos modelos priorizaron la precisión sobre la sensibilidad, logrando un 0.9231 en Precision pero dejando pasar dos falsos negativos (FN=2) en sus respectivas matrices (Figuras 4.14). Esto indica que son modelos sumamente cautelosos; cuando diagnostican a un paciente como "positivo", la probabilidad de acierto es muy alta, aunque pueden ser menos sensibles ante daños pulmonares incipientes.

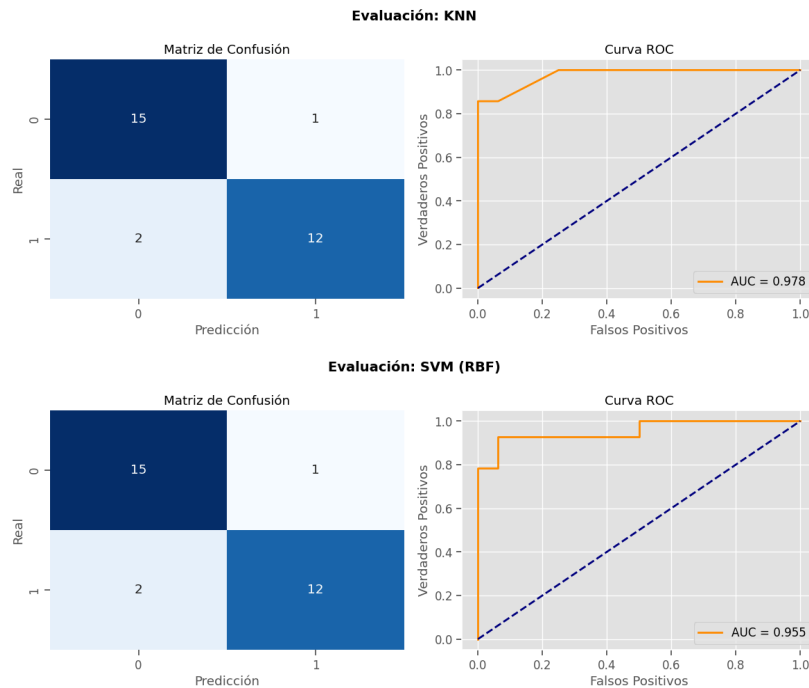


Figura 4.14.- Rendimiento del modelo KNN y SVM mostrando sus matrix de confusión y curva ROC

Se permite concluir que, si bien XGBoost es el modelo superior en términos de exactitud y sensibilidad clínica, la metodología basada en 14 biomarcadores morfológicos y texturales es lo suficientemente robusta como para permitir que cualquier arquitectura de ML alcance resultados de alta fidelidad diagnóstica.

Para evaluar a profundidad las capacidades y los límites de generalización del sistema unificado, se procedió a realizar una auditoría diagnóstica de las instancias donde el mejor algoritmo (XGBoost) incurrió en una clasificación errónea. Durante la validación sobre el conjunto test set, el modelo registró únicamente un Falso Positivo, correspondiente al paciente indexado con el número 23, cuyo identificador de origen en la base de datos de contraste (TCIA) es MED_LYMPH_030.

Al realizar la extracción de sus biomarcadores computacionales y contrastarlos con los promedios globales de la muestra, se identificó la razón técnica de la falla diagnóstica. El paciente 23 exhibió un volumen del árbol vascular de apenas 75,006 mm³, una cifra críticamente alejada de la media poblacional de la clase No-COVID $\mu_0 = 248,183.32$ mm³ e incluso sensiblemente inferior al promedio de los pacientes diagnosticados con daño

por SARS-CoV-2 $u_1 = 89,474.27 \text{ mm}^3$. De igual forma, su descriptor textural de energía de co-ocurrencia de niveles de gris (GLCM) se ubicó en un valor bajo de 0.7887, reflejando una pérdida de uniformidad y una heterogeneidad tisular que se aproxima al comportamiento de la clase patológica ($u_1 = 0.7618$).

Desde una perspectiva de bioingeniería médica, este comportamiento anómalo encuentra su plena justificación en el contexto clínico del estudio original de contraste. Al pertenecer a la cohorte CT Lymph Nodes, este sujeto fue sometido a tomografía computarizada debido a la presencia de linfadenopatías mediastinales o masas tumorales. La presencia de estas estructuras ocupativas en el mediastino ejerce un efecto de compresión mecánica intratorácica que podría deformar la geometría pulmonar, restringe el llenado vascular y altera la uniformidad de la densidad del tejido circundante. En consecuencia, el algoritmo XGBoost detectó posibles rasgos macroestructurales de colapso vascular y rugosidad que correlacionan directamente con la signatología del COVID-19 grave, provocando un veredicto erróneo debido a la similitud fenotípica de los biomarcadores, a pesar de que el paciente presentaba un índice de lesión bajo ($\text{ROI} = 0.0066$).

4.5. Análisis de importancia de variables e interpretabilidad

Para este análisis se utilizaron dos técnicas complementarias: la Importancia por Permutación, que mide la dependencia global del modelo hacia cada variable, y los valores SHAP, que cuantifican la contribución exacta de cada biomarcador al valor de la predicción final basándose en la teoría de juegos cooperativos.

4.5.1. Interpretabilidad en el modelo líder: XGBoost

El modelo XGBoost basa su alta precisión en una jerarquía clara de biomarcadores. Como se observa en la gráfica de Importancia por Permutación y en el SHAP Global (Figura 4.15), el factor más determinante para este algoritmo es el Volumen $[\text{mm}^3]_{\text{arbol}}$.

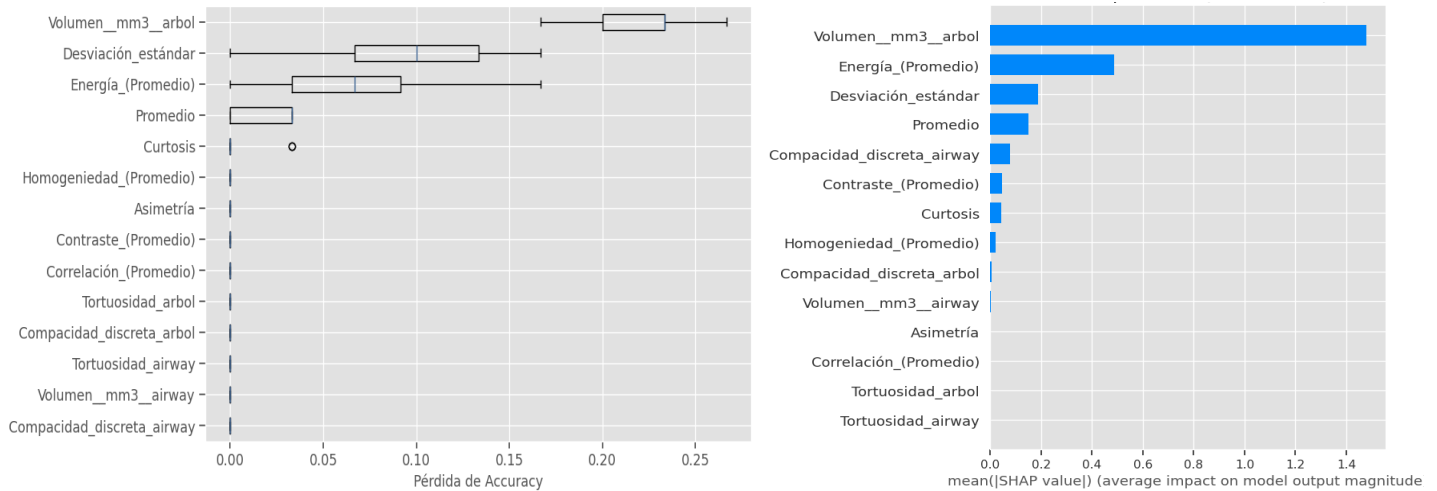


Figura 4.15.- Del lado derecho se muestra la dependencia y del lado izquierdo el impacto de cada variable para el modelo XGBoost

- **Dirección del Impacto:** El gráfico de resumen SHAP (Figura 4.16) revela una relación clínica fundamental: los valores bajos de volumen vascular (puntos azules) presentan una correlación directa con un impacto positivo en la salida del modelo ($f(x)$), lo que confirma que la reducción del lecho vascular es el principal indicador de la patología COVID-19 en esta cohorte.

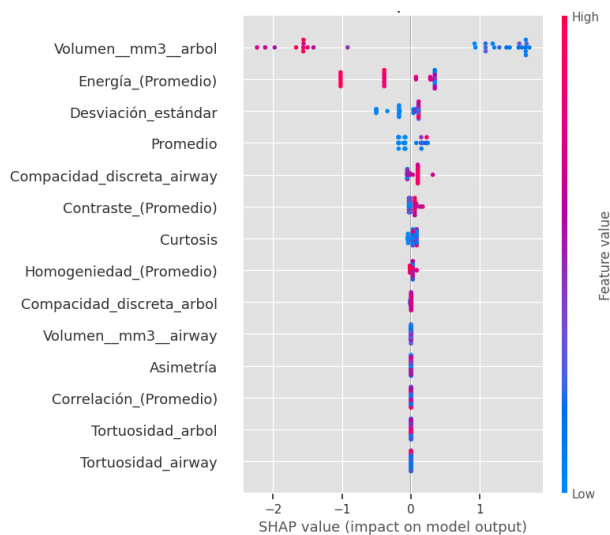


Figura 4.16.- Gráfica de dirección de impacto de cada variable para el modelo XGBoost

- **Factores Secundarios:** La Energía (Promedio) y la Desviación estándar de la lesión siguen en orden de importancia. Una baja energía textural (pérdida de uniformidad) y una alta variabilidad en el ratio de lesión actúan como confirmadores del daño parenquimatoso.
- **Decisión Local (Paciente #2):** En la cascada de decisión (Figura 4.17), se observa un caso positivo donde el modelo predice una probabilidad de enfermedad muy alta ($f(x)=2.304$). La contribución más fuerte proviene de la reducción del volumen vascular (+1.57), seguida por la pérdida de energía textural (+0.34), demostrando cómo el algoritmo suma evidencias morfológicas y estadísticas para emitir un juicio.

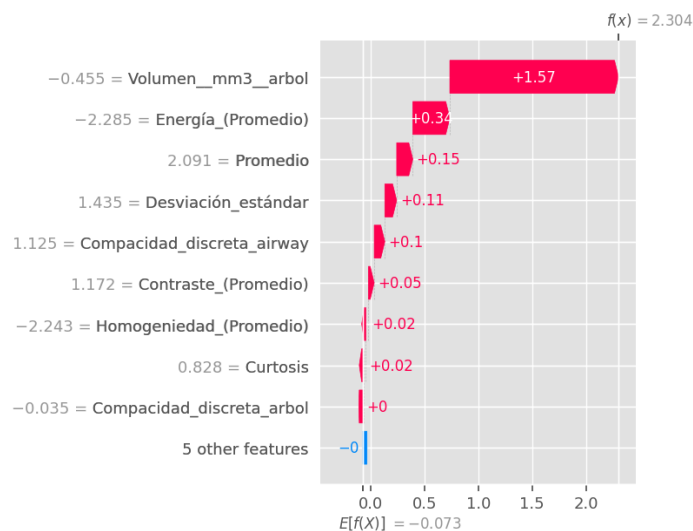


Figura 4.17.- Gráfica de toma de decisión del modelo XGBoost para clasificar en el paciente 2

4.5.2. Interpretabilidad en la red neuronal: MLP

A pesar de ser un modelo de arquitectura distinta, la Red Neuronal converge en hallazgos similares a XGBoost, lo que otorga una mayor validez científica a los biomarcadores seleccionados.

- **Jerarquía de Variables:** El gráfico de importancia global para MLP (Figura 4.18) sitúa nuevamente al Volumen [mm3]_arbol y a la Energía (Promedio) en la cima.

Sin embargo, este modelo otorga un peso mayor a la Asimetría y a la Correlación (Promedio), sugiriendo que la red neuronal aprovecha mejor las relaciones de segundo orden en la textura para compensar su estructura no lineal.

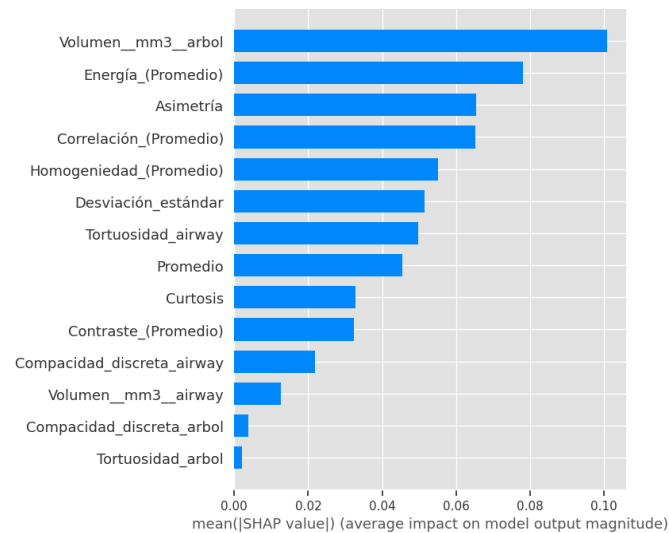


Figura 4.18.- Gráfica de importancia de cada variable para el modelo MLP

- **Consistencia Clínica:** El gráfico de impacto (Figura 4.19) muestra que la red identifica la enfermedad mediante la detección de valores bajos de volumen y energía, junto con valores anómalos de asimetría en la distribución de la lesión. Esto refuerza la idea de que el daño por COVID-19 no solo es volumétrico, sino que altera la simetría de la densidad pulmonar.

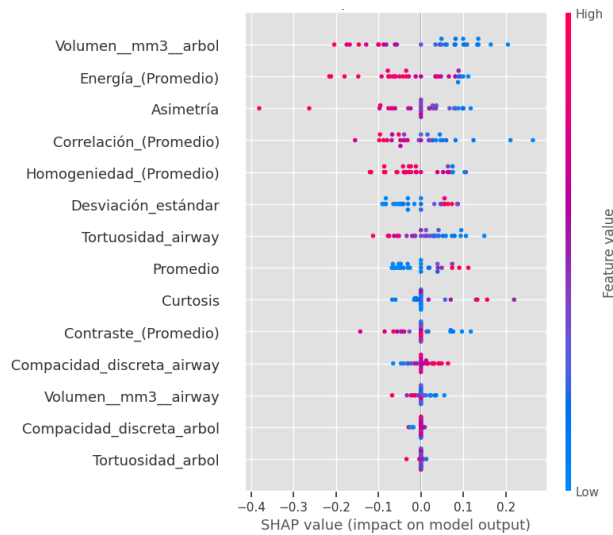


Figura 4.19.- Gráfica de dirección de impacto de cada variable para el modelo MLP

- **Decisión Local (Paciente #4):** En el análisis individual para el paciente #4 (Figura 4.20), el modelo predice correctamente la clase 1 ($f(x)=1$). Aquí, el principal impulsor fue el Promedio de la lesión (+0.11), seguido por la Energía textural (+0.1) y la Asimetría (+0.09), ilustrando cómo la red neuronal prioriza los descriptores estadísticos de la lesión para confirmar casos donde la morfología vascular podría ser menos evidente.

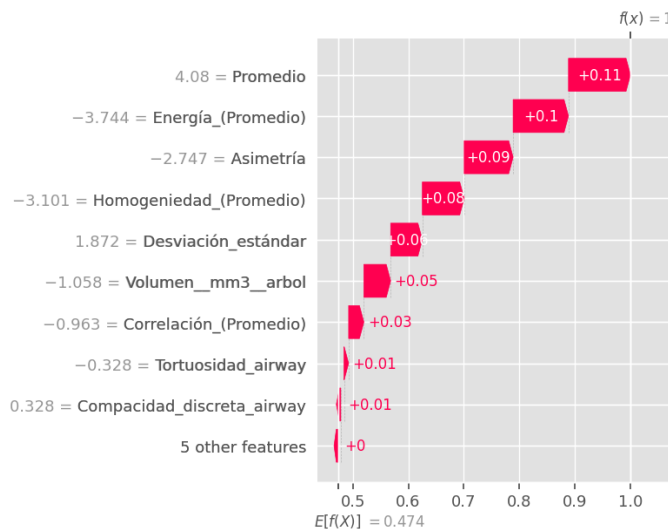


Figura 4.20.- Gráfica de toma de decisión del modelo MLP para clasificar en el paciente 4

4.5.3. Interpretabilidad en el modelo KNN

El modelo KNN basa su clasificación en la similitud local dentro del espacio de características de 14 dimensiones. Al "abrir" este modelo, se observa una dependencia equilibrada entre la morfología vascular y la organización textural del parénquima.

- **Jerarquía Global:** Según el análisis SHAP Global (Figura 4.21) y la importancia por Permutación (Figura 4.21), el Volumen [mm3]_arbol es el predictor más influyente, seguido de cerca por la Energía (Promedio) y la Homogeneidad (Promedio). Esto indica que KNN identifica a los pacientes con COVID-19 buscando "vecinos" que compartan una reducción en la red vascular y una pérdida de uniformidad en la textura del tejido.

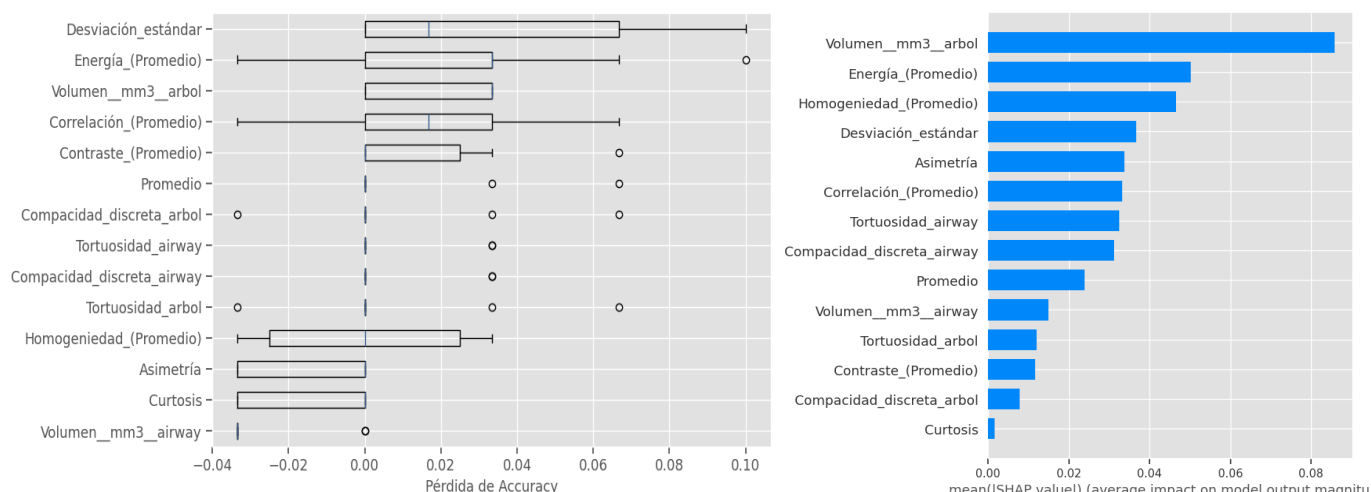


Figura 4.21.- Del lado derecho se muestra la dependencia y del lado izquierdo el impacto de cada variable para el modelo KNN

- **Dirección del Impacto:** El gráfico de Dirección del Impacto (Figura 4.22) muestra que los valores bajos de volumen (puntos azules) incrementan significativamente la probabilidad de pertenecer a la clase patológica. Asimismo, la alta concentración de puntos rojos (valores altos) en el lado negativo de los biomarcadores de textura confirma que un pulmón sano se caracteriza por niveles elevados de homogeneidad y energía.

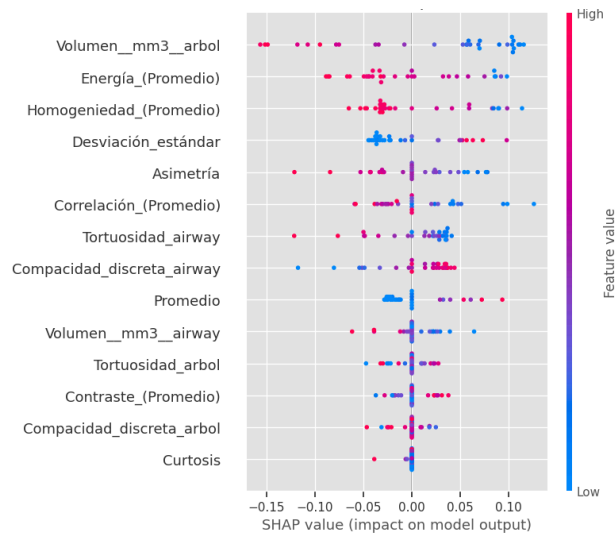


Figura 4.22.- Gráfica de dirección de impacto de cada variable para el modelo KNN

- **Decisión Local (Paciente #2):** En la cascada de decisión para el Paciente #2 (Figura 4.23), el modelo predice con éxito la clase 1 ($f(x)=1$). Es interesante notar que, para este caso específico, la Homogeneidad (+0.09) y la Energía (+0.09) aportaron más al diagnóstico que el volumen vascular solo (+0.05), demostrando la sensibilidad del KNN para capturar daños texturales sutiles que definen la vecindad del paciente patológico.

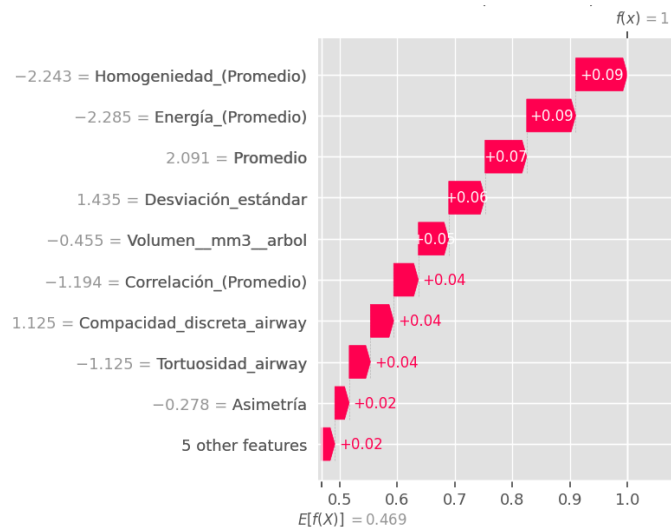


Figura 4.23.- Gráfica de toma de decisión del modelo KNN para clasificar en el paciente 2

4.5.4. Interpretabilidad en el modelo SVM

La Máquina de Soporte Vectorial con núcleo RBF proyecta los biomarcadores a un espacio de mayor dimensionalidad para encontrar el hiperplano de separación óptimo. El análisis XAI revela cómo este modelo utiliza los márgenes de decisión para separar las clases.

- **Variables Determinantes:** El SHAP Global para SVM (Figura 4.24) confirma una vez más que el Volumen [mm3]_arbol encabeza la importancia, pero otorga un peso mayor a la Correlación (Promedio) en comparación con los modelos anteriores. Esto sugiere que SVM aprovecha la dependencia lineal entre los niveles de gris para definir la frontera entre tejido sano y enfermo.

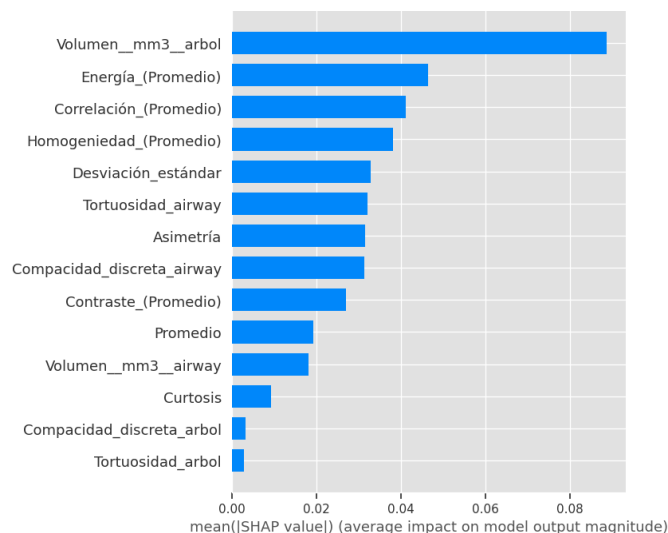


Figura 4.24.- Gráfica de importancia de cada variable para el modelo SVM

- **Consistencia en el Impacto:** La Figura 4.25 refleja un patrón coherente con XGBoost y MLP: la disminución del volumen vascular y de la energía textural son los motores que "empujan" al paciente hacia el lado patológico del hiperplano. La dispersión de los valores de Desviación estándar de la lesión también muestra un impacto positivo claro cuando el daño es heterogéneo.

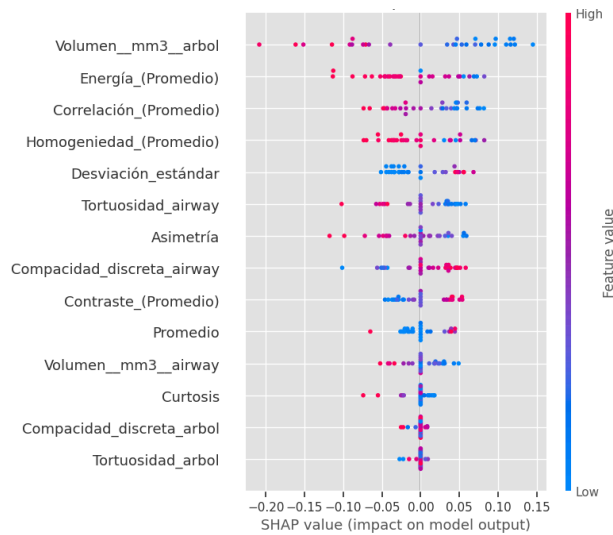


Figura 4.25.- Gráfica de dirección de impacto de cada variable para el modelo SVM

- Decisión Local (Paciente #18):** En el análisis del Paciente #18 (Figura 4.26), el modelo alcanza un valor de decisión de $f(x)=0.98$. La reducción del Volumen del árbol fue el principal factor impulsor (+0.1), seguido de la pérdida de Energía (+0.08) y Homogeneidad (+0.07). Este desglose prueba que SVM requiere la concurrencia de múltiples biomarcadores alterados para emitir un diagnóstico de alta certeza.

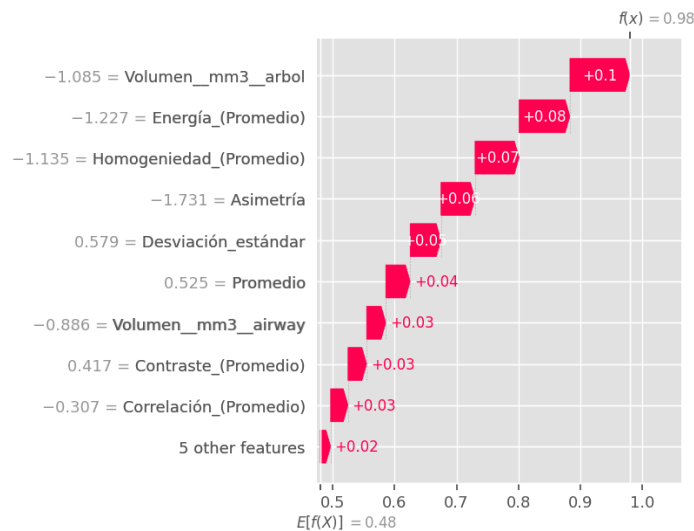


Figura 4.26.- Gráfica de toma de decisión del modelo SVM para clasificar en el paciente 18

4.5.5. Interpretabilidad en el modelo Random Forest

El modelo Random Forest utiliza una multitud de árboles de decisión para votar por la clase final. Al analizar su estructura mediante técnicas XAI, se observa una alta dependencia de los rasgos geométricos y la densidad lesional.

- **Variables Determinantes:** El análisis SHAP Global (Figura 4.27) y la Importancia por Permutación (Figura 4.27) posicionan al Volumen [mm3]_arbol como el predictor soberano. Sin embargo, a diferencia de otros modelos, Random Forest otorga una importancia significativamente mayor al Promedio de la lesión, situándose como el segundo factor más relevante para la toma de decisiones.

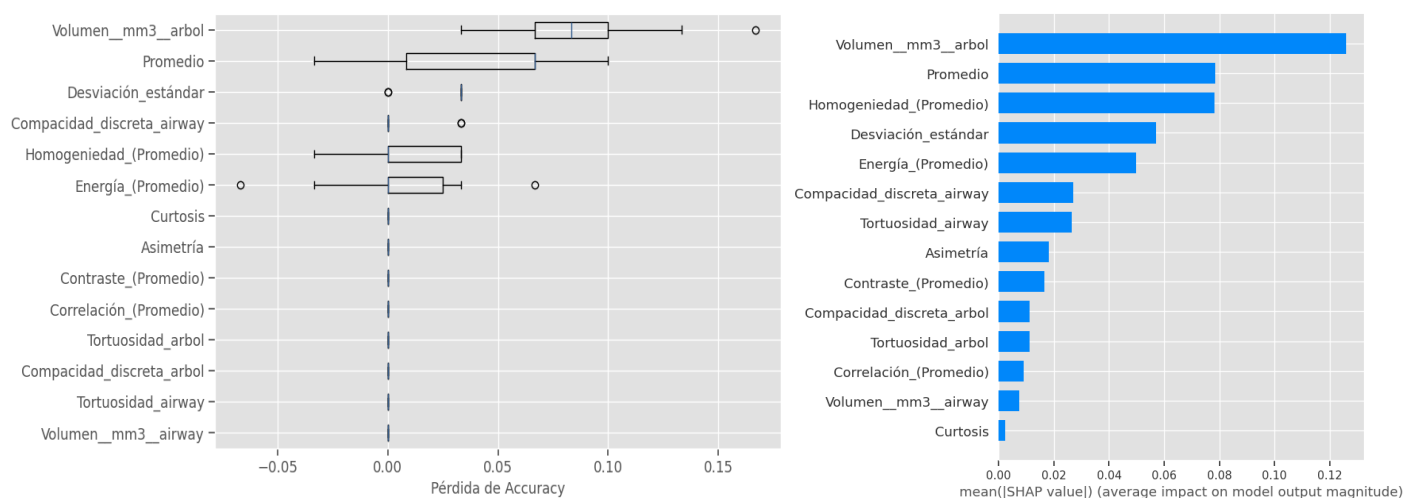


Figura 4.27.- Del lado derecho se muestra la dependencia y del lado izquierdo el impacto de cada variable para el modelo RF

- **Dirección del Impacto:** En el gráfico de Dirección del Impacto (Figura 4.28), se observa que valores bajos de volumen vascular (puntos azules) y valores altos de promedio lesional (puntos rojos) "empujan" la predicción hacia la clase COVID-19. La Homogeneidad y la Energía también muestran un impacto claro: su disminución es un síntoma inequívoco de la patología capturado por el ensamble.

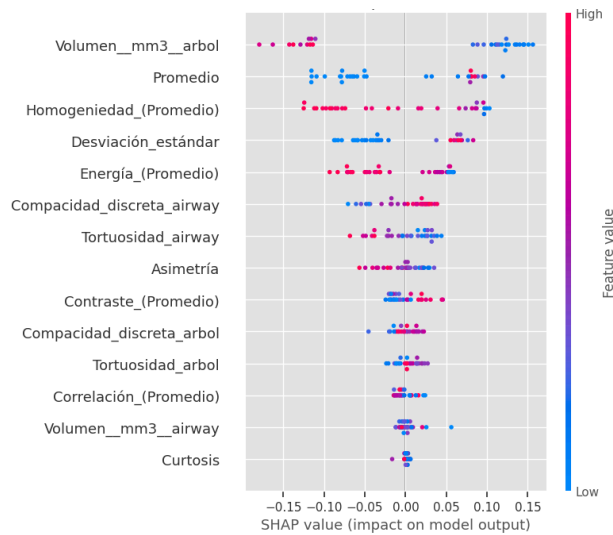


Figura 4.28.- Gráfica de dirección de impacto de cada variable para el modelo RF

- Decisión Local (Paciente #22):** La cascada de decisión para el Paciente #22 (Figura 4.29) muestra una predicción exitosa de la clase 1 ($f(x)=1$). El modelo justifica este diagnóstico principalmente por la reducción del Volumen del árbol (+0.12) y el incremento del Promedio lesional (+0.1), demostrando cómo la arquitectura de árboles suma evidencias para superar el umbral de clasificación.

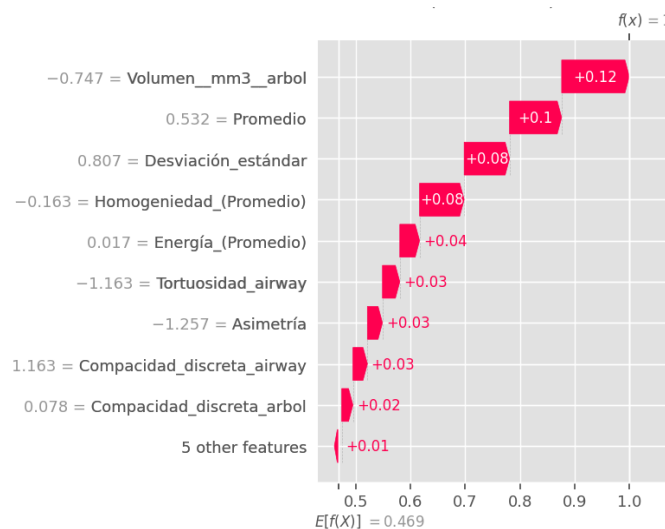


Figura 4.29.- Gráfica de toma de decisión del modelo RF para clasificar en el paciente 22

4.5.6. Interpretabilidad en el modelo Regresión Logística

La regresión logística, por su naturaleza lineal, ofrece una visión directa de cómo cada biomarcador altera las "probabilidades" del diagnóstico. Este modelo es especialmente útil para confirmar la relación causa-efecto de las variables seleccionadas.

- **Factores de Riesgo:** El SHAP Global (Figura 4.30) confirma que el Volumen [mm3]_arbol es la variable con mayor peso. No obstante, la Regresión Logística destaca la Energía (Promedio) y la Homogeneidad (Promedio) con pesos mucho más altos que los modelos no lineales, sugiriendo que la pérdida de uniformidad textural es un indicador linealmente muy potente de la enfermedad.

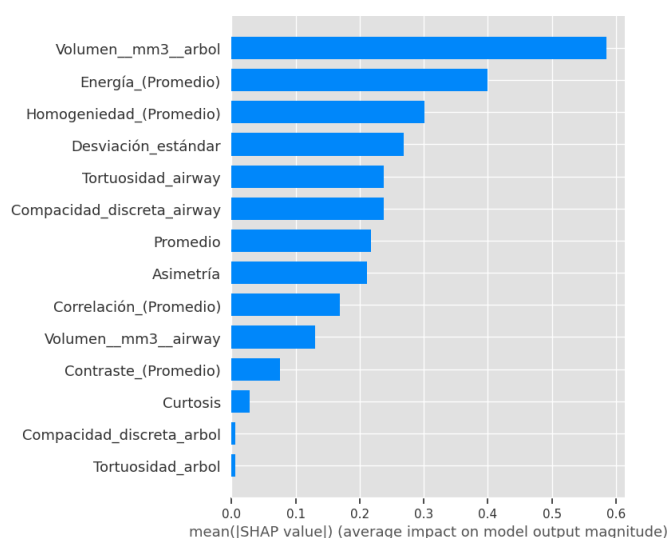


Figura 4.30.- Gráfica de importancia de cada variable para el modelo RL

- **Consistencia de la Frontera:** El gráfico de Dirección del Impacto (Figura 4.31) presenta una separación muy nítida. Los puntos azules de volumen y textura se alinean agresivamente hacia valores SHAP positivos, validando que el modelo asocia linealmente el deterioro estructural con la presencia de COVID-19.

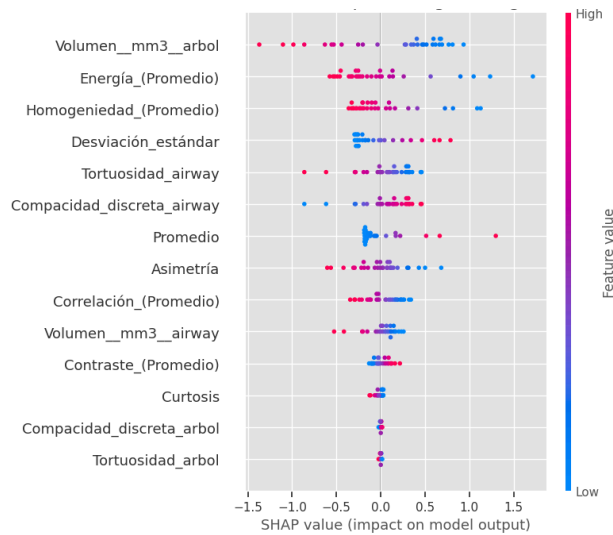


Figura 4.31.- Gráfica de dirección de impacto de cada variable para el modelo RL

- Decisión Local (Paciente #4):** En el análisis del Paciente #4 (Figura 4.32), el modelo emite un valor de log-odds de $f(x)=6.875$ (una probabilidad altísima de enfermedad). Es notable que, para este paciente, los impulsores principales fueron la Energía (+1.72) y la Homogeneidad (+1.13), por encima incluso del volumen vascular. Esto demuestra que la Regresión Logística es sumamente sensible a los cambios texturales del parénquima cuando estos son muy pronunciados.

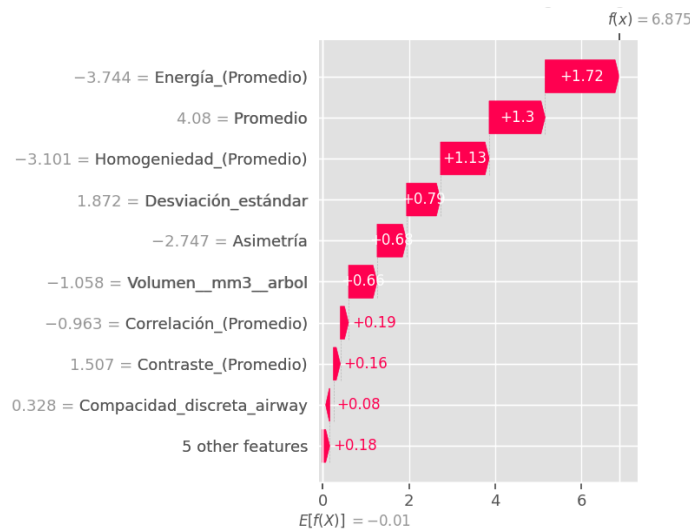


Figura 4.32.- Gráfica de toma de decisión del modelo RL para clasificar en el paciente 4

4.6. Discusión

A pesar de los altos niveles de exactitud alcanzados, es imperativo reconocer que este estudio presenta limitaciones intrínsecas que deben ser consideradas para la interpretación de los resultados. Una de las principales restricciones reside en el tamaño de la cohorte, compuesta por 150 pacientes; si bien esta cifra es estadísticamente sólida para el entrenamiento de algoritmos de ML, una muestra más extensa y diversa permitiría capturar una mayor variabilidad fenotípica de las secuelas post-COVID-19. Además, al tratarse de un estudio de centro único, existe la posibilidad de que los modelos presenten un sesgo hacia los protocolos de adquisición de imagen específicos de la institución de origen, lo que podría afectar la generalización de los resultados en tomógrafos de diferentes fabricantes o con distintos parámetros de reconstrucción.

Otra limitación significativa es la especificidad del diagnóstico. La investigación se centró exclusivamente en la clasificación binaria entre sujetos sanos y pacientes con COVID-19, sin abordar el diagnóstico diferencial con otras patologías pulmonares de apariencia radiológica similar, como la fibrosis pulmonar idiopática o neumonías virales de otra índole. Esta distinción es crucial para la utilidad clínica del sistema en un entorno hospitalario general, donde la coexistencia de múltiples enfermedades respiratorias exige que los biomarcadores morfológicos y texturales demuestren una especificidad diagnóstica que trascienda la infección por SARS-CoV-2.

No obstante, la importancia de este trabajo radica en haber demostrado que un enfoque basado en biomarcadores diseñados e interpretabilidad puede superar las barreras de adopción de la inteligencia artificial en la medicina. Mientras que las tendencias actuales se inclinan hacia modelos masivos de datos, esta investigación prueba que la ingeniería de características fundamentada en la fisiopatología pulmonar permite obtener diagnósticos robustos con una fracción del costo computacional y una total transparencia en la toma de decisiones.

Al comparar el desempeño obtenido con el estado del arte en DL, los resultados son altamente competitivos. Modelos de DL referenciados en la literatura suelen reportar

precisiones en rangos del 90% al 98%, pero requieren miles de imágenes anotadas y operan como "cajas negras". En contraste, el modelo XGBoost desarrollado en esta tesis alcanzó una exactitud del 96.67% y una sensibilidad del 100% utilizando un conjunto de datos significativamente menor. Esto demuestra que el uso de biomarcadores morfológicos, textura y daño capturan la esencia del daño pulmonar de forma tan eficaz como las redes neuronales profundas, con la ventaja añadida de que cada predicción es rastreable y explicable para el especialista clínico. No obstante, los altos niveles de desempeño obtenidos deben interpretarse con cautela debido al tamaño acotado del conjunto de prueba independiente ($n = 30$). En una muestra de estas dimensiones, un solo error de clasificación altera la exactitud global en un 3.33%, lo que amplía los intervalos de confianza del sistema. Por lo tanto, el recall perfecto de 1.0 alcanzado por XGBoost debe tipificarse como un rendimiento altamente competitivo y prometedor dentro de este dataset específico, en lugar de una frontera de decisión absoluta. Para confirmar la capacidad de generalización del pipeline e implementar de forma segura este modelo en un entorno clínico real, ser imprescindible una validación externa cruzada con una cohorte multicéntrica sustancialmente mayor.

La unanimidad con la que el volumen vascular del árbol se consolidó como el descriptor de mayor peso predictivo de manera simultánea en algoritmos estructuralmente disímiles (XGBoost, Random Forest, MLP, KNN y SVM) constituye uno de los hallazgos más robustos de esta investigación. Esta convergencia metodológica elimina la posibilidad de un sesgo del clasificador y encuentra un sustento directo en la fisiopatología pulmonar. Los resultados apoyan la hipótesis planteada y son consistentes con la línea de investigación de Torres Guzmán et al. (2024), quienes documentaron una alteración geométrica profunda y una reducción del lecho capilar menor en el árbol vascular pulmonar debido a los procesos microtrombóticos e inflamatorios característicos del SARS-CoV-2. Esto demuestra que las métricas macroestructurales capturan la esencia de la patología de forma tan eficaz como los descriptores locales.

Finalmente, esta investigación rompe la brecha entre la complejidad técnica y la utilidad práctica. Al lograr métricas de desempeño equivalentes a las del DL mediante un

paradigma de XAI, se garantiza no sólo la precisión del diagnóstico, sino también la confianza necesaria para su futura integración en protocolos de triaje y seguimiento de pacientes post-pandemia.

Conclusiones

En conclusión, la presente investigación cumplió satisfactoriamente con el objetivo general al desarrollar un sistema de clasificación de pacientes con lesiones pulmonares por COVID-19 mediante algoritmos de inteligencia artificial. La integración de técnicas de segmentación robustas, capaces de gestionar casos críticos de fusión pulmonar, junto con la cuantificación sistémica de 14 biomarcadores, permitió obtener una herramienta de apoyo al diagnóstico con una exactitud máxima del 96.67% en el conjunto de prueba. Este resultado demuestra que es posible transformar datos tomográficos complejos en vectores de características altamente discriminantes que facilitan la toma de decisiones clínicas objetivas y precisas.

La hipótesis planteada se validó de manera contundente, demostrando que los modelos de ML fundamentados en biomarcadores morfológicos alcanzan una precisión diagnóstica comparable, y en muchos casos superior, a los enfoques de DL. El uso de métricas como la compacidad discreta, el volumen y la tortuosidad permitió capturar la esencia de la patología sin recurrir a arquitecturas de "caja negra". Este hallazgo confirma que el conocimiento experto y la ingeniería de rasgos específicos pueden rivalizar con el DL en términos de rendimiento, manteniendo una estructura mucho más ligera y eficiente desde el punto de vista del uso de recursos computacionales.

En cuanto al desarrollo técnico, se optimizó con éxito un método de segmentación automatizada que garantizó la precisión de los contornos pulmonares incluso en cuadros severos de consolidación. La implementación de la lógica de respaldo para abordar eventos de fusión pulmonar fue un pilar esencial para la extracción de un vector de características robusto y confiable. Este vector no solo integró los biomarcadores

estructurales del árbol vascular y de las vías respiratorias, sino que se complementó con métricas de textura y estadísticas de lesión, proporcionando una caracterización sistémica de la arquitectura pulmonar que fue determinante para el éxito del entrenamiento de los modelos.

El proceso de entrenamiento y optimización mediante búsqueda en cuadrícula permitió identificar a XGBoost como el modelo líder, logrando una sensibilidad perfecta del 100%. La comparativa de rendimiento entre los seis algoritmos supervisados reveló que el sistema posee una alta capacidad de generalización, superando el 90% de exactitud en todos los casos al evaluarse en conjuntos de datos independientes. El uso de métricas rigurosas como el F1-score y un AUC-ROC superior a 0.98 confirma que la clasificación no sólo es precisa, sino equilibrada y fiable para su aplicación potencial en escenarios epidemiológicos reales.

Finalmente, se cumplió el objetivo de validar la interpretabilidad clínica a través de la XAI. Mediante la aplicación de valores SHAP e importancia por permutación, se logró rastrear cada diagnóstico hasta biomarcadores específicos, identificando al volumen vascular y la energía textural como los predictores clave alineados con la fisiopatología del COVID-19. Este paradigma no solo optimiza los recursos tecnológicos necesarios para el entrenamiento, sino que ofrece una transparencia diagnóstica superior a los modelos convencionales, garantizando que la tecnología actúe como un aliado comprensible y auditable para el personal médico en la práctica clínica.

En cuanto a las recomendaciones, se sugiere la realización de estudios multicéntricos que validen la robustez del vector de características en poblaciones con diversas características étnicas y demográficas. La inclusión de bases de datos internacionales permitiría recalibrar los hiperparámetros de modelos líderes como XGBoost para asegurar que la sensibilidad del 100% observada en este estudio se mantenga consistente ante variaciones en la calidad de la imagen y el grosor de corte de la TC. Asimismo, se recomienda el desarrollo de modelos de clasificación multiclase que

integren biomarcadores específicos para diferentes estadios de severidad y tipos de neumonía intersticial.

Una recomendación fundamental para futuros trabajos es la integración de datos clínicos y funcionales. Correlacionar los biomarcadores extraídos (como la tortuosidad vascular y la energía textural) con pruebas de función pulmonar, tales como la espirometría y la capacidad de difusión de monóxido de carbono (DLCO), proporcionaría una validación clínica más profunda. Esta sinergia entre la morfología computacional y la fisiología respiratoria permitiría no solo clasificar la presencia de la enfermedad, sino predecir el grado de discapacidad funcional del paciente, transformando el sistema de una herramienta de triaje en un modelo de pronóstico integral.

Finalmente, se propone la implementación de estudios longitudinales para monitorear la evolución de los biomarcadores morfológicos a lo largo del tiempo. Observar si la reducción del volumen vascular o la alteración de la compacidad de la vía aérea son reversibles o si persisten como secuelas permanentes permitiría a los clínicos ajustar los esquemas de rehabilitación pulmonar de manera personalizada. El uso de herramientas de XAI en estos seguimientos temporales facilitará la comprensión de los mecanismos de recuperación del parénquima, consolidando a los biomarcadores cuantitativos como el estándar de oro para la medicina de precisión en neumología.

Referencias bibliográficas

Ahmad, I. S., Dai, J., Xie, Y., & Liang, X. (2025). Deep learning models for CT image classification: A comprehensive literature review. *Quantitative Imaging in Medicine and Surgery*, 15(1), 962–1011.

Akbari, Y., Hassen, H., Al-Maadeed, S., & Zughaier, S. M. (2021). COVID-19 Lesion Segmentation Using Lung CT Scan Images: Comparative Study Based on Active Contour Models. *Applied Sciences*, 11(17), 8039.

Albataineh, Z., Aldrweesh, F., & Alzubaidi, M. A. (2024). COVID-19 CT-images diagnosis and severity assessment using machine learning algorithm. *Cluster Computing*, 27, 547–562.

Bodduluri, S., Puliyakote, A. K., Nakhmani, A., Charbonnier, J. P., Reinhardt, J. M., & Bhatt, S. P. (2021). Computed tomography–based airway surface area-to-volume ratio for phenotyping airway remodeling in chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 203(8), 970–981.

Brocki, L., & Chung, N. C. (2023). Integration of radiomics and tumor biomarkers in interpretable machine learning models. *Cancers*, 15(9), 2459.

Cai, L., Gao, J., & Zhao, D. (2020). A review of the application of deep learning in medical image classification and segmentation. *Annals of Translational Medicine*, 8(11), 713.

Campos, M. A., & Diaz, A. A. (2018). The role of computed tomography for the evaluation of lung disease in alpha-1 antitrypsin deficiency. *Chest*, 153(5), 1240–1248.

Cellina, M., et al. (2023). Artificial intelligence in lung cancer screening: The future is now. *Cancers*, 15(17), 4344.

Chae, K. J., et al. (2024). Central role of CT in management of pulmonary fibrosis. *RadioGraphics*, 44(1), e230165.

Clark, K., et al. (2013). The Cancer Imaging Archive (TCIA): Maintaining and operating a public information resource. *Journal of Digital Imaging*, 26(6), 1045–1057.

Dubois, C., et al. (2024). Deep learning in medical image analysis: introduction to underlying principles and reviewer guide using diagnostic case studies in paediatrics. *The BMJ*, 387, e076703.

Elicker, B. M., Kallianos, K. G., & Henry, T. S. (2017). The role of high-resolution computed tomography in the follow-up of diffuse lung disease. *European Respiratory Review*, 26(144), 170008.

Fan, D. P., et al. (2020). Inf-Net: Automatic COVID-19 lung infection segmentation from CT images. *IEEE Transactions on Medical Imaging*, 39(8), 2626-2637.

Guhan, B., et al. (2022). Automated system for classification of COVID-19 infection from lung CT images based on machine learning and deep learning techniques. *Scientific Reports*, 12, Artículo 17417.

Harmon, S. A., et al. (2020). Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nature Communications*, 11(1), 4080.

Hou, Y., & Navarro-Cia, M. (2022). A computationally-inexpensive strategy in CT image data augmentation for robust deep learning classification of COVID-19. *Applied Intelligence*, 52(14), 16298–16312.

Huidrom, R., Chanu, Y. J., & Singh, K. M. (2018). Automated Lung Segmentation on Computed Tomography Image for the Diagnosis of Lung Cancer. *Computación y Sistemas*, 22(3), 907-915.

Iftly, T. T., et al. (2024). Explainable lung disease classification from chest X-ray images utilizing deep learning and XAI. *arXiv*.

Labaki, W. W., et al. (2017). The role of chest computed tomography in the evaluation and management of the patient with chronic obstructive pulmonary disease. *AJRCCM*, 196(11), 1372–1379.

Lee, M.-H., et al. (2023). Deep learning methods for interpretation of pulmonary CT and X-ray images in patients with COVID-19-related lung involvement: A systematic review. *Journal of Clinical Medicine*, 12(10), 3446.

Li, M., et al. (2023). Medical image analysis using deep learning algorithms. *Frontiers in Public Health*, 11, 1273253.

Li, Y. (2023). COVID-19 diagnosis and detection based on deep learning models. *Proceedings of the 2023 International Conference on Software Engineering and Machine Learning*.

Mansoor, A., et al. (2014). A generic approach to pathological lung segmentation. *IEEE Transactions on Medical Imaging*, 33(12), 2293-2310.

Poola, R. G., et al. (2023). COVID-19 diagnosis: A comprehensive review of pre-trained deep learning models based on feature extraction algorithm. *Results in Engineering*, 18, 101020.

Pu, J., et al. (2020). Any unique image biomarkers associated with COVID-19? *European Radiology*, 30(11), 6221–6227.

Pu, J., et al. (2008). Adaptive border marching algorithm: Automatic lung segmentation on chest CT images. *Computerized Medical Imaging and Graphics*, 32(6), 452-462.

Rahaghi, F. N., et al. (2021). Quantification of arterial and venous morphologic markers in pulmonary arterial hypertension using CT imaging. *Chest*, 160(2), 657–668.

Raju, S., et al. (2017). Chest CT signs in pulmonary disease: A pictorial review. *Chest*, 151(6), 1356–1374.

Roberts, M., et al. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, 3(3), 199–217.

Rueda, S., et al. (2008). Local curvature scale: a new concept of shape description. *Proceedings of SPIE*, 6914, 69144Q.

Sarsembayeva, T., et al. (2025). Enhancing U-Net Segmentation Accuracy Through Comprehensive Data Preprocessing. *Journal of Imaging*, 11(2), 50.

Sarv Ahrabi, S., et al. (2021). An accuracy vs. complexity comparison of deep learning architectures for the detection of COVID-19 disease. *Computation*, 9(1), 3.

Singadkar, G., et al. (2021). Automatic lung segmentation for the inclusion of juxtapleural nodules and pulmonary vessels using curvature based border correction. *JKSU-CIS*, 33(8), 975–987.

Soyyigit, I. (2020). COVID-19 Chest CT Scan Dataset (Version 1) [Data set]. Harvard Dataverse.

Suganyadevi, S., et al. (2022). A review on deep learning in medical image analysis. *International Journal of Multimedia Information Retrieval*, 11(1), 19–38.

Torres Guzmán, D., et al. (2024). Evaluation of CT imaging biomarkers in the trachea and pulmonary vascular tree in patients Post-COVID-19 [Preprint]. Research Square.

Turk, M. A., & Kılıçaslan, Y. (2025). Lung image segmentation with improved U-Net, V-Net, and Seg-Net techniques. *PeerJ Computer Science*, 11, e2700.

Yan, Q., et al. (2021). COVID-19 Chest CT Image Segmentation Network by Multi-Scale Fusion and Enhancement Operations. *IEEE Transactions on Big Data*, 7(1), 13-24.

Yu, T., et al. (2024). Vascular biomarkers for pulmonary nodule malignancy: Arteries vs. veins. *Cancers*, 16(19), 3274.

Zeng, W., et al. (2024). COVID-19 chest CT image segmentation network by multi-scale fusion and enhancement operations. *IEEE/CAA Journal of Automatica Sinica*, 11(1), 1-3.

Zhang, J. D., et al. (2023). Interpretable machine learning on metabolomics data reveals biomarkers for Parkinson's disease. *ACS Central Science*, 9(6), 1035–1045.