



**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES
TÍTULO DEL TRABAJO DE TITULACIÓN**

**DESARROLLO DE UN ALGORITMO BASADO EN TRANSFER
LEARNING PARA EL ANÁLISIS DE IMÁGENES MÉDICAS DE
CÁNCER DE MAMA**

AUTOR

Carrión Vásquez, Daniel Alejandro

PROYECTO DE UNIDAD DE INTEGRACIÓN CURRICULAR

**Previo a la obtención del grado académico en
INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN**

TUTOR

Rosero Vásquez, Shendry, Ms.CC.


Santa Elena, Ecuador

Año 2025




**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**

TRIBUNAL DE SUSTENTACIÓN



Ing. José Sánchez Aquino. Mgt.
DIRECTOR DE LA CARRERA



Ing. Shendry Rosero Vásquez. Mgt.
TUTOR



Ing. Jaime Orozco Iguasnia. Mgt.
DOCENTE ESPECIALISTA



Ing. Marjorie Coronel Suárez. Mgt.
DOCENTE GUÍA UIC



**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**

CERTIFICACIÓN

Certifico que luego de haber dirigido científica y técnicamente el desarrollo y estructura final del trabajo, este cumple y se ajusta a los estándares académicos, razón por el cual apruebo en todas sus partes el presente trabajo de titulación que fue realizado en su totalidad por **CARRIÓN VÁSQUEZ DANIEL ALEJANDRO**, como requerimiento para la obtención del título de Ingeniero en Tecnologías de la Información.

La Libertad, a los 23 días del mes de junio del año 2025

TUTOR

**Shendry
Rosero**

Firmado digitalmente por Shendry
Rosero
DN: cn=Shendry Rosero
gn=Shendry Rosero c=ES Spain
l=ES Spain
e=shendry.rosero.v@gmail.com
Motivo: Soy el autor de este
documento
Ubicación:
Fecha: 2025-06-23 11:33:05:00

Ing. Shendry Rosero Vásquez. Ms.CC.



UPSE

**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES
DECLARACIÓN DE RESPONSABILIDAD**

Yo, CARRIÓN VÁSQUEZ DANIEL ALEJANDRO

DECLARO QUE:

El trabajo de Titulación, **Desarrollo de un algoritmo basado en transfer learning para el análisis de imágenes médicas de cáncer de mama**, previo a la obtención del título en Ingeniero en Tecnologías de la Información, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

La Libertad, a los 23 días del mes de junio del año 2025

EL AUTOR

Daniel Carrión V

Daniel Alejandro Carrión Vásquez



**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA**

FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

CERTIFICACIÓN DE ANTIPLAGIO

Certifico que después de revisar el documento final del trabajo de titulación denominado **Desarrollo de un algoritmo basado en transfer learning para el análisis de imágenes médicas de cáncer de mama**, presentado por el estudiante, **CARRÓN VÁSQUEZ DANIEL ALEJANDRO** fue enviado al Sistema Antiplagio, presentando un porcentaje de similitud correspondiente al 6%, por lo que se aprueba el trabajo para que continúe con el proceso de titulación.



TUTOR

**Shendry
Rosero**

Firmado digitalmente por Shendry Rosero
DN: cn=Shendry Rosero gn=Shendry Rosero c=ES Spain l=ES Spain e=shendry.rosero.v@gmail.com
Motivo: Soy el autor de este documento
Ubicación:
Fecha: 2025-06-23 11:34:05:00

Ing. Shendry Rosero Vásquez. Ms.CC.



**UNIVERSIDAD ESTATAL PENÍNSULA
DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES**

AUTORIZACIÓN

Yo, **CARRÓN VÁSQUEZ DANIEL ALEJANDRO**

Autorizo a la Universidad Estatal Península de Santa Elena, para que haga de este trabajo de titulación o parte de él, un documento disponible para su lectura consulta y procesos de investigación, según las normas de la Institución.

Cedo los derechos en línea patrimoniales del presente trabajo de titulación con fines de difusión pública, además apruebo la reproducción dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor.

Santa Elena, a los 23 días del mes de junio del año 2025

EL AUTOR

Daniel Carrion V

Daniel Alejandro Carrión Vásquez

ÍNDICE GENERAL

TITULO DEL TRABAJO DE TITULACIÓN	I
CERTIFICACIÓN	III
DECLARACIÓN DE RESPONSABILIDAD	IV
CERTIFICACIÓN DE ANTIPLAGIO	V
AUTORIZACIÓN	VI
ÍNDICE GENERAL	VII
ÍNDICE DE TABLAS	X
ÍNDICE DE FIGURAS	XI
RESUMEN	XIII
ABSTRACT	XIII
INTRODUCCIÓN	1
1. CAPÍTULO I FUNDAMENTACIÓN	2
1.1. Antecedentes	2
1.2. Descripción del Proyecto	5
1.3. Objetivos del Proyecto	7
1.4. Justificación del Proyecto	8
1.5. Alcance del Proyecto	9
1.6. Metodología del Proyecto	10
1.6.1 Metodología de Investigación	10
1.6.2 Beneficiarios del Proyecto	11
1.6.3 Variable	11
1.6.4 Análisis de recolección de datos	12
1.7. Metodología de desarrollo	14
2. CAPÍTULO II PROPUESTA	17
2.1. Marco Contextual	17
2.1.1 Diagnóstico Médico por Imágenes	17
2.1.2 Uso de la Inteligencia Artificial en el diagnóstico Médico	17
2.2. Marco Conceptual	18
2.2.1 Fundamentos del aprendizaje profundo	18
2.2.1.1 Machine Learning	18

2.2.1.2 Transfer Learning	19
2.2.2 Redes Neuronales	19
2.2.2.1 Redes Neuronales Convolucionales (CNN)	20
2.2.3 Técnicas de mejora de imágenes médicas	22
2.2.3.1 CLAHE (Contrast Limited Adaptive Histogram Equalization)	22
2.2.3.2 Gamma Correction	22
2.2.3.3 Histogram Equalization	23
2.2.4 Técnicas de optimización para modelos de Deep Learning	23
2.2.4.1 Optimizadores Adam y SGD	24
2.2.4.2 Schedulers de Tasa de Aprendizaje	24
2.2.4.3 Budgeted Training	24
2.2.4.4 Parada Temprana (Early Stopping)	24
2.2.5 Conjuntos de datos para imágenes médicas	25
2.2.5.1 CBIS-DDSM	25
2.2.6 Metodología CRISP-DM	26
2.3. Marco Teórico	27
2.3.1 Redes Neuronales Convolucionales (CNN) en el Diagnóstico Médico	27
2.3.2 Aprendizaje Profundo y Transfer Learning en Aplicaciones Médicas	27
2.3.3 Mejora de la Calidad de Imágenes Médicas para Diagnóstico asistido	28
2.4. Componente de la Propuesta	29
2.4.1. Arquitectura del sistema	29
2.4.2. Enfoque Metodológico	30
2.4.3. Obtención y preparación de datos	31
2.4.4. Configuración de Modelos	33
2.4.5. Proceso de Entrenamiento	34
2.4.6. Evaluación de modelos y Análisis de resultados	38
2.4.6.1. SGD con MultiStepLR en VGG19 y Gamma	38
2.4.6.2. SGD con MultiStepLR en EfficentNetB0 y CLAHE	39
2.4.6.3. SGD con MultiStepLR en ResNet50 y Gamma	40
2.4.6.4. SGD con StepLR en VGG19 y HE	40
2.4.6.5. SGD con StepLR en VGG19 y CLAHE	41

2.4.6.6. Matriz de Confusión: 5 mejores combinaciones	42
2.4.7. Diseño de Interfaces	46
2.4.8. Pruebas	49
CONCLUSIONES	58
RECOMENDACIONES	60
REFERENCIAS	61
ANEXOS	68

ÍNDICE DE TABLAS

Tabla 1: Resultados de las mejores Redes Neuronales con Diferentes Técnicas de Mejora de Imagen.	38
--	----

ÍNDICE DE FIGURAS

Figura 1: Diagrama de flujo de la metodología empleada en el estudio	7
Figura 2: Flujo de Transfer Learning	19
Figura 3: Gráfico del funcionamiento de una CNN	20
Figura 4: Arquitectura de ResNet50	20
Figura 5: Arquitectura VGGNet19	21
Figura 6: Arquitectura EfficientNetB0	21
Figura 7: Ajuste del contraste a una mamografía aplicando CLAHE	22
Figura 8: Ajuste de luminosidad a una mamografía aplicando Corrección de Gamma	23
Figura 9: Ajuste del histograma una mamografía aplicando HE	23
Figura 10: Diagrama de la validación empleada en Early Stopping	25
Figura 11: Ejemplo del conjunto de datos del CBIS-DDSM	26
Figura 12: Ciclo CRISP-DM	26
Figura 13: Arquitectura del sistema en Jupyter y Anaconda.	30
Figura 14: Estructura original del Dataset.	32
Figura 15: Estructuración final del Dataset.	33
Figura 16: Previsualización de imágenes después de aplicar las transformaciones.	34
Figura 17: Algoritmo de Entrenamiento y evaluación para la mejor combinación de Redes Neuronales con Transfer Learning y Técnicas de Optimización.	36
Figura 18: Diagrama de combinaciones de optimizadores y planificadores.	37
Figura 19: Resultados de todos los experimentos ejecutados en el algoritmo.	37
Figura 20: SGD con MultiStepLR.	38
Figura 21: SGD con MultiStepLR y EfficientNetB0.	39
Figura 22: SGD con MultiStepLR y ResNet50.	40

Figura 23: SGD con StepLR y VGG19	40
Figura 24: SGD con StepLR y VGG19	41
Figura 25: Matrices de Confusión de las mejores combinaciones.	42
Figura 26: Comparaciones globales de modelo, técnica de mejora, optimizadores y schedulers.	43
Figura 27: Archivos generados por el algoritmo a partir de la evaluación de cada experimento.	45
Figura 28: Algoritmo mostrando cual es la mejor combinación.	45
Figura 29: Estructura de la aplicación.	46
Figura 30: Pantalla Principal.	47
Figura 31: Pantalla de historial de predicciones.	48
Figura 32: Archivos CSV y PDF con los resultados del análisis.	49
Figura 33: Imágenes listas para los 3 tipos de prueba.	49
Figura 34: Función para verificar si la imagen es una mamografía.	50
Figura 35: Cargar imagen.	51
Figura 36: Imágenes malignas a probar.	51
Figura 37: Pantalla principal con resumen de las 50 muestras malignas analizadas.	52
Figura 38: Pantalla de resultados con el historial de predicciones.	53
Figura 39: Reporte PDF con el resumen de la primera prueba.	53
Figura 40: Imágenes benignas a probar.	54
Figura 41: Pantalla principal con resumen de las 50 muestras benignas analizadas.	55
Figura 42: Pantalla de resultados con el historial de predicciones benignas.	56
Figura 43: Reporte PDF con el resumen de la segunda prueba.	56

RESUMEN

El presente trabajo desarrolla un algoritmo unificado basado en Transfer Learning para la clasificación de imágenes médicas de cáncer de mama, usando mamografías del dataset CBIS-DDSM. Su objetivo es mostrar qué combinación de técnicas de mejora de imagen, como CLAHE, HE y Gamma Correction, junto con modelos de redes neuronales convolucionales, incluyendo ResNet50, VGG19 y EfficientNetB0 es la más adecuada para detección de malignidad a partir de mamografías. Aplicando la metodología CRISP-DM, también se examinaron las consecuencias de diferentes configuraciones de procesamiento en la clasificación de imágenes benignas y malignas.

Palabras claves: Transfer Learning, Redes Neuronales, Mamografías, Procesamiento.

ABSTRACT

This study develops a unified algorithm based on Transfer Learning for the classification of medical images of breast cancer, using mammograms from the CBIS-DDSM dataset. Its objective is to establish which combination of image enhancement techniques such as CLAHE, HE, and Gamma Correction, along with the optimization of convolutional neural network models, including ResNet50, VGG19, and EfficientNetB0 is the most suitable for detecting anomalies from mammograms. Following the CRISP-DM methodology, the effects of various preprocessing configurations on the classification of benign and malignant images also were evaluated.

Keywords: Transfer Learning, Neural Networks, Mammograms, Process.

INTRODUCCIÓN

Este proyecto, con título “Desarrollo de un algoritmo basado en Transfer Learning para el análisis de imágenes médicas de cáncer de mama”, tiene como fin hacer un algoritmo de clasificación de imágenes médicas que elija la combinación óptima de técnicas de mejora de imagen y redes neuronales convolucionales (CNNs) para la detección de casos benignos o malignos de cáncer de mama. Ya que la detección a tiempo es clave para mejorar la supervivencia, este estudio busca mejorar la estructura de los modelos de inteligencia artificial en el diagnóstico médico, ayudando a una evaluación más precisa y quitando un poco la dependencia de criterios únicamente humanos en el vistazo de mamografías.

Para el desarrollo del modelo, se emplean Transfer Learning con arquitecturas preentrenadas como ResNet50, VGG19 y EfficientNetB0, configuradas con técnicas de optimización de imágenes como CLAHE (Contrast Limited Adaptive Histogram Equalization), Histogram Equalization (HE) y Gamma Correction. Estas mejoras permiten mejorar la calidad de las imágenes médicas, ayudando a la identificación de puntos importantes en los tejidos mamarios. La metodología usada en este proyecto se basa en el enfoque CRISP-DM, reconocido en el campo del aprendizaje automático, lo que asegura una forma clara y repetible para el procesamiento y entendimiento de información.

El Capítulo 1 pone en contexto sobre la problemática del diagnóstico de cáncer de mama, mostrando a detalle la importancia de herramientas computacionales que guíen a los especialistas en la vista de mamografías. Se presentan los antecedentes del uso de inteligencia artificial en la medicina, la justificación del estudio y los objetivos que asisten al desarrollo del algoritmo. También, se explica la metodología de Transfer Learning usada y el porqué de integrar técnicas de mejora de imagen para aumentar el desempeño de los modelos de clasificación.

En el Capítulo 2, se detalla la puesta en escena del algoritmo, detallando las configuraciones de las redes neuronales y los métodos de procesamiento utiliza

dos. Se incluyen las métricas utilizadas para evaluar el rendimiento del sistema, como precisión, sensibilidad, especificidad y F1-score, con el fin de elegir qué combinación de técnicas logra los mejores resultados en la clasificación de mamografías del dataset CBIS-DDSM. Se muestran además comparaciones entre las varias configuraciones, identificando los planes más efectivos para mejorar la capacidad diagnóstica del modelo.

Este proyecto aporta al desarrollo de herramientas computacionales avanzadas para el diagnóstico médico, dando una alternativa interesante para las bases de detección de cáncer de mama mediante el uso de inteligencia artificial. Los resultados obtenidos pueden sentar las bases para futuras investigaciones en el campo de la visión por computadora aplicada a la medicina. Igualmente, se tiene en cuenta que este estudio inspire a la adopción de tecnologías basadas en aprendizaje profundo, incitando su aplicación en entornos clínicos para mejorar la toma de decisiones médicas.

1. CAPÍTULO I FUNDAMENTACIÓN

1.1. Antecedentes

El cáncer ha sido una enfermedad que ha dañado a muchas personas a lo largo de la historia. En Ecuador, es la segunda causa de muerte después de enfermedades del corazón, con 29,273 casos nuevos en 2020 [1]. Demostrando una incidencia en aumento a nivel nacional [2]. Las limitaciones de los sistemas actuales a pesar de los avances en la tecnología médica hacen que la rapidez de realizar diagnósticos precisos sea complicada. En el campo del análisis de imágenes médicas se puede observar esto, ya que el médico hace una interpretación manual que podría ser lenta y estar sujeta a errores [2].

Para darle solidez al proyecto, se realizó una entrevista a un profesional de la salud especializado en diagnósticos médicos ([Ver Anexo 1](#)), con el objetivo de recolectar información que aclare los parámetros que tiene en cuenta un médico para analizar una mamografía y en base a eso dictar el diagnóstico. El Doctor explica que la herramienta de IA debe tener algunas características a tener en cuenta para su utilidad en la clínica diaria. La capacidad de integración con sistemas de historias clínicas, la confiabilidad y precisión son algunos aspectos que el profesional destaca. El juicio humano sigue siendo preferencial para los diagnósticos clínicos y aunque el uso de la IA puede reducir errores, hay elementos que no pueden capturarse a detalle y necesitan de una vista física.

Técnicas de mejora de imagen como contraste adaptativo o corrección de gama, en los casos donde hay parecidos en las patologías distintas como quistes o tejidos hiperplásicos pueden mejorar la calidad de las imágenes diagnosticas, resaltó el entrevistado. No obstante, se subrayó la importancia de los aspectos éticos, como la transparencia en los resultados y la responsabilidad en caso de errores, lo que representa un reto para el uso de IA en el diagnóstico médico. Si bien la IA es vista como una herramienta adicional y no como un remplazo del conocimiento médico, su aplicación tiene varios desafíos relacionados con los costos, la infraestructura y la aceptación por parte del personal médico.

Dada la intensa carga de trabajo que los médicos enfrentan al analizar un elevado

número de imágenes médicas cada día, surge la necesidad de esta investigación, ya que dicha presión puede comprometer la rapidez y precisión en sus diagnósticos. Con el propósito de agilizar la detección del cáncer de mama, se propone diseñar un algoritmo que identifique la técnica de optimización de imágenes más adecuada, combinada con una red neuronal apropiada, para determinar qué conjunto de parámetros logra las predicciones con mayor exactitud. Esto se pondrá en práctica gracias al uso de técnicas de Transfer Learning y la optimización de imágenes, logrando un mayor dinamismo, sensibilidad y especificidad en la clasificación de imágenes malignas y benignas.

Por eso este proyecto se centra en la evaluación de tres arquitecturas avanzadas de CNNs: ResNet50, VGGNet19 y EfficientNetB0, mediante la técnica de Transfer Learning para el análisis de mamografías en el contexto de la detección de cáncer de mama. Utilizando el CBIS-DDSM: Breast Cancer Image Dataset, un repositorio reconocido a nivel mundial, se aplicaron diversas técnicas de optimización de imágenes, como CLAHE, HE y Gamma Correction, para mejorar la visibilidad de las características relevantes en las mamografías y aumentar la precisión diagnóstica.

A nivel mundial, en la Universidad Politécnica de Madrid, se desarrolló la tesis titulada “Clasificación de Anomalías en Mamografías mediante Redes Neuronales Convolucionales”, cuyo objetivo fue implementar un sistema automatizado que permita clasificar las anomalías presentes en mamografías mediante el uso de Redes Neuronales Convolucionales (CNN) [3]. Para mejorar la precisión diagnóstica y minimizar errores humanos, se utilizaron técnicas de Deep Learning. El entrenamiento del modelo utilizó herramientas como TensorFlow y conjuntos de datos de Kaggle. Los resultados mostraron que las CNN alcanzaron un alto nivel de exactitud en la detección de anomalías, lo que representó una notable mejora en la fiabilidad de los diagnósticos [3].

En la Universidad Distrital Francisco José de Caldas, ubicada en Bogotá, Colombia, se desarrolló en Latinoamérica una tesis titulada “Machine Learning para el Diagnóstico del Cáncer de Mama”. Este trabajo tuvo como objetivo brindar a los profesionales médicos una herramienta que permitiera realizar

diagnósticos de cáncer de mama con mayor rapidez y precisión [4]. Se usaron modelos como Support Vector Machines (SVM) y árboles de decisión (Decision Trees), logrando 100% de precisión con este último. Se creó la aplicación web BreastApp para generar informes automáticos basados en datos de pacientes y mamografías. Se concluyó que la automatización mejoró la precisión diagnóstica, redujo errores humanos y optimizó los tiempos de atención médica [4].

En Ecuador, en la Universidad Católica de Santiago de Guayaquil, se desarrolló la tesis titulada “Modelo de Machine Learning para la Detección de Tumores Cerebrales mediante Imágenes de Resonancia Magnética (MRI)”, cuyo objetivo fue crear un sistema basado en Machine Learning que permitiera detectar tumores cerebrales a través del análisis automatizado de imágenes de resonancia magnética [5]. Para ello, se utilizó una metodología de desarrollo en cascada y se implementaron Redes Neuronales Convolucionales (CNN) utilizando plataformas como TensorFlow. La tesis concluyó que el modelo mejoró la rapidez y precisión en la detección de tumores, brindando a los médicos una herramienta fiable para tomar decisiones más informadas en entornos clínicos [5].

Se examinaron tres estudios previos relacionados con la aplicación de Machine Learning en el diagnóstico médico, lo que permitió establecer una base sólida para determinar las tecnologías y mejores prácticas adaptables a esta propuesta. Al analizar dichas investigaciones, se identificaron métodos que optimizan la velocidad del diagnóstico mediante el procesamiento de imágenes médicas. Esto facilitó la comprensión de cómo los sistemas automatizados pueden asistir a los médicos en la toma de decisiones clínicas y reducir los tiempos de diagnóstico [6]. Por ello, se plantea el desarrollo de un algoritmo basado en transfer learning para el análisis de imágenes médicas, buscando agilizar la generación de resultados en escenarios clínicos e integrando técnicas de mejora de imagen.

La revisión de estos estudios subraya la necesidad apremiante de incorporar soluciones de inteligencia artificial en la práctica clínica actual. Con el avance constante de la tecnología y el aumento en el volumen de datos médicos, resulta crucial que los sistemas de salud adopten herramientas que no solo incrementen

la eficiencia, sino que también eleven la calidad de la atención al paciente [6]. Mi proyecto, “Desarrollo de un algoritmo basado en transfer learning para el análisis de imágenes médicas de cáncer de mama“ busca precisamente abordar estos desafíos mediante el desarrollo de un algoritmo que analice y determine las técnicas más efectivas de mejora de imágenes para identificar qué combinación de elementos produce las predicciones más precisas.

Esta iniciativa propone una solución innovadora y efectiva para enfrentar los desafíos del diagnóstico de cáncer de mama, aprovechando las capacidades de la inteligencia artificial y las redes neuronales profundas. El objetivo es desarrollar sistemas que apoyen a los profesionales médicos, aliviando su carga laboral y mejorando la precisión diagnóstica. El uso de estos modelos y técnicas busca reducir los tiempos de análisis y minimizar cargas de trabajo, proporcionando una herramienta poderosa para el diagnóstico temprano de enfermedades críticas.

1.2. Descripción del Proyecto

El proyecto se centra en el desarrollo de un algoritmo unificado basado en Transfer Learning para mejorar la rapidez y precisión en la clasificación de imágenes médicas, específicamente mamografías, a través de la implementación de técnicas de mejora de imagen y la optimización de modelos de redes neuronales convolucionales (CNN). El sistema se diseñará para identificar y evaluar las combinaciones más efectivas de técnicas de mejora de imagen, como CLAHE, Histogram Equalization (HE) y Corrección de Gamma, con el objetivo de aumentar la capacidad diagnóstica del modelo en la clasificación de imágenes benignas y malignas. Las fases del proyecto siguen la estructura de CRISP-DM, un marco ampliamente utilizado en minería de datos y aprendizaje automático [7].

Estas Fases se dividen en 5 y son:

Fase 1: Preparación de Datos

- **A1.** Recolección y procesamiento de las imágenes de mamografías del dataset CBIS-DDSM.
- **A2.** Se transforman las imágenes para incluir redimensionamiento, conversión a tensor y normalización y desviaciones estándar.

- **A3.** Data set dividido en conjunto de entrenamiento (80%) y conjunto de prueba (20%).

Fase 2: Configuración de Modelos

- **A1.** Evaluación de las tres arquitecturas de CNN: ResNet50, VGGNet19 y EfficientNetB0.
- **A2.** Se preentrenan para realizar una clasificación binaria entre imágenes benignas y malignas.
- **A3.** Ajuste de la última capa de cada modelo maximizando la eficiencia y precisión.
- **A4.** Implementación de distintas configuraciones para cada modelo, ajustando hiperparámetros relevantes.

Fase 3: Proceso de Entrenamiento

- **A1.** Se define un número total de 3 épocas (epochs) para el entrenamiento, con la opción de ajuste global.
- **A2.** Se emplean técnicas como Budgeted Training y Parada temprana para optimizar los recursos computacionales y alcanzar un criterio de precisión para evitar sobreajuste.
- **A3.** Se utilizan los optimizadores Adam y SGD con tasas de aprendizaje 0.001 y 0.01 relativamente.
- **A4.** Convergencia rápida y mejora de la precisión y estabilidad en la optimización.

Fase 4: Evaluación de Modelos

- **A1.** Aplicación de métricas de evaluación como precisión del modelo, generación de matrices de confusión y reportes de clasificación para las técnicas de mejora de imagen.
- **A2.** Uso de validación cruzada para asegurar la robustez de los resultados obtenidos.
- **A3.** Técnicas aplicadas: CLAHE, HE (Histogram Equalization), Gamma Correction.
- **A4.** Comparación del impacto de cada técnica en la precisión de la

clasificación de imágenes benignas y malignas.

Fase 5: Visualización y Almacenamiento de Resultados

- **A1.** Se crean gráficos y visualizaciones de matrices de confusión para una interpretación clara del rendimiento del modelo.
- **A2.** Los modelos entrenados y los informes de evaluación se almacenan en disco.
- **A3.** Documentación completa del proceso de investigación para análisis y verificaciones futuras.

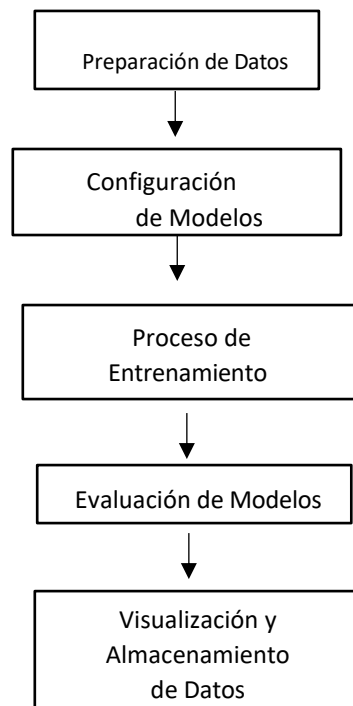


Figura 1: Diagrama de flujo de la metodología empleada en el estudio

1.3. Objetivos del Proyecto

Objetivo General:

Desarrollar un algoritmo unificado basado en Transfer Learning que facilite la implementación de mejoras gráficas en imágenes médicas, así como la integración de schedulers y optimizadores.

Objetivos específicos:

- Analizar el rendimiento de diferentes arquitecturas de redes neuronales convolucionales mediante Transfer Learning, en la clasificación de imágenes de mamografías del dataset CBIS-DDSM.
- Integrar técnicas de mejoras gráficas para aumentar la calidad de las imágenes diagnósticas, facilitando la interpretación médica y reduciendo la necesidad de intervenciones invasivas.
- Determinar la combinación óptima de técnicas que genere resultados concluyentes y optimizados a través del algoritmo, maximizando la precisión en la clasificación.

1.4. Justificación del Proyecto

En los últimos diez años, los desarrollos en inteligencia artificial y aprendizaje profundo han potenciado notablemente la capacidad del Machine Learning (ML) para procesar grandes cantidades de datos, incluidas imágenes médicas [8]. Estas tecnologías ofrecen un enorme potencial para incrementar la precisión y acelerar los diagnósticos médicos, un aspecto especialmente vital en la detección precoz de enfermedades como el cáncer de mama [7]. No obstante, persisten desafíos en el ámbito médico para integrar sistemas de inteligencia artificial de manera efectiva en los procesos clínicos diarios [9].

Este proyecto está diseñado para convertirse en una herramienta que transforme la carga de trabajo médico en una oportunidad de optimización, reduciendo los tiempos de diagnóstico y facilitando la toma de decisiones médicas de manera más dinámica y fundamentada. El algoritmo basado en Transfer Learning integra técnicas avanzadas de mejora gráfica, incrementando la calidad de las imágenes médicas y minimizando la variabilidad en los resultados, lo cual disminuye el riesgo de errores en la interpretación.

Al usar modelos de redes neuronales convolucionales (CNN) entrenados con técnicas avanzadas de inteligencia artificial, se garantiza una clasificación de imágenes precisa y consistente, proporcionando a los profesionales de la salud una herramienta complementaria que disminuye la dependencia de juicios subjetivos

[9]. Esta solución no solo agiliza el proceso diagnóstico, sino que también mejora el acceso a una atención médica eficiente, permitiendo a los pacientes recibir tratamientos oportunos y aumentando sus probabilidades de recuperación [9].

Este proyecto está alineado “Plan de Desarrollo para el Nuevo Ecuador”, especificado en los siguientes ejes:

Eje 1.- Eje Social

Objetivo 1.- Mejorar las condiciones de vida de la población de forma integral, promoviendo el acceso equitativo a salud, vivienda y bienestar social [10].

Política 1.3.- Mejorar la prestación de los servicios de salud de manera integral, mediante la promoción, prevención, atención primaria, tratamiento, rehabilitación y cuidados paliativos, con talento humano suficiente y fortalecido, enfatizando la atención a grupos prioritarios y todos aquellos en situación de vulnerabilidad [10].

Objetivo 5.- Fomentar de manera sustentable la producción mejorando los niveles de productividad [10].

Al mejorar la eficiencia y precisión del diagnóstico médico en instituciones de salud, se apoya el desarrollo de un sistema de salud más accesible y de calidad, promoviendo el bienestar de los ciudadanos y reduciendo las disparidades en el acceso a diagnósticos avanzados.

1.5. Alcance del Proyecto

El alcance de esta investigación es analítico debido a que busca identificar y evaluar las combinaciones más efectivas de técnicas de mejoras de imagen y modelos de redes neuronales para mejorar la rapidez y precisión en el diagnóstico de cáncer de mama mediante mamografías. Los resultados esperados de este estudio incluyen características como la identificación de las configuraciones óptimas para lograr una alta precisión, sensibilidad y especificidad en la clasificación de imágenes malignas y benignas. Estos hallazgos serán el fundamento para concluir cuál combinación de técnicas y modelos es la más efectiva para el contexto clínico y de investigación médica [11].

Mediante este enfoque, se busca generar datos detallados sobre el impacto de cada combinación en el desempeño del algoritmo, facilitando una comprensión más profunda de las relaciones causales entre las técnicas de mejora de imagen y el tipo de red neuronal utilizada. Este análisis no solo proporcionará resultados detallados sobre la eficacia del modelo en esta tarea, sino que también ofrecerá conocimientos valiosos para futuras investigaciones en diagnóstico asistido por inteligencia artificial, destacando la influencia de cada técnica en la precisión diagnóstica. Así, el método estará condicionado por un enfoque analítico que incluirá una evaluación cuantitativa exhaustiva de los resultados, junto con una comparación de métricas entre las distintas combinaciones analizadas [12].

El proyecto culminará con un conjunto de configuraciones validadas que podrán ser utilizadas como referencia para investigaciones futuras o aplicaciones clínicas. Esta iniciativa contribuirá al campo del diagnóstico médico asistido, proporcionando a investigadores y desarrolladores de inteligencia artificial una base sólida para el diseño y perfeccionamiento de sistemas diagnósticos. De esta manera, el enfoque analítico asegura que los métodos de validación y análisis empleados en el estudio se centren en entender las interacciones entre las variables y su impacto en el desempeño del sistema [13].

1.6. Metodología del Proyecto

1.6.1 Metodología de Investigación

Este proyecto sigue una metodología de búsqueda sistemática de hiperparámetros basada en un método de investigación hipotético-deductivo, ya que el desarrollo del algoritmo se basa en la formulación de hipótesis específicas sobre cuáles combinaciones de técnicas de mejora de imagen y modelos de redes neuronales optimizarán la precisión en la clasificación de imágenes médicas. A través del método hipotético-deductivo, se organiza el análisis de forma sistemática y coherente, planteando hipótesis que serán evaluadas mediante experimentos para determinar su influencia en la precisión diagnóstica, buscando así corroborar o descartar los supuestos iniciales [14].

El enfoque cuantitativo resulta idóneo para este proyecto, ya que permite medir y analizar variables de forma precisa y objetiva, un aspecto fundamental en investigaciones que evalúan la eficacia de diferentes técnicas o modelos basados en datos numéricos [16]. Este proceso de validación y optimización es esencial, pues posibilita el perfeccionamiento continuo del algoritmo según los resultados obtenidos, consolidando una base robusta que respalda la efectividad del modelo y fomenta el progreso del conocimiento en el campo del diagnóstico médico asistido por inteligencia artificial [15].

1.6.2 Beneficiarios del Proyecto

La investigación no se orienta a una institución de salud específica, sino que tiene fines experimentales y de optimización de algoritmos, con el objetivo de contribuir al conocimiento sobre el diagnóstico médico asistido por IA [13]. Sin embargo, el resultado del rendimiento del algoritmo podría ser evaluado por médicos y técnicos encargados de la interpretación de resultados en radiografías.

Este proyecto se enfoca en el desarrollo de un algoritmo para la clasificación de mamografías, probándolo en un entorno experimental con el dataset CBIS-DDSM. Se experimentará con diversas técnicas de mejora de imagen (como CLAHE, Histogram Equalization y Gamma Correction) y configuraciones de redes neuronales avanzadas, bajo la orientación y asesoramiento del Dr. especialista en diagnóstico médico, con quien se ha realizado una entrevista para comprender mejor las necesidades del análisis de imágenes en un entorno clínico.

1.6.3 Variable

En este caso, la variable principal es la efectividad de la combinación de técnicas de mejora de imagen y arquitecturas de redes neuronales en la precisión y rapidez para proporcionar resultados diagnósticos. Este planteamiento facilita el establecimiento de métricas claras, como precisión, sensibilidad y especificidad, cruciales para medir el desempeño del algoritmo e identificar la combinación que

genera las predicciones más exactas al clasificar mamografías como benignas o malignas [12].

Los datos estadísticos obtenidos permitirán analizar la efectividad de cada configuración y establecer la combinación más adecuada para mejorar la precisión diagnóstica. Al final, el enfoque cuantitativo de este estudio permite extraer conclusiones basadas en datos objetivos y replicables, lo que fortalece la validez de los hallazgos y sienta una base robusta para investigaciones futuras en diagnóstico médico asistido por inteligencia artificial [12].

1.6.4 Análisis de recolección de datos

Población y muestra

La población de estudio de esta investigación está formada por un total de 3.568 imágenes médicas provenientes del dataset CBIS-DDSM (Curated Breast Imaging Subset of DDSM), un conjunto de mamografías específicamente recopilado para estudios de detección de cáncer de mama. Dado que este dataset ha sido cuidadosamente curado para fines de investigación médica, refleja con precisión los patrones patológicos presentes en mamografías, lo que lo convierte en una herramienta ideal para desarrollar y evaluar modelos de aprendizaje automático en el diagnóstico de imágenes médicas [17]. La representatividad de los casos clasificados en este conjunto de datos facilita un análisis preciso del desempeño de los modelos de inteligencia artificial.

Para el desarrollo experimental, se optó por emplear la totalidad del dataset CBIS-DDSM, descartando el uso de técnicas de muestreo. El conjunto completo de datos se distribuirá en un 80% destinado al entrenamiento del modelo y un 20% para las fases de validación y prueba. Esta estrategia maximiza la utilización de los datos disponibles, asegurando un conjunto representativo para el aprendizaje del modelo y una evaluación robusta de su desempeño. La división en estas proporciones es una práctica habitual en proyectos de inteligencia artificial, ya que permite optimizar el modelo al aprovechar la abundancia de información durante el entrenamiento y garantizar la fiabilidad de los datos en la evaluación.

Técnicas e instrumentos de recolección de datos

En el proyecto se utilizarán diversas técnicas e instrumentos para la recolección de datos, incluyendo entrevistas a expertos, análisis de datasets y revisión de bibliografía especializada. Estos instrumentos proporcionan una base sólida de datos y conocimientos esenciales para el desarrollo del modelo del algoritmo basado en Transfer Learning.

En primer lugar, se realizó una entrevista ([Ver Anexo 1](#)) un profesional de la salud, especialista en diagnóstico médico, con el propósito de entender mejor los criterios clínicos que influyen en el análisis de imágenes médicas, como las mamografías. Esta entrevista permitió obtener información cualitativa relevante sobre las características fundamentales que debe tener una herramienta de IA para ser útil en la práctica clínica. Además, se abordaron las necesidades y desafíos en el análisis de imágenes diagnósticas, las cuales sirvieron de guía para el desarrollo de los objetivos del proyecto y la optimización del algoritmo.

El análisis cuantitativo de este proyecto se basa en el dataset CBIS-DDSM (Curated Breast Imaging Subset of DDSM), que incluye aproximadamente 3.568 imágenes de mamografías clasificadas como normales, benignas o malignas. Este conjunto de datos proporciona una base robusta para entrenar y evaluar modelos de redes neuronales convolucionales (CNN), siendo ideal para iniciativas de clasificación en el diagnóstico de cáncer de mama [6]. Las imágenes, que contienen regiones de interés y diagnósticos patológicos, permiten una evaluación precisa del algoritmo, consolidando al CBIS-DDSM como una fuente primaria que garantiza un análisis sólido y reproducible en la clasificación de imágenes [13].

Para sustentar el desarrollo del proyecto, se consultaron diversas fuentes bibliográficas relevantes en los campos de inteligencia artificial y diagnóstico médico. Entre ellas destaca el libro *Inteligencia artificial: conceptos básicos y aplicaciones en el desarrollo*, que fundamenta el diseño de la investigación, junto con estudios especializados sobre el empleo de redes neuronales y técnicas de Transfer Learning en el análisis de imágenes médicas [18]. Estas referencias proporcionan el marco teórico y metodológico del proyecto, además de respaldar el uso de técnicas avanzadas de mejora de imagen, como CLAHE y Corrección de

Gamma, que se implementarán durante la fase experimental [17].

Estos instrumentos y técnicas de recolección de datos aseguran que el desarrollo del proyecto esté fundamentado en una comprensión exhaustiva del contexto clínico y técnico, facilitando un enfoque integral en el diseño del modelo de diagnóstico asistido por inteligencia artificial. Al considerar tanto estos aspectos, se maximiza la efectividad y relevancia del sistema, permitiendo que el modelo se adapte a las necesidades y desafíos específicos del futuro entorno médico.

1.7. Metodología de desarrollo

Este proyecto se estructura bajo el modelo CRISP-DM, un marco ampliamente reconocido en iniciativas de Machine Learning e inteligencia artificial [18]. La flexibilidad y el carácter iterativo de CRISP-DM lo hacen ideal para este desarrollo, ya que permite organizar de manera eficiente las tareas necesarias para entrenar y perfeccionar modelos de Transfer Learning aplicados al análisis de imágenes médicas [19]. Todas las fases del proyecto están alineadas con las etapas del ciclo CRISP-DM, garantizando un proceso de desarrollo metódico. En el contexto del diagnóstico de cáncer de mama, este enfoque promueve la precisión y rapidez en la interpretación de imágenes, proporcionando una guía clara para la preparación de datos, entrenamiento, evaluación y registro de resultados [7].

La metodología consta de cinco fases, cada una orientada a desarrollar y optimizar un aspecto fundamental del sistema de diagnóstico médico. A continuación, se describen las fases en detalle:

- **Fase 1: Preparación de Datos**

La primera fase, se centra en la obtención y preparación de las imágenes de mamografías provenientes del dataset CBIS-DDSM. Para garantizar la uniformidad y calidad de los datos antes del entrenamiento, las imágenes se someten a varias transformaciones, como normalización, conversión a tensor y redimensionamiento. Específicamente, cada imagen se ajusta a una resolución de 256x256 píxeles, seguida de un recorte centrado a 224x224 píxeles. Posteriormente, las imágenes se normalizan utilizando medias y desviaciones

estándar específicas para estabilizar el aprendizaje del modelo. Finalmente, el dataset se divide en 80% para entrenamiento y 20% para prueba, asegurando un entrenamiento equilibrado y representativo.

- **Fase 2: Configuración de Modelos**

En la segunda fase, se procede a la implementación de tres arquitecturas de redes neuronales convolucionales (CNN) preentrenadas: VGGNet19, ResNet50 y EfficientNetB0. Estas redes se configuran mediante Transfer Learning para realizar una clasificación binaria, distinguiendo entre imágenes benignas y malignas. La última capa de cada modelo se modifica para adaptarse a la tarea específica de diagnóstico, y se definen hiperparámetros clave, como funciones de pérdida y tasas de aprendizaje, personalizados para cada arquitectura. Además, se realizan pruebas iniciales para validar la correcta configuración de los modelos. Al concluir esta fase, los modelos están preparados para el entrenamiento, optimizados para maximizar su rendimiento en el análisis de imágenes médicas.

- **Fase 3: Proceso de Entrenamiento**

En la tercera etapa, se procede al entrenamiento de los modelos, empleando técnicas avanzadas de optimización. Se establece un máximo de tres épocas (epochs) para el entrenamiento, con la opción de interrumpirlo anticipadamente si se alcanza un criterio de precisión, utilizando métodos como la parada temprana (Early Stopping) para prevenir el sobreajuste. Con el objetivo de optimizar los recursos computacionales limitados, se usa el budgeted training, que maximiza la eficiencia en tiempo y procesamiento durante el entrenamiento. Para garantizar una convergencia más balanceada, se utilizan los optimizadores Adam y SGD, ambos con una tasa de aprendizaje de 0.001. Al concluir esta fase, los modelos estarán entrenados y optimizados, listos para su evaluación.

- **Fase 4: Evaluación de Modelos**

Durante la cuarta fase, se lleva a cabo una evaluación detallada de los modelos entrenados. Se analiza el desempeño de cada modelo en combinación con las técnicas de mejora de imagen aplicadas, como CLAHE, Histogram Equalization y Gamma Correction, mediante el cálculo de métricas esenciales, incluyendo

precisión, sensibilidad, especificidad y la elaboración de matrices de confusión. Para garantizar la fiabilidad de los resultados, se emplea validación cruzada, y se compara la eficacia de cada modelo en la clasificación precisa de imágenes como benignas o malignas. Este análisis permitirá determinar qué combinación de técnica de mejora de imagen y modelo de red neuronal convolucional (CNN) ofrece los mejores resultados en términos de exactitud diagnóstica.

- **Fase 5: Visualización y Almacenamiento de Resultados**

En la fase final, se crean visualizaciones de las matrices de confusión para proporcionar una representación gráfica clara del desempeño de clasificación de cada modelo, facilitando una interpretación directa de su eficacia en la tarea diagnóstica. Los modelos entrenados, junto con los informes de evaluación, se guardan en disco, asegurando una documentación estructurada y completa del proceso de desarrollo. Este almacenamiento permite análisis futuros y la posibilidad de emplear los modelos en pruebas clínicas o como base para optimizaciones posteriores.

La metodología CRISP-DM resulta especialmente usable en este ambiente, porque la preparación y calidad de los datos son importantes para que las redes neuronales convolucionales (CNN) den predicciones precisas [20]. También, este enfoque inspira una orientación hacia resultados prácticos, asegurando que los sistemas desarrollados no solo muestren avances técnicos, sino que sean aplicables y valiosos en entornos clínicos.

2. CAPÍTULO II PROPUESTA

2.1. Marco Contextual

2.1.1 Diagnóstico Médico por Imágenes

En la medicina moderna, el diagnóstico por imágenes es una utilidad importante, porque ayuda a los especialistas a evaluar internamente los esquemas del cuerpo humano sin hacer procedimientos físicos apresurados. La radiografía, tomografía computarizada (CT), resonancia magnética (MRI) y mamografía se muestran como las principales técnicas para la detección a tiempo de cáncer de mama [21]. La disminución de muertes en esta enfermedad es una consecuencia de la detección pronta de cáncer de mama, porque se ven lesiones en estados prematuros [21]. Sin embargo, el análisis de estas imágenes depende mucho de la experiencia del radiólogo, algo que puede generar aspectos cambiantes en los diagnósticos y posibles errores en la detección de cosas raras.

Con el avance de la Inteligencia artificial y el aprendizaje profundo, se han creado modelos con el volumen de guía a los doctores en la evaluación de imágenes médicas, mejorando la precisión y reduciendo el tiempo necesario a emitir un diagnóstico [2]. Este proyecto trata de acercarse a este aspecto usando el estudio de redes neuronales convolucionales (CNNs) y técnicas de mejora de imagen para mejorar la detección de cáncer de mama.

2.1.2 Uso de la Inteligencia Artificial en el diagnóstico Médico

Los adelantos importantes en el procesamiento y análisis de imágenes diagnósticas se han permitido gracias a la unión de la Inteligencia Artificial en el campo Médico [6]. La precisión en la detección de enfermedades no solo ha sido una mejora de la Inteligencia Artificial, también ha aminorado los tiempos de evaluación para disminuir falsos positivos o negativos [22].

La inteligencia artificial se ha usado para ver microcalcificaciones y masas raras, permitiendo una examinación más rápida y sin juicio, dentro del entorno de la mamografía [22]. Combinando modelos de aprendizaje profundo y técnicas de mejora de imagen se puede aumentar la calidad de las mamografías, aminorando el análisis por parte de los radiólogos según recientes estudios[23].

En cuanto a la regulación del uso de bases de datos para investigaciones de Machine Learning, Transfer Learning o Deep Learning utilizadas en la medicina, la adopción de normativas éticas que protejan la privacidad los datos es elemental tenerlas en cuenta. El dataset elegido para la prueba es el CBIS-DDSM (Curated Breast Imaging Subset of DDSM), el que con los principios de acceso abierto y cumpliendo con los estándares de anonimización de datos está disponible para estudios científicos [24]. Esta base de datos tiene las mamografías con etiquetas puestas por especialistas permiten evaluar el rendimiento de modelos dentro de ambientes controlados y aprobados por la comunidad científica.

El proyecto se construye con la guía de un especialista en diagnóstico médico, quien ha dado información clave sobre las necesidades en la interpretación de mamografías en ambientes clínicos. Es importante tener en cuenta que la investigación en esta área es determinante para seguir mejorando la detección pronta del cáncer de mama mediante Inteligencia Artificial, incentivando el uso de tecnologías nuevas en el campo de la salud para mejorar la toma de decisiones clínicas y aumentar los procesos de análisis de mamografías [6].

2.2. Marco Conceptual

2.2.1 Fundamentos del aprendizaje profundo

El aprendizaje profundo (Deep Learning) es una categoría inferior del aprendizaje automático (Machine Learning) que usa redes neuronales para quitar características junto con la realización de predicciones en datos difíciles [25]. El alcance de estas redes para crear relaciones no lineales junto con el aprendizaje de copias jerárquicas iniciando con datos de entrada, han definido el éxito de esta ciencia [25].

2.2.1.1 Machine Learning

Machine Learning es familia de la IA y se enfoca en la construcción de algoritmos y modelos computacionales listos para aprender muestras y tomar decisiones con base en información no programada antes. Mientras más información se trabaje los modelos de ML pueden mejorar su rendimiento, ayudando en la automatización de tareas difíciles como el reconocimiento de imágenes, detección de enfermedades o

procesamiento de lenguaje. La destreza de aplicar lo aprendido a nuevos ambientes marcan el éxito de los modelos porque no hay intervención humana [26]. Existen varias técnicas de ML, como el aprendizaje supervisado, no supervisado y por refuerzo, todas con aplicaciones individuales en el análisis de datos y la salud con IA [26].

2.2.1.2 Transfer Learning

Transfer Learning es un enfoque de aprendizaje profundo que ayuda a usar en una actividad parecida un modelo anteriormente entrenado. Los modelos pueden utilizar su conocimiento previo si se aplica esta técnica, aumentando la precisión y aminorando el tiempo de entrenamiento. Cuando los datos son limitados o difíciles de conseguir, como las imágenes médicas, Transfer Learning es útil, ya que no hay que entrenar desde 0 los modelos y hace más sencilla una adaptación [27].

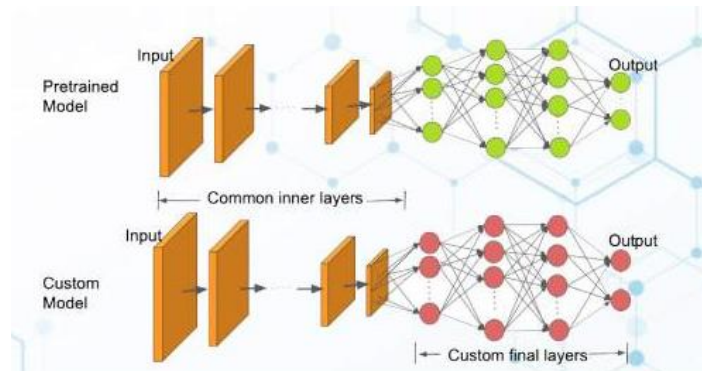


Figura 2: Flujo de Transfer Learning

2.2.2 Redes Neuronales

Las Redes Neuronales con base en el cerebro humano, son un conjunto de modelos computacionales[27]. Las neuronas artificiales son parte de su anatomía y se agrupan en capas bien conectadas que trabajan la información por operaciones matemáticas. La separación de imágenes, análisis de texto o reconocimiento de voz, son tareas que hacen parte su composición gracias al aprendizaje de representaciones jerárquicas de los datos [27]. Con algoritmos de optimización y técnicas de ajuste de pesos sinápticos, las ANN pueden aumentar el desempeño con la exposición a nuevos datos.

2.2.2.1 Redes Neuronales Convolucionales (CNN)

Las redes neuronales convolucionales (CNN) están cambiadas para el procesamiento y separación de imágenes, siendo un tipo de arquitectura de redes neuronales. Las CNN para el análisis de ejemplos difíciles, como patrones de imágenes médicas, son una gran opción. Sin necesidad de un preparamiento humano estas pueden aprender características desde los datos [11].

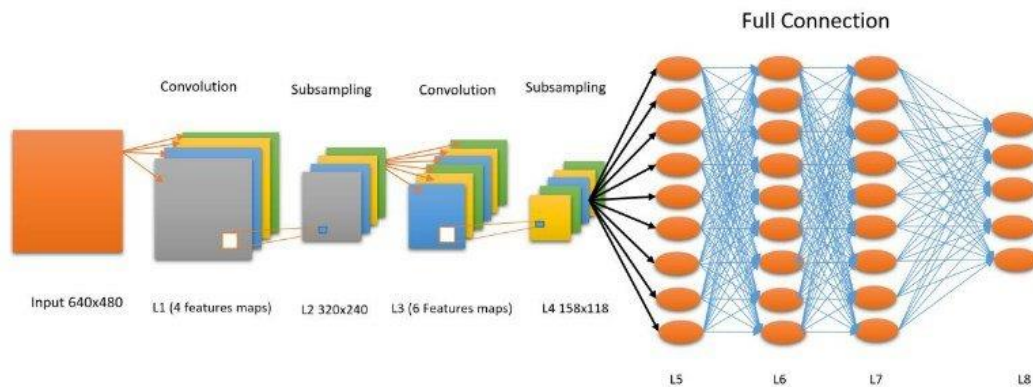


Figura 3: Gráfico del funcionamiento de una CNN

2.2.2.1.2 ResNet50

ResNet50 es una arquitectura de red neurona convolucional, tiene 50 capas y usa bloques sobrantes para evitar que en redes muy profundas haya degradación de la precisión [29]. En tareas de clasificación y detección, esta arquitectura ha mostrado gran desempeño, además de utilizarse en trabajos de diagnóstico por imágenes con Transfer Learning [29].

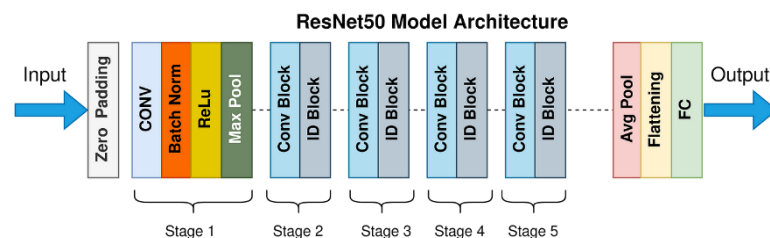


Figura 4: Arquitectura de ResNet50

2.2.2.1.3 VGG19

VGGNet19 es una arquitectura de red neuronal profunda, creada en el Visual Geometry Group, tiene 19 capas [30]. El diseño ajustable y a profundidad de la red la hacen una opción famosa para tareas de Transfer Learning, más para el rendimiento de clasificación de imágenes médicas mejorado que permite mayor ajuste de capacidad [31].

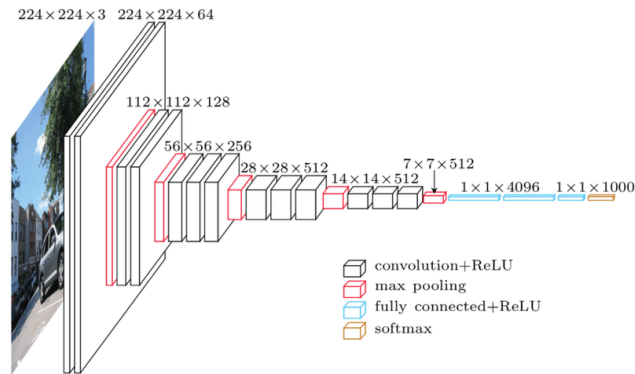


Figura 5: Arquitectura VGGNet19

2.2.2.1.4 EfficientNetB0

EfficientNetB0 es un modelo de red neuronal convolucional mejorado para obtener una alta precisión y eficiencia computacional gracias a un escalado igualado de resolución de la red, ancho y profundidad [32]. Este modelo es muy usado en diagnósticos médicos por su disposición de ofrecer precisión en la clasificación con recursos computacionales limitados [32].



Figura 6: Arquitectura EfficientNetB0

2.2.3 Técnicas de mejora de imágenes médicas

En el diagnóstico por imágenes, la mejora de imágenes es importante porque se aumenta la vista de estructuras suaves y ayuda a resaltar características a tener en cuenta [4]. Desde procesamiento digital hasta ciclos basados en inteligencia artificial se consiente el aumento de la calidad de las imágenes médicas adaptativamente. La mejora en la precisión de detección de enfermedades se ha mejorado gracias a la aplicación de estas técnicas en ambientes de apoyo [4].

2.2.3.1 CLAHE (Contrast Limited Adaptive Histogram Equalization)

CLAHE es una técnica de mejora de imagen que en regiones pequeñas ajusta el contraste de manera adaptativa. CLAHE ayuda a mejorar la visibilidad de estructuras delicadas cuando se aplica a la medicina, además de indicar posibles cosas raras [33]. Un ajuste local de contraste es ideal para señalar cosas específicas en imágenes complejas, facilitando un diagnóstico preciso y confiable [33].

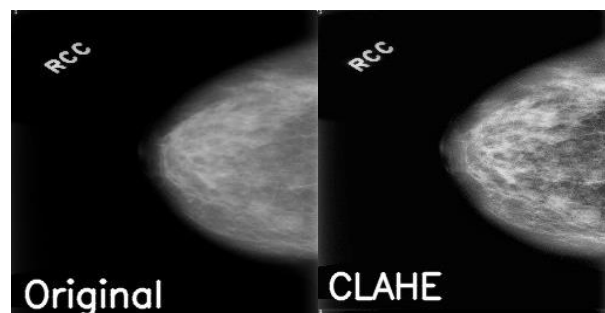


Figura 7: Ajuste del contraste a una mamografía aplicando CLAHE

2.2.3.2 Gamma Correction

Gamma Correction incrementa la vista en partes oscuras o iluminadas gracias al ajuste no lineal de valores de luminosidad en una imagen. En ambiente médico para facilitar el trabajo a los especialistas, esta técnica es usable para destacar cosas específicas en imágenes diagnósticas [34]. Además, Corrección de Gamma para que una imagen tenga calidad adecuada para su análisis, nivela las diferencias de iluminación [34].

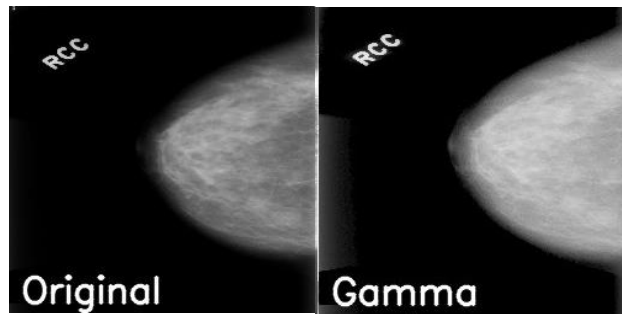


Figura 8: Ajuste de luminosidad a una mamografía aplicando Corrección de Gamma

2.2.3.3 Histogram Equalization

Histogram Equalization (HE) es una técnica de mejora de imágenes que reasigna los niveles de intensidad de los píxeles en toda la imagen con el fin de mejorar el contraste [35]. Se alcanza una mayor organización del histograma, aumentando el rango de intensidades y mostrando cosas que pueden ser difíciles de visualizar en imágenes médicas. HE ha sido útil para mejorar claridad de radiografías, tomografías y mamografías en diagnóstico por imágenes [35]. Por su enfoque global, HE podría hacer sobreajustes para algunas zonas y para lograr mejores puntajes se suele combinar con técnicas como CLAHE [35].

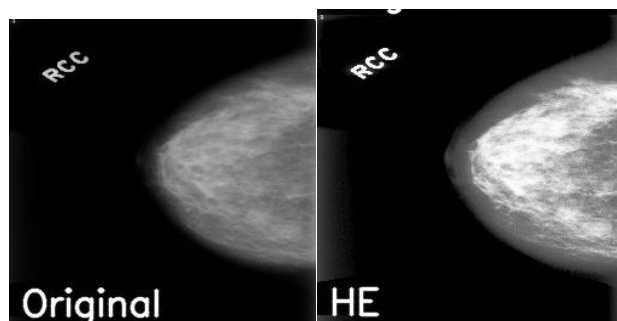


Figura 9: Ajuste del histograma a una mamografía aplicando HE

2.2.4 Técnicas de optimización para modelos de Deep Learning

Para entrenar modelos de aprendizaje profundo (Deep Learning) de manera adecuada se debe aplicar técnicas modernas de mejoría en tareas de clasificación de imágenes médicas que adecuen la convergencia del rendimiento [30]. Las técnicas van desde la selección de optimizadores adecuados, ajustes de tasa de aprendizaje y el uso de estrategias de regularización. Aplicando dichas estrategias el modelo puede adelantar su entrenamiento y mejorar la forma de generalización

de los modelos [20].

2.2.4.1 Optimizadores Adam y SGD

Adam usa dos métodos, el de decaimiento de gradiente estocástico (SGD) y técnicas de costumbre de tasas de aprendizaje, mejorando la convergencia en el entrenamiento de redes profundas. Adam y SGD son optimizadores importantes en modelos de Transfer Learning, dejando en imágenes médicas configurar los parámetros del modelo para tener mejores resultados [36].

2.2.4.2 Schedulers de Tasa de Aprendizaje

Los schedulers de tasa de aprendizaje con base en el entrenamiento, permiten ajustar la tasa de aprendizaje del modelo. En transfer Learning esto es importante debido a que se permite un ajuste en aumento del modelo y así mejorar la estabilidad y precisión [37]. Con la adaptación gradual de la tasa de aprendizaje, los schedulers no dejan que exista el sobreajuste y el modelo con el conjunto de datos se configura de manera aceptable [37].

2.2.4.3 Budgeted Training

Budgeted Training es una técnica usada cuando hay pocos recursos computacionales al momento de entrenar modelos de aprendizaje profundo [38]. Los hiperparámetros se configuran aplicando métodos como parada temprana para tener en cuenta el tiempo y recursos de procesamiento, ayudando a los modelos a entrenar de manera adecuada [38].

2.2.4.4 Parada Temprana (Early Stopping)

Parada Temprana o Early Stopping es una técnica regularizada que evita el sobreajuste cuando se entrena modelos de aprendizaje profundo [39]. La técnica actúa cuando el rendimiento del modelo ya no presenta mejoras después de un número de épocas seguidas y sobre un conjunto de validación detiene el entrenamiento automáticamente. Si se evita el sobreajuste de parámetros es posible que exista un actuar bueno del modelo sobre los datos de entrenamiento porque ya se ha alcanzado su mejor rendimiento en validación y habrá nuevos datos que no se vieron por la mayor capacidad de generalización [39].

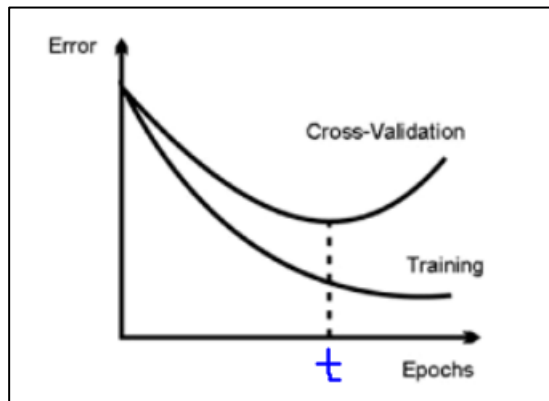


Figura 10: Diagrama de la validación empleada en Early Stopping

2.2.5 Conjuntos de datos para imágenes médicas

La utilización de bases de datos especiales para entrenar y evaluar modelos de detección de imágenes es un parte muy importante de cualquier proyecto. Hay diferentes tipos de conjuntos de datos creados para hacer herramientas de diagnóstico de enfermedades combinado con IA y entrenamiento de modelos de aprendizaje automático [40].

2.2.5.1 CBIS-DDSM

El CBIS-DDSM (Curated Breast Imaging Subset of DDSM) es un dataset de mamografías accesible en la web usando para entrenamiento y validación de inteligencia artificial para análisis de cáncer [41]. Imágenes de alta calidad con sus etiquetas de anomalías son algunas características que incluyen el conjunto de datos, esto ayuda a ver ejemplos en el aprendizaje profundo para el diagnóstico de cáncer [41]. CBIS-DDSM ha sido de las bases de datos más importantes en investigación biomédica para una predicción prematura de cáncer de mama por el desarrollo de algoritmos de diagnóstico [41].

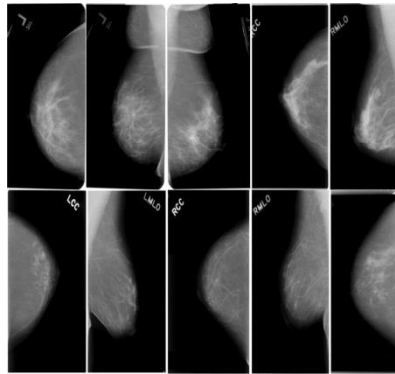


Figura 11: Ejemplo del conjunto de datos del CBIS-DDSM

2.2.6 Metodología CRISP-DM

CRISP-DM (Cross Industry Standard Process for Data Mining) es un enfoque usado en minería de datos y análisis predictivo estandarizado [42]. Originalmente cuenta con seis fases como, comprensión del negocio, modelado de datos, preparación de datos, evaluación y despliegue de información [42]. Este modelo ha ayudado a construir trabajos de análisis de datos de manera más estructurada, teniendo en cuenta la duplicidad y obtención de conclusiones confiables. Concretamente, en el ambiente clínico con IA, CRISP-DM hace más fácil organizar el área de trabajo que va desde la obtención de datos hasta la evaluación de los modelos construidos [42].

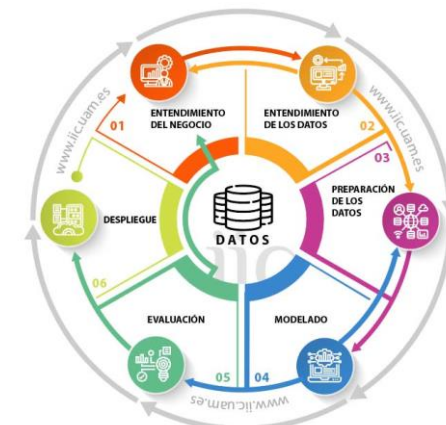


Figura 12: Ciclo CRISP-DM

2.3. Marco Teórico

2.3.1 Redes Neuronales Convolucionales (CNN) en el Diagnóstico Médico

El ambiente de la visión por computadora ha sido transformado por las redes neuronales convolucionales (CNN) que sacan características importantes de forma automática gracias a la capacidad de procesar grandes cantidades de datos [43]. En concreto estas redes están hechas para el análisis de imágenes y su arquitectura tiene una base de capas convolucionales que aprenden ejemplos específicos de datos de datos y, en imágenes médicas esto es importante porque identifica características difíciles [43]. Las CNN son muy utilizadas en el diagnóstico médico ya que, son buenas para clasificar, segmentar y detectar rarezas en varias imágenes, como mamografías o resonancias magnéticas, dando en menor tiempo resultados precisos y seguros [13].

Las CNN pueden acoplar un modelo previamente entrenado en conjunto de datos parecido, ayudando en tareas de clasificación de imágenes médicas sin muchos datos, es decir, que hay una capacidad de Transfer Learning adecuada [20]. Esta característica es destacada en el diagnóstico de enfermedades como cáncer debido a que los datos con etiqueta pueden ser escasos. Es real la mejora de modelos por técnicas de mejora de imagen y configuración de hiperparámetros, gracias a las CNN, esto deja que se hagan sistemas de diagnóstico rápidos y seguros que aumenten los resultados en pacientes al detectarles a tiempo enfermedades [32].

2.3.2 Aprendizaje Profundo y Transfer Learning en Aplicaciones Médicas

Como un área inferior del aprendizaje automático, está el aprendizaje profundo o Deep Learning que en tareas de diagnóstico y análisis de imágenes médicas enseña un rendimiento excelente [44]. En la teoría del aprendizaje profundo se habla de los patrones difíciles y no lineales en los datos, idóneo para el análisis de imágenes biomédicas donde hay cosas suaves que obligatoriamente deben diagnosticarse. Estos patrones difíciles son identificados gracias a las redes neuronales de múltiples capas [44].

En el campo de la medicina se ha optimizado el rendimiento de los modelos de aprendizaje profundo porque se reutiliza el conocimiento de modelos entrenados anteriormente en conjuntos de datos parecidos, todo esto gracias al uso de Transfer Learning [44]. Así no solo se disminuye el tiempo de un diagnóstico más preciso, sino que, reduce costos computacionales también [45].

2.3.3 Mejora de la Calidad de Imágenes Médicas para Diagnóstico asistido

Para la precisión de los modelos de aprendizaje profundo la calidad de las imágenes médicas para entrenamiento deben ser adecuadas, ya que, es un factor bastante importante en los sistemas de diagnóstico por IA [40]. En entornos clínicos, la baja resolución, el ruido o artefactos pueden disminuir la interpretación de parte de algoritmos de IA o especialistas [46]. Para superar esa dificultad, se hicieron varias técnicas de mejora de imagen que ayudan a aumentar la calidad de las imágenes previo al análisis, entre las técnicas más usadas, están reducción de ruido, mejora del contraste o filtros adaptativos [46].

En cuanto a la Inteligencia Artificial en la medicina, el uso de estas herramientas han mejorado la identificación de ejemplos clínicos en imágenes médicas, hablamos del procesamiento de imágenes y la superresolución [8]. Enfoques como la interpolación bicúbica, transformadas de dominio de frecuencia y redes neuronales convolucionales especializadas en reconstrucción de imágenes, para incrementar la fiabilidad de diagnósticos automatizados y corrección de diferencias a la vista, se han aplicado a la medicina [46]. También, técnicas de aprendizaje profundo con algoritmos de mejora de imagen, han ayudado a crear sistemas cruzados que pueden rehacer aspectos anatómicos con alta calidad, incrementando la detección de rarezas en resonancias magnéticas, radiografías y tomografías [47].

Para el ámbito de los sistemas de comunicación, las imágenes transmitidas pueden verse afectadas por interferencias o pérdida de datos, por ejemplo, en OFDM (Multiplexación por División de Frecuencia Ortogonal) [48]. Estos malos efectos de calidad visual se han disminuido porque cuando los modelos de IA trabajan con imágenes precisas y adecuadas a los entornos clínicos reales, se aplican las técnicas de corrección de errores para reconstrucción de imágenes [48].

2.4. Componente de la Propuesta

2.4.1. Arquitectura del sistema

La arquitectura del sistema usada para este proyecto es la de procesamiento de flujo local, en donde los datos que en este caso son las imágenes mamográficas, previamente limpiadas y organizadas en la carpeta “my_dataset” se procesan en el ambiente de desarrollo de anaconda y JupyterNotebook utilizando Python como lenguaje de programación. A partir de ahí, se trabaja con un algoritmo que carga, procesa, entrena y evalúa diferentes combinaciones de hiper parámetros experimentales. El sistema opera localmente sin la intervención de comunicaciones remotas o acceso a internet, en contraparte con una arquitectura cliente servidor, garantizando la eficiencia en la ejecución y un mejor control sobre los recursos, además de una facilidad para tratar datos sensibles [49].

En la figura 13 se puede observar que la estructura ordenada de las imágenes forman parte del punto de partida del sistema. Aquí las muestras se dividen por tipo de lesión y categoría (banignas y malignas) posteriormente se procesan con técnicas de mejora de imagen. Luego se hace una evaluación de los hiper parámetros que conforman, tres arquitecturas de redes neuronales convolucionales (ResNet50, VGG19, EfficientNetB0), dos optimizadores y dos schedulers. Después el algoritmo trabaja con cada configuración y guarda los modelos, junto con métricas como precisión, AUC, sensibilidad y F1-score. Por último, el algoritmo analiza todas las combinaciones y determina cuales tienen el mejor desempeño, automatizando en términos diagnósticos la elección del modelo más eficiente.

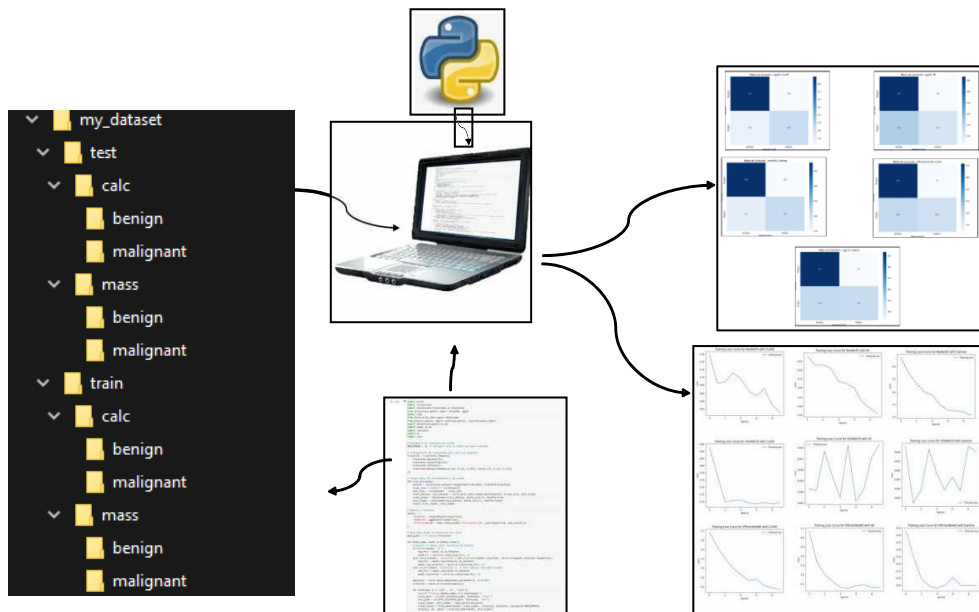


Figura 13: Arquitectura del sistema en Jupyter y Anaconda.

2.4.2. Enfoque Metodológico

El presente proyecto conlleva la metodología CRISP-DM como se mencionó en los apartados de metodología empleada para a estructuración del algoritmo a construir. Este enfoque ha sido usado ampliamente en la minería de datos y aprendizaje profundo debido a que es iterativo, flexible y centrado en la comprensión progresiva del problema mediante análisis de datos [42].

Entendido esto se emplearán las 5 fases del enfoque para la realización de este trabajo, las cuales son, preparación de datos, configuración de modelos, proceso de entrenamiento, evaluación de modelos y visualización y almacenamiento de resultados. Estas fases están adaptadas para cumplir con los objetivos de la investigación del desarrollo de un algoritmo de transfer learning para el análisis de imágenes médicas y a continuación se detallará el desarrollo de cada una de ellas.

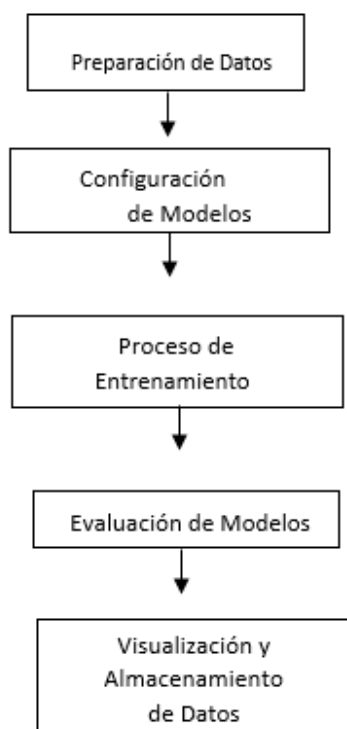


Figura 1: Diagrama de flujo de la metodología empleada en el estudio.

2.4.3. Obtención y preparación de datos

Fase 1

Para la primera parte del proyecto se lleva a cabo la obtención, organización y limpieza del conjunto de datos CBIS-DDSM, que conlleva la base del entrenamiento y evaluación del algoritmo desarrollado. Este conjunto incluye mamografías digitalizadas con características de anotaciones clínicas que categorizan las imágenes en tipos de lesión (“masa” o “calcificación”) y patología (“benigna” o “maligna”).

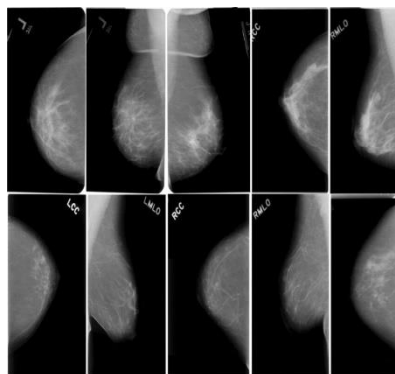


Figura 11: Conjunto de datos del CBIS-DDSM.

Una vez descargado el dataset desde la página oficial del Breast Cancer Image Dataset (kaggle), se continuó con la estructuración en carpetas jerárquicas según el conjunto que utilizaría para entrenamiento o prueba, etiquetado con su tipo de lesión y clasificación patológica. Para ello se implementó un script en Python que automatice el proceso de la lectura de los archivos CSV proporcionados por el dataset y la reubicación de las imágenes JPG relacionadas, las cuales originalmente estaban distribuidas en subdirectorios identificados como DICOM. Este algoritmo extrae los identificadores presentes en las rutas DICOM, localiza las imágenes relacionadas y hace una copia de la estructura más eficaz para el proyecto con nombres unificados que incluyen la vista de la mamografía y tipo de seno.

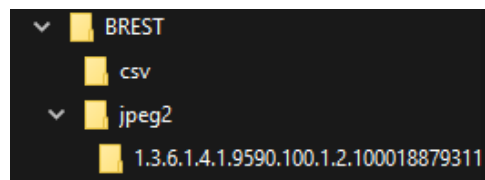


Figura 14: Estructura original del Dataset.

Durante la reorganización de los datos, se tomaron en cuenta cuatro aspectos importantes: la división entre entrenamiento y prueba (train/test), el tipo de lesión (mass o calc), la patología (benigna o maligna) y la vista anatómica. Se aplicaron validaciones cruzadas para detectar las filas que contenían identificadores nulos o patologías no clasificadas, las cuales se destacaron para asegurar la integridad de los datos en el conjunto final. El resultado de esta fase fue la creación de un dataset limpio y organizado con un total de 3568 imágenes que se distribuyen de la siguiente manera:

❖ **Entrenamiento (80%):**

- ✓ mass/benign: 681 imágenes
- ✓ mass/malignant: 637 imágenes
- ✓ calc/benign: 1002 imágenes
- ✓ calc/malignant: 544 imágenes

❖ **Prueba (20%):**

- ✓ mass/benign: 231 imágenes
- ✓ mass/malignant: 147 imágenes

- ✓ calc/benign: 197 imágenes
- ✓ calc/malignant: 129 imágenes

Esta reestructuración hace que sea mucho más sencillo montar los modelos y manejar los datos en PyTorch, además de conseguir que las clases y subtipos estén bien equilibrados para entrenar y evaluar. Poner tanto esmero en preparar los datos es clave para proyectos de clasificación de imágenes, porque dependiendo de la calidad de su limpieza y preprocesamiento se obtendrá cómo van a funcionar los modelos al final.

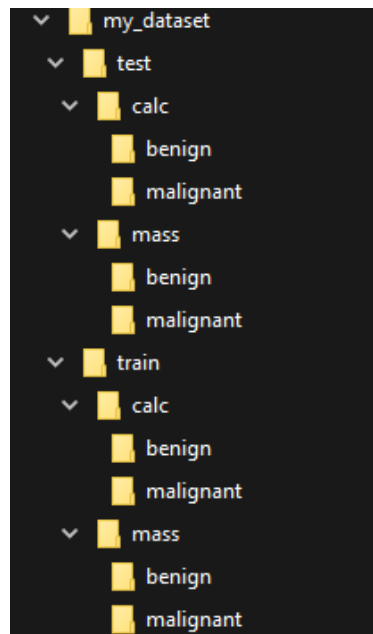


Figura 15: Estructuración final del Dataset.

2.4.4. Configuración de Modelos

Fase 2

Una vez se ha armado adecuadamente el dataset, continúa la configuración y aplicación de los modelos de redes neuronales convolucionales. En esta etapa se seleccionan las tres arquitecturas ampliamente exploradas dentro del campo científico: ResNet50, VGG19 y EfficientNetB0, preentrenadas sobre ImageNet y adaptadas para poder hacer la clasificación binaria entre mamografías benignas y malignas.

La estrategia que sigue la segunda fase se basa en Transfer Learning, donde hay un reemplazo en la última capa de cada uno de los modelos para configurar la nueva

tarea de clasificación de imágenes médicas. Se incluyeron transformaciones previas a las imágenes como redimensionamiento, normalización y conversión a tensor, con el fin de crear un estándar en las entradas y mejorar el entrenamiento. También, se probaron distintas configuraciones de hiperparámetros, incluyendo dos tipos de optimizadores (Adam y SGD), dos compendios de planificadores (StepLR y MultiStepLR) además de variaciones en tasa de aprendizaje y tamaño de batch.

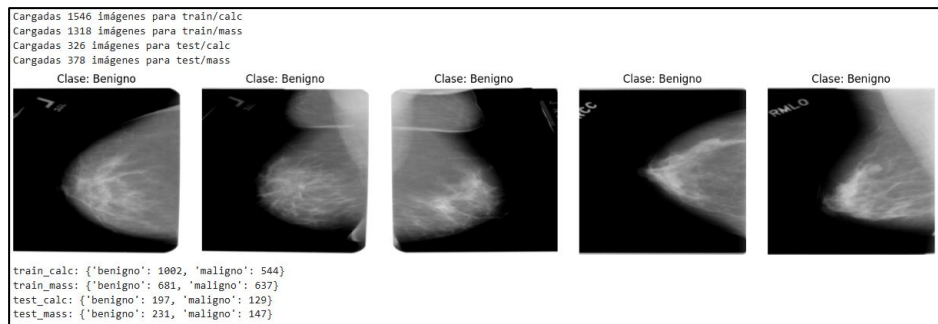


Figura 16: Previsualización de imágenes después de aplicar las transformaciones.

Posterior a este proceso se entrenó cada modelo por un número determinado de épocas y se evaluó el mismo con métricas como sensibilidad, precisión, curva ROC y matriz de confusión. Esto permitió analizar qué combinaciones demostraban un rendimiento superior. Esta fase es la representación del núcleo del en el enfoque utilizado ya que es el punto de partida hacia la experimentación comparativa de las técnicas de mejora de imagen y ajuste fuerte del algoritmo.

2.4.5. Proceso de Entrenamiento

Fase 3

Para la construcción del algoritmo basado en transfer learning que determine cuál es la mejor combinación de hiperparámetros para la clasificación de mamografías, comenzamos con el proceso de entrenamiento, el cuál como mencionamos es el núcleo del proyecto. Se configuró un entorno en Python basado en PyTorch con el fin de probar las diversas combinaciones de modelos, técnicas de mejora de imagen, optimizadores, planificadores (schedulers) y tasas de aprendizaje. Las arquitecturas de redes neuronales convolucionales usadas fueron preentrenadas e inicializadas con pesos en ImageNet. Para mostrar una idea de la magnitud del trabajo hay que entender que en total se ejecutaron 36 experimentos divididos en:

- Modelos de redes neuronales: 3 ((ResNet50, VGG19, EfficientNetB0)
- Técnicas de mejora de imagen: 3 (CLAHE, HE, Gamma)
- Optimizadores: 2 (Adam, SGD)
- Schedulers 2: (StepLR, MultiStepLR)
- Batch size: 1(16)
- Tasa de aprendizaje: 1 (0.001)

Total: $3 \times 3 \times 2 \times 1 \times 1 \times 1 = 36$ *experimentos*.

```

1: Input: Dataset CBIS-DDSM - Breast Cancer Image Dataset
2: Output: Métricas de clasificación, modelos entrenados, gráficos y tabla comparativa
3: Modelos CNN:
4:  ResNet50, VGG19, EfficientNetB0
7: Técnicas de Mejora de Imagen:
8:  CLAHE, HE (Histogram Equalization), Gamma Correction
11: Optimizadores:
12:  Adam, SGD (momentum=0.9)
14: Schedulers:
15:  StepLR (step_size=5, gamma=0.1), MultiStepLR (milestones=[5,10,15], gamma=0.1)
17: Configuración Global:
18:  - batch_size = 16
19:  - learning_rate = 0.001
20:  - num_epochs = 3
21:  - early_stopping = 3
22:  - dispositivo = GPU/CPU automático
23:  - split de validación interno (90% train / 10% val)
24: Preprocesamiento:
25: for cada imagen en el dataset:
26:  Redimensionar a 224x224
27:  Aplicar técnica de mejora de imagen (CLAHE, HE, Gamma)
28:  Convertir a tensor y normalizar con media y desviación de ImageNet
29: end for
30: for cada modelo CNN in {ResNet50, VGG19, EfficientNetB0}:
31:  for cada técnica de mejora TOI in {CLAHE, HE, Gamma}:
32:    for cada combinación (optimizador, scheduler):
33:      Inicializar el modelo con pesos de ImageNet
34:      Ajustar la última capa para clasificación binaria
35:      Crear dataloaders con las imágenes preprocesadas
36:      Entrenar durante 3 épocas con validación interna
37:      Aplicar Early Stopping si no mejora val_loss
38:      Evaluar el modelo en test set:
39:        - Matriz de Confusión
40:        - Curva ROC y AUC
41:        - Precision, Recall, F1-score
42:      Guardar modelo, métricas y gráficos
43:    end for

```

```
44: end for
45: end for
46: Evaluación Final:
47: - Guardar resultados en CSV
48: - Ordenar mejores por Precisión y AUC
49: - Generar gráficos comparativos por Modelo, Técnica, Optimizador y Scheduler
50: - Mostrar mejores 5 configuraciones por Precisión Global y AUC
```

Figura 17: Algoritmo de Entrenamiento y evaluación para la mejor combinación de Redes Neuronales con Transfer Learning y Técnicas de Optimización.

Las configuraciones que se aplicaron en el proceso de entrenamiento empiezan con un número fijo de 3 épocas por experimento, se eligió este número de épocas para poder aplicar la estrategia de “budgeted training” debido a las limitaciones en recursos computacionales al utilizar únicamente el poder del CPU para procesar los datos. Esta decisión se completó con la técnica de Early Stopping en paciencia de 3 épocas, permitiendo detener el entrenamiento de forma anticipada si no se muestran mejoras en la pérdida de validación, optimando el uso del tiempo y los recursos computacionales. También, en cada época se evaluaron pérdidas y métricas encima de pequeños subconjuntos de validación y entrenamiento, lo cual ayudó a realizar ajustes iterativos y precisos.

En cuanto a los parámetros de experimentación antes mencionados se emplearon algunas configuraciones. El tamaño de lote (batch size) fue de 16, la tasa de aprendizaje (learning rate) de 0.001 y las dos técnicas de optimización como Adam, con convergencia rápida y adaptativa, junto con SGD con momentum de 0.9 que facilita escapar de mínimos locales. El par de optimizadores se combinaron con los planificadores (shcedulers), StepLR que disminuye la tasa de aprendizaje cada 5 épocas, y MultiStepLR que lo hace en hitos críticos con el objetivo de mejorar el equilibrio en al final del aprendizaje. Gracias a esta combinación de hiperparámetros se pudo explorar las diferentes trayectorias de optimización y su efecto en el rendimiento del modelo.

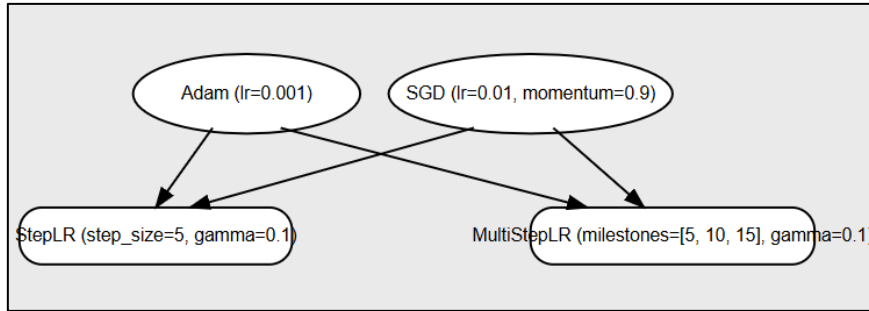


Figura 18: Diagrama de combinaciones de optimizadores y planificadores.

Para la evaluación, dentro del algoritmo se hizo una reserva del 10% del conjunto de entrenamiento como subconjunto de validación interna, validando el modelo en datos que no se vieron durante cada época. Una vez completado el entrenamiento, cada modelo fue evaluado en el conjunto de prueba y se generaron los resultados en matrices de confusión, curvas ROC y parámetros de puntuación. Todo este proceso se automatizó dentro del algoritmo y se registró para cada combinación, asegurando resultados puestos a analizarse y compararse entre sí.

- efficientnet_b0_CLAHE_Adam_MultiStepLR_bs16_lr0.001
- efficientnet_b0_CLAHE_Adam_StepLR_bs16_lr0.001
- efficientnet_b0_CLAHE_SGD_MultiStepLR_bs16_lr0.001
- efficientnet_b0_CLAHE_SGD_StepLR_bs16_lr0.001
- efficientnet_b0_Gamma_Adam_MultiStepLR_bs16_lr0.001
- efficientnet_b0_Gamma_Adam_StepLR_bs16_lr0.001
- efficientnet_b0_Gamma_SGD_MultiStepLR_bs16_lr0.001
- efficientnet_b0_Gamma_SGD_StepLR_bs16_lr0.001
- efficientnet_b0_HE_Adam_MultiStepLR_bs16_lr0.001
- efficientnet_b0_HE_Adam_StepLR_bs16_lr0.001
- efficientnet_b0_HE_SGD_MultiStepLR_bs16_lr0.001
- efficientnet_b0_HE_SGD_StepLR_bs16_lr0.001
- resnet50_CLAHE_Adam_MultiStepLR_bs16_lr0.001
- resnet50_CLAHE_Adam_StepLR_bs16_lr0.001
- resnet50_CLAHE_SGD_MultiStepLR_bs16_lr0.001
- resnet50_CLAHE_SGD_StepLR_bs16_lr0.001
- resnet50_Gamma_Adam_MultiStepLR_bs16_lr0.001
- resnet50_Gamma_Adam_StepLR_bs16_lr0.001
- resnet50_Gamma_SGD_MultiStepLR_bs16_lr0.001
- resnet50_Gamma_SGD_StepLR_bs16_lr0.001

Figura 19: Resultados de todos los experimentos ejecutados en el algoritmo.

2.4.6. Evaluación de modelos y Análisis de resultados

Fase 4 y 5

Los resultados del algoritmo evaluaron las mejores 5 combinaciones de técnicas de optimización de imágenes con optimizadores como Adam y SGD junto con redes neuronales, generando una vía más accesible para el manejo de la data debido a que ya no se analizan los 36 experimentos individualmente sino los 5 óptimos en métricas de precisión global y AUC. En la tabla 1 se muestra un resumen breve de los resultados generados por las 5 combinaciones y analizaremos el comportamiento individual de cada una.

Red	Técnica	VP	VN	FN	PG	AUC
VGG19	Gamma	134	367	142	71.16%	0.766
EfficientNetB0	CLAHE	130	358	146	69.31%	0.726
ResNet50	Gamma	155	328	121	68.60%	0.741
VGG19	HE	114	368	162	68.46%	0.720
VGG19	CLAHE	156	324	120	68.16%	0.745

Tabla 1: Resultados de las mejores Redes Neuronales con Diferentes Técnicas de Mejora de Imagen.

2.4.6.1. SGD con MultiStepLR en VGG19 y Gamma

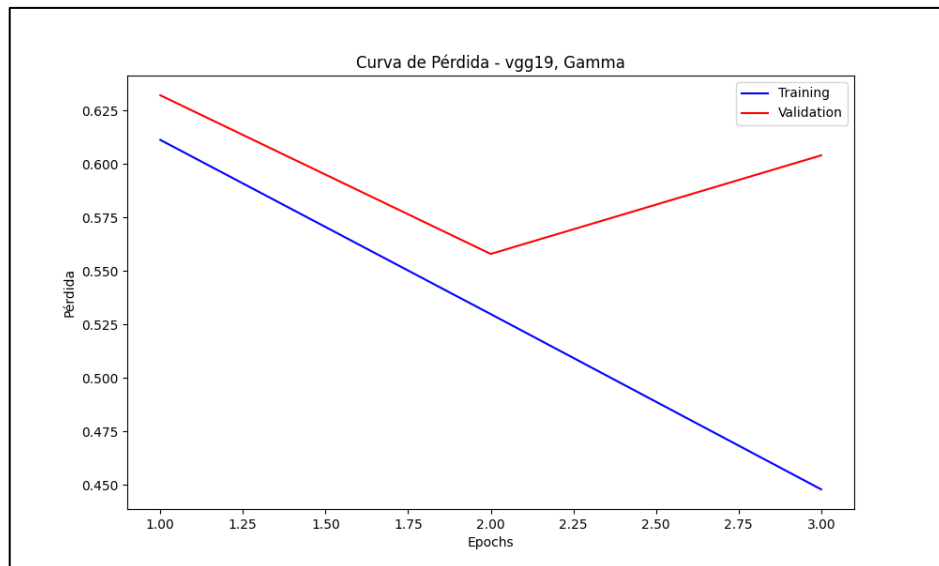


Figura 20: SGD con MultiStepLR.

Analizando la curva de pérdida en SGD con MultiStepLR aplicado a la combinación de VGG19 con la técnica de optimización de imagen Gamma, se

puede observar una reducción de la pérdida dentro del conjunto de entrenamiento a lo largo de las épocas, lo que significa una mejora progresiva del modelo en la adaptación de datos. En cuanto al conjunto de validación se refleja que después de una disminución inicial de la pérdida, la misma vuelve a incrementarse en la última época. La divergencia presente en las curvas de entrenamiento y validación destaca la necesidad de aplicar un mecanismo como EarlyStopping, que de hecho fue implementado para evitar que el modelo siga entrenando en malas condiciones y por eso se convirtió en el más adecuado para el estudio.

2.4.6.2. SGD con MultiStepLR en EfficientNetB0 y CLAHE

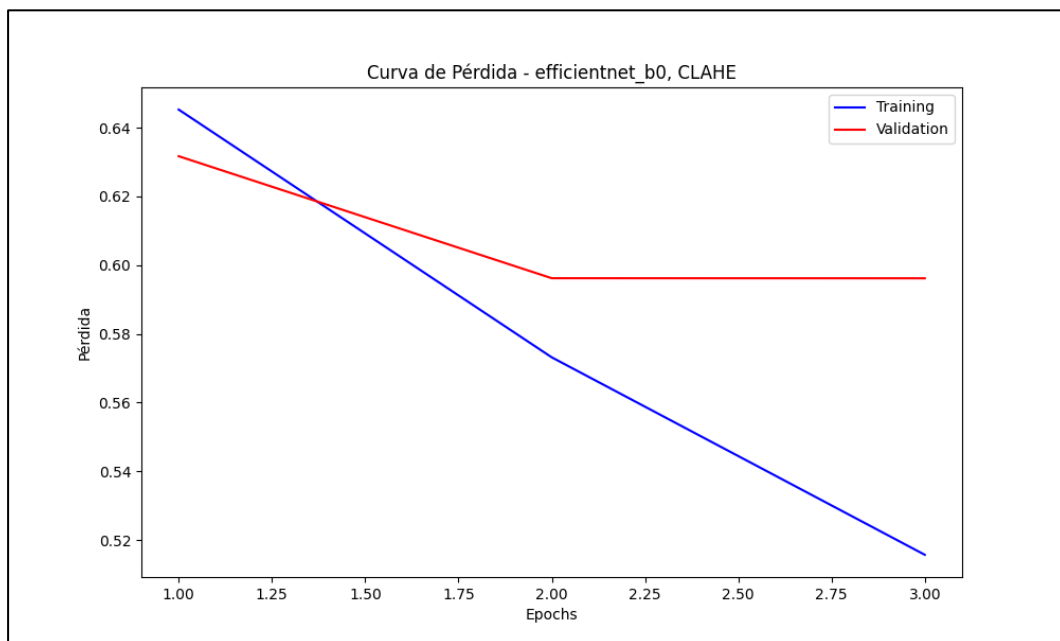


Figura 21: SGD con MultiStepLR y EfficientNetB0.

En la segunda mejor combinación de EfficientNetB0 con CLAHE y SGD con MultiStepLR, se observa que también se combinaron el mismo optimizador y scheduler analizados en la configuración anterior. Se evidencia durante el entrenamiento la disminución sostenida y progresiva de la pérdida, mientras que en la validación hay un balance temprano. No hay indicios de sobreajuste por lo que el entrenamiento fue controlado y estable, demostrando que la técnica CLAHE puede resaltar efectivamente las imágenes mamográficas y combinándola con Adam y MultiStepLR la optimización se mantuvo fluida, aun cuando el entorno de entrenamiento tenía épocas limitadas.

2.4.6.3. SGD con MultiStepLR en ResNet50 y Gamma

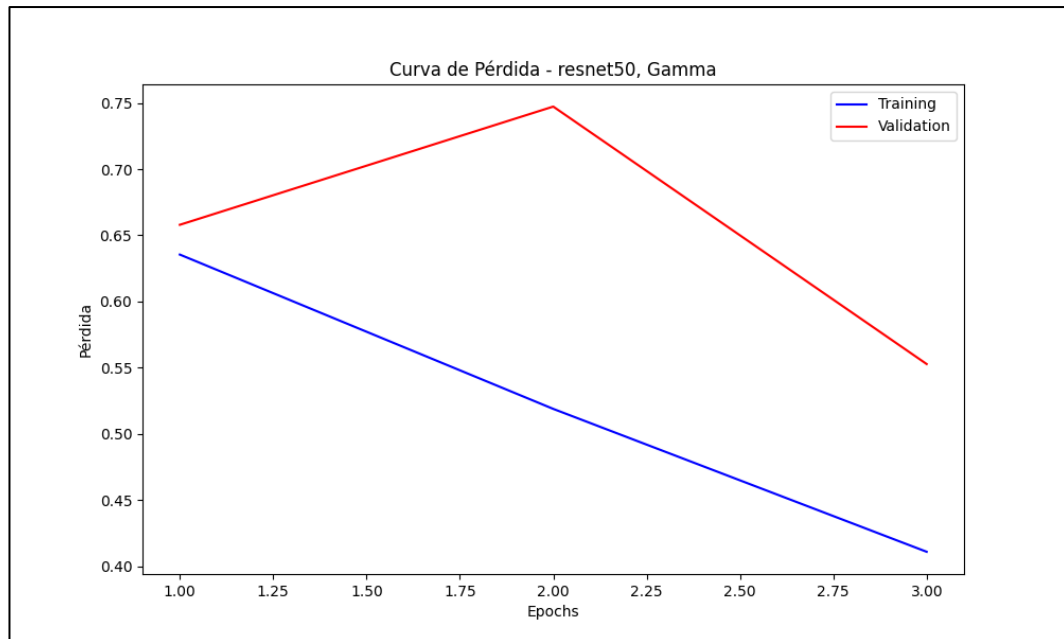


Figura 22: SGD con MultiStepLR y ResNet50.

En la tercera mejor combinación de ResNet50 y Gamma combinados igualmente con SGD y MultiStepLR, se evidencia que la pérdida de validación tiene una inestabilidad que aumenta durante la segunda época generando una fuerte caída. Esto podría estar relacionado con la sensibilidad del modelo a los ajustes de la tasa de aprendizaje del scheduler. En cambio, en la pérdida de entrenamiento hay una disminución constante lo que sugiere que la corrección de gamma mantiene un aprendizaje efectivo para estabilizarse al final del modelo.

2.4.6.4. SGD con StepLR en VGG19 y HE

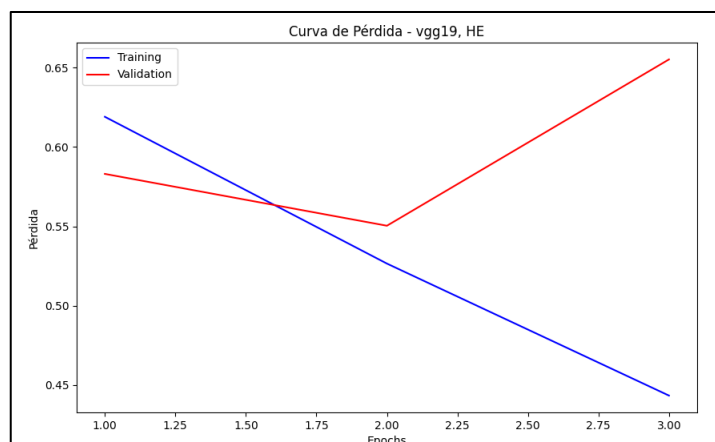


Figura 23: SGD con StepLR y VGG19

Para la combinación de VGG19 con HE y el optimizador SGD con e scheduler StepLR, se representa un comportamiento divergente entre las curvas de validación y entrenamiento. En la pérdida de entrenamiento hay una bajada constante desde 0.62 hasta 0.44 a lo largo de las 3 épocas, mientras que en la pérdida de validación hay un evento raro, ya que al principio disminuye hasta 0.55 en la segunda época, pero después sube exponencialmente hasta sobrepasar 0.65 en la última época, lo que podría indicar un sobreajuste del modelo.

2.4.6.5. SGD con StepLR en VGG19 y CLAHE

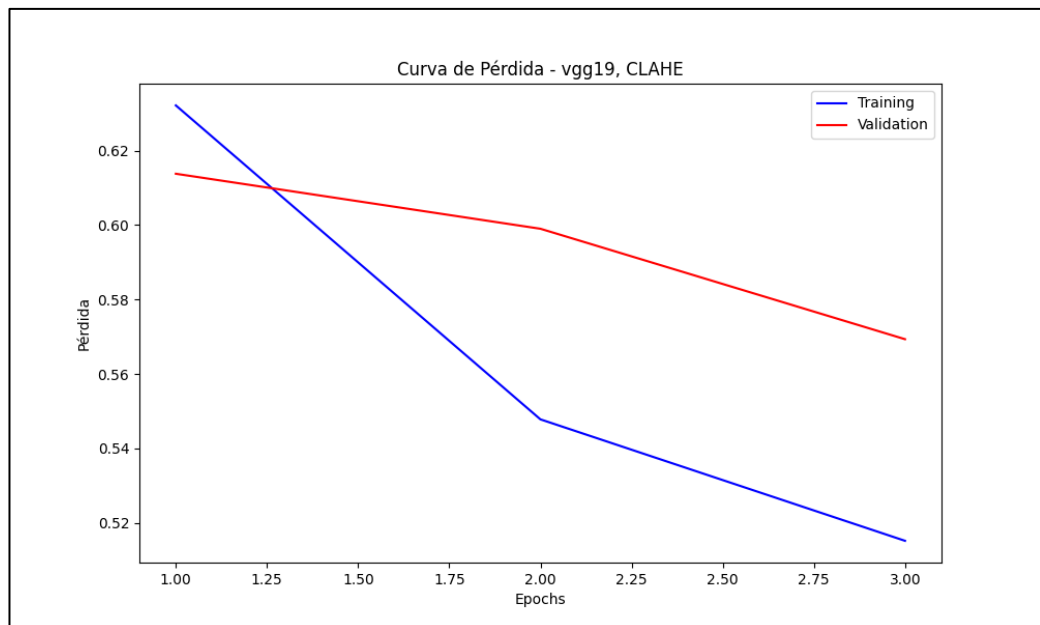


Figura 24: SGD con StepLR y VGG19

Por último, la 5ta mejor combinación es SGD con StepLR en VGG19 Y CLAHE, demostrando consistencia en ambas curvas de pérdida. Hay una bajada más pronunciada en la pérdida de entrenamiento, que decrece desde 0.63 hasta 0.51 a lo largo de todas las épocas. En cambio, en la pérdida de validación hay una bajada de 0.61 hasta 0.57 pero con menor pronunciamiento de la pendiente más ligero. El comportamiento paralelo en ambas curvas sugiere un buen equilibrio de aprendizaje y generalización, concluyendo que la técnica de mejora CLAHE está dando características más fuertes en la optimización del modelo sin llegar a sobreajustarlo.

2.4.6.6. Matriz de Confusión: 5 mejores combinaciones

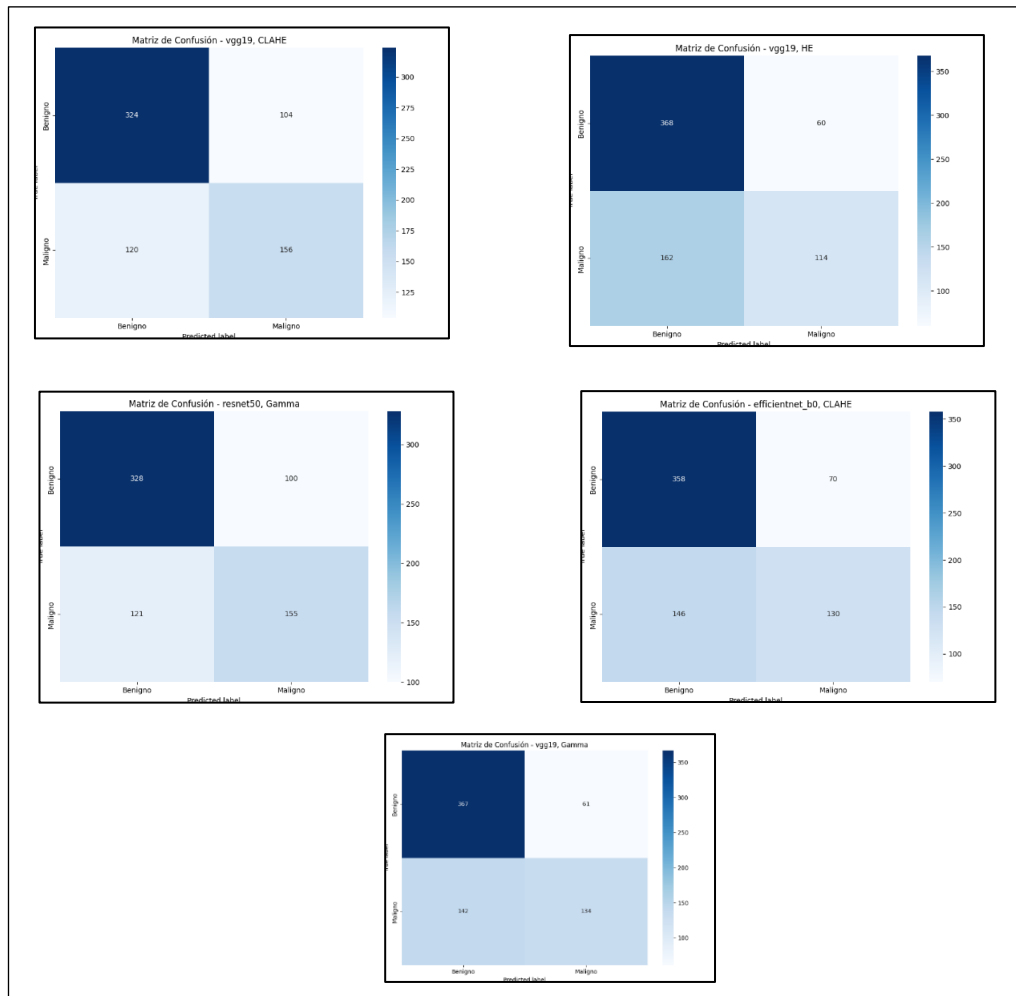


Figura 25: Matrices de Confusión de las mejores combinaciones.

El análisis de las matrices de confusión demuestran importantes aspectos sobre el rendimiento de las 5 mejores configuraciones, las cuales incluyen las 3 redes neuronales propuestas, VGG19, ResNet50 y EfficientNetB0, con las 3 técnicas de mejora de imagen, CLAHE, Gamma y HE, todas ellas optimizadas con SGD y el scheduler StepLR o MultiStepLR dependiendo de la combinación.

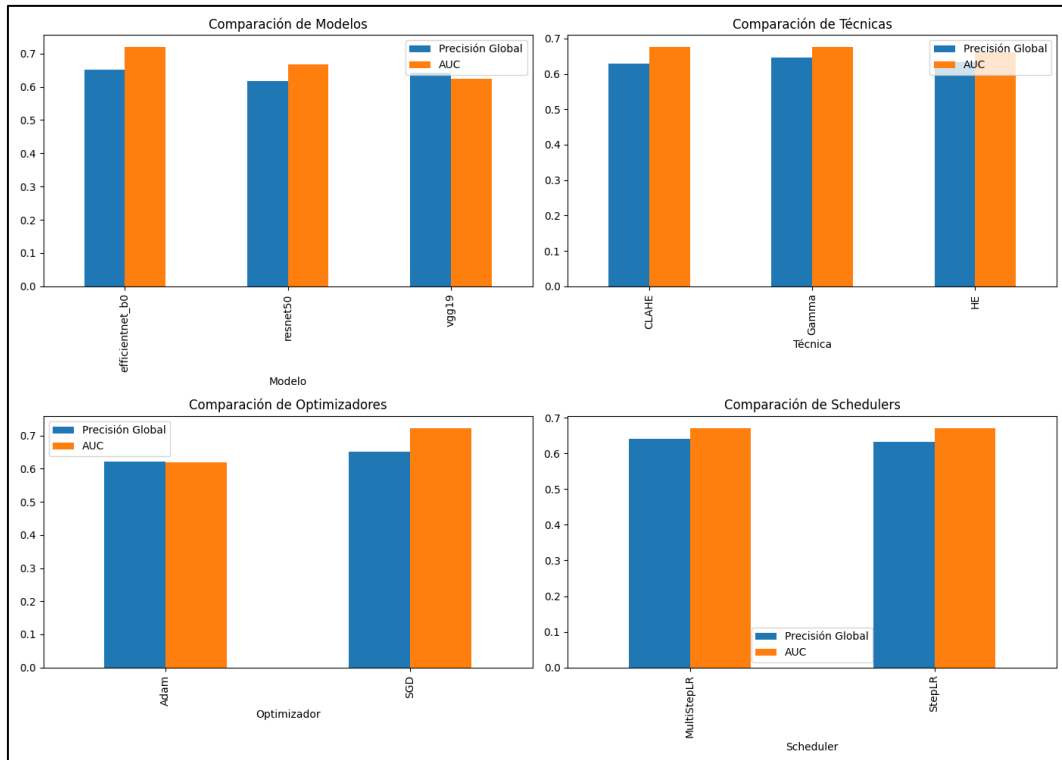


Figura 26: Comparaciones globales de modelo, técnica de mejora, optimizadores y schedulers.

En el puesto 5 de la mejor combinación está VGG19 con CLAHE, se observa una precisión global aceptable (0.681) y un AUC bueno (0.7454). En cuanto a su rendimiento la clasificación de muestras benignas, tenemos una buena precisión (VN=324), pero presenta una tasa media de falsos negativos (FN=120). También muestra una tasa moderada de falsos positivos (FP=104) sin embargo los verdaderos positivos (VP=156) superan a los falsos negativos. Hay un balance competitivo muy bueno en esta combinación.

VGG19 con HE entran en el puesto 4, obteniendo una precisión global un poco más alta que la combinación anterior (0.684) y un AUC aceptable de (0.72). Esta configuración demostró igualmente una tasa alta de clasificación de muestras benignas (VN=368) y una tasa de falsos positivos muy baja (FP=60). En su rendimiento para las muestras malignas tenemos una tasa más alta de falsos negativos (FN=162) y las correctamente clasificadas con malignidad bastante moderado (VP=114). El modelo es robusto, pero en la detección de malignos tiene

muy poca sensibilidad por lo que el uso de otra técnica de mejora de imagen podría beneficiar su rendimiento.

En el puesto 3 está ResNet50 con Gamma tiene una precisión global ligeramente más alta que las dos combinaciones anteriores (0.686) con un AUC bueno (0.741). Muestra un alta en falsos positivos (FP=100) en comparación con el modelo anterior y una tasa moderada de falsos negativos (FN=121) pero no lo suficientemente diferente como para ser la mejor opción. Igual que en las otras configuraciones tenemos alta detección de positivos (VN=328) y en negativos presenta buen rendimiento (VP=155). Se demuestra que puede ser buena opción de clasificación binaria ya que tiene el mejor rendimiento en detección de malignos sin embargo presentó más falsos positivos.

EfficientNetB0 con CLAHE se encuentra en el puesto 2 presentando una precisión global a punto de topar un puntaje aceptable (0.69) y un AUC igual que las anteriores configuraciones (0.726). La tasa de falsos positivos es relativamente baja (FP=70) y las muestras benignas clasificadas correctamente es bastante alta igual a los modelos anteriores (VN=358). En la malignidad vemos un menor rendimiento con una disminución de sensibilidad en verdaderos positivos (VP=130) y mayores falsos negativos (FN=146). A pesar de esto se puede ver un comportamiento balanceado, controlando mejor los falsos positivos que otras configuraciones.

En el puesto número 1 llega VGG19 con Gamma, mostrando la mayor precisión global (0.71) y mayor AUC (0.76). Este modelo muestra un buen manejo de falsos positivos (FP=61), por lo que las muestras benignas con corrección clasificadas son igualmente altas (VN=367). Con respecto a las muestras malignas obtenemos un rendimiento aceptable y con menores falsos negativos que la combinación del segundo puesto (FN=142), las muestras malignas clasificadas correctamente son estables (VP=134). Esta configuración fue seleccionada por el algoritmo gracias al mejor desempeño en todas las métricas generales y si bien tiene cierta dificultad en malignos, para la última parte del estudio será útil.

Cada configuración ha sido evaluada con graficas de entrenamiento como curvas ROC y las matrices de confusión antes analizadas que se generaron automáticamente como resultado del algoritmo. Gracias a estas herramientas visuales se pudo identificar las fortalezas, patrones y debilidades de cada combinación. Además de estos resultados medibles, el algoritmo también generó archivos CSV y de texto con las métricas evaluadas por modelo, técnica, optimizador y scheduler completando la fase 5 del proyecto.

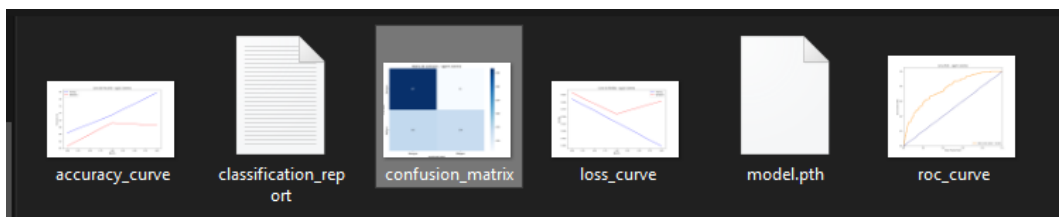


Figura 27: Archivos generados por el algoritmo a partir de la evaluación de cada experimento.

En conclusión, el algoritmo confirma que la combinación VGG19 + Gamma + SGD + MultiStepLR es la más efectiva para la clasificación binaria de mamografías dentro de este estudio, mientras que otras configuraciones como EfficentNetB0 mostraron una versatilidad en el momento de ser acompañados con las técnicas de optimización de imagen adecuadas.

Mejores 5 configuraciones por precisión global:						
	Modelo	Técnica	Optimizador	Scheduler	Precisión Global	\
23	vgg19	Gamma	SGD	MultiStepLR	0.711648	
27	efficientnet_b0	CLAHE	SGD	MultiStepLR	0.693182	
11	resnet50	Gamma	SGD	MultiStepLR	0.686080	
18	vgg19	HE	SGD	StepLR	0.684659	
14	vgg19	CLAHE	SGD	StepLR	0.681818	
AUC						
23	0.766304					
27	0.726153					
11	0.741111					
18	0.720049					
14	0.745403					

Figura 28: Algoritmo mostrando cual es la mejor combinación.

2.4.7. Diseño de Interfaces

Una vez elegida la mejor combinación evaluada por el algoritmo, se procedió a crear una aplicación en Python que probara el rendimiento del modelo con el fin de documentar cuantos aciertos y errores puede tener el mismo al analizar tanto mamografías benignas o como malignas. En este caso se trabajará con la combinación que incluye la red Neuronal VGG19 con la técnica de optimización de imagen de corrección de gama debido a que está obtuvo el primer lugar en la evaluación del algoritmo. La aplicación se ha estructurado en diferentes módulos que individualmente manejan, el procesamiento de la imagen, diseño de la interfaz, integración del modelo, resultados y demás configuraciones para una mejor gestión del código.

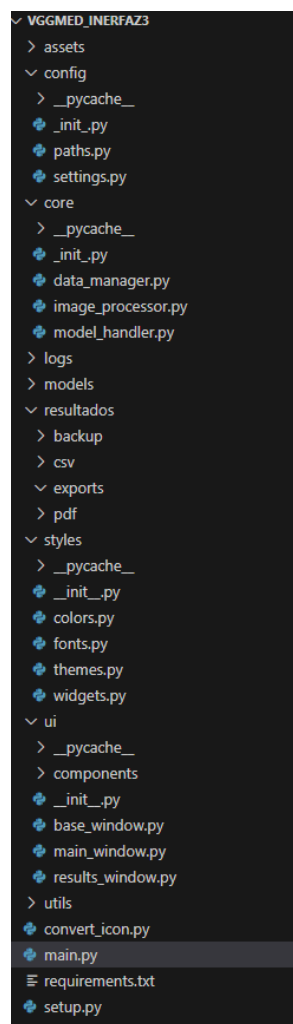


Figura 29: Estructura de la aplicación.

Pantalla principal

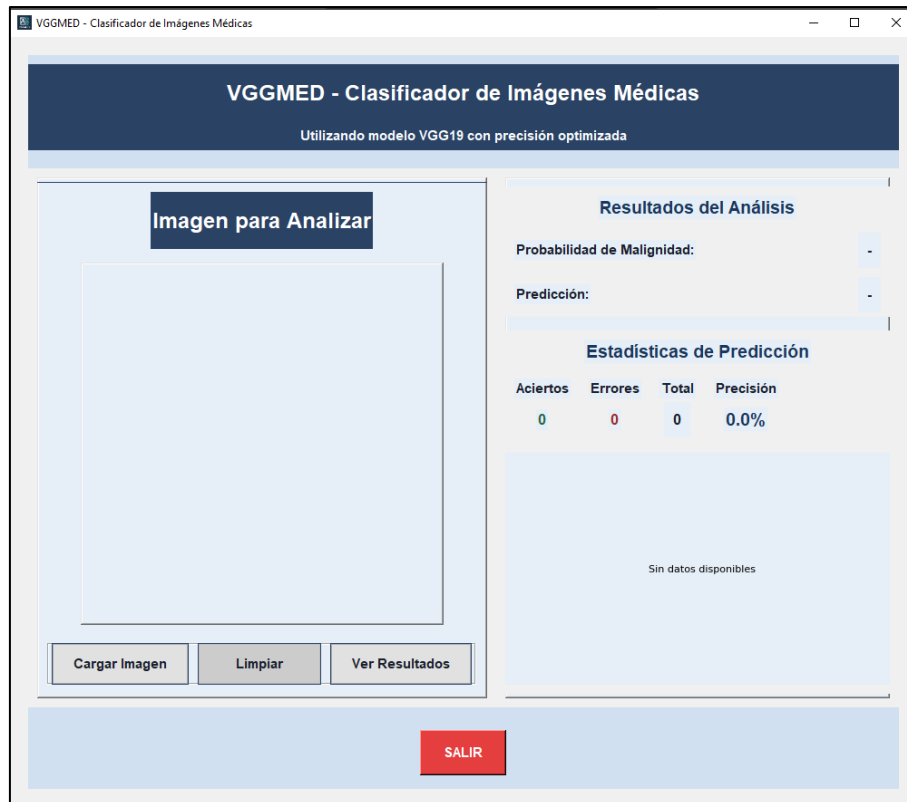


Figura 30: Pantalla Principal.

La aplicación cuenta con dos pantallas, la primera es la pantalla principal en donde se podrán subir las imágenes y el programa hará la predicción, después mostrará los resultados y las estadísticas de predicción. En los resultados están la probabilidad de que la imagen cargada sea maligna y la predicción que el sistema hizo. En cuanto a las estadísticas de predicción encontramos los aciertos, errores y total de imágenes analizadas, por último, la precisión de todos los análisis hechos que promedia en base a las imágenes que se carguen $\frac{\text{Predicciones correctas}}{\text{Total de predicciones}} \times 100$. Como ultima parte el resumen del rendimiento se mostrará en un gráfico pastel para indicar los resultados de manera interactiva y si queremos ver el historial de todas las predicciones podremos ir al botón de “ver resultados”.

Pantalla de historial de predicciones

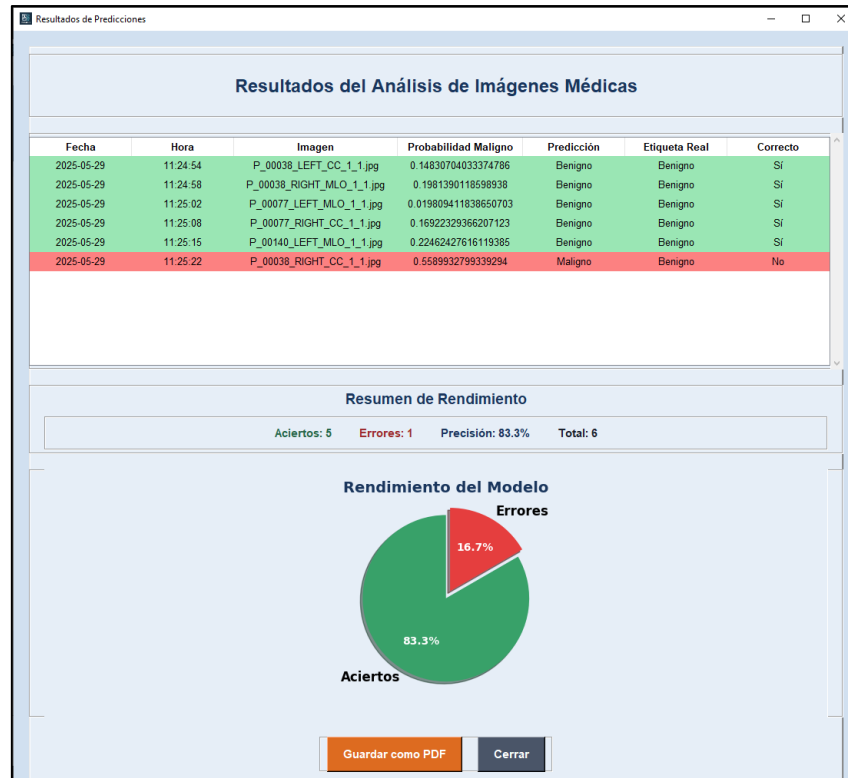


Figura 31: Pantalla de historial de predicciones.

Esta es la segunda pantalla, aquí se muestran cada de una de las predicciones que se han hecho, acopladas en una tabla con diferentes columnas que se dividen en, fecha de la predicción, hora, nombre y tipo de la imagen, probabilidad de malignidad, el tipo de predicción que hizo el programa, la etiqueta real y si acertó o tuvo equivocaciones. Debajo de la tabla se encuentra otra vez las estadísticas de predicción con el grafico pastel para que en conjunto con la tabla sea más fácil observar a detalle el rendimiento de los análisis. Por último, tenemos dos botones, el de cerrar y el de exportar todos los datos como PDF.

Archivos exportados

```
resultados > csv | prediction_results.20250529.110704.csv > data
1 Fecha,Hora,Imagen,Probabilidad Maligno,Predicción,Etiqueta Real,Correcto
2 2025-05-29,11:24:54,P_00038_LEFT_CC_1_1.jpg,0.14830704033374786,Benigno,Benigno,Sí
3 2025-05-29,11:24:58,P_00038_RIGHT_MLO_1_1.jpg,0.1981390118598938,Benigno,Benigno,Sí
4 2025-05-29,11:25:02,P_00077_LEFT_MLO_1_1.jpg,0.019809411838650703,Benigno,Benigno,Sí
5 2025-05-29,11:25:08,P_00077_RIGHT_CC_1_1.jpg,0.16922329366207123,Benigno,Benigno,Sí
6 2025-05-29,11:25:15,P_00140_LEFT_MLO_1_1.jpg,0.22462427616119385,Benigno,Benigno,Sí
7 2025-05-29,11:25:22,P_00038_RIGHT_CC_1_1.jpg,0.5589932799339294,Maligno,Benigno,No
8
```

Resultados del Análisis de Imágenes Médicas						
MODELO VGG19						
Fecha	Hora	Imagen	Probabilidad Maligno	Predicción	Etiqueta Real	Correcto
2025-05-29	11:24:54	P_00038_LEFT_CC_1_1.jpg	0.14830704033374786	Benigno	Benigno	Si
2025-05-29	11:24:58	P_00038_RIGHT_MLO_1_1.jpg	0.1981390118598938	Benigno	Benigno	Si
2025-05-29	11:25:02	P_00077_LEFT_MLO_1_1.jpg	0.019809411838650703	Benigno	Benigno	Si
2025-05-29	11:25:08	P_00077_RIGHT_CC_1_1.jpg	0.16922329366207123	Benigno	Benigno	Si
2025-05-29	11:25:15	P_00140_LEFT_MLO_1_1.jpg	0.22462427616119385	Benigno	Benigno	Si
2025-05-29	11:25:22	P_00038_RIGHT_CC_1_1.jpg	0.5589932799339294	Maligno	Benigno	No

Resumen de Rendimiento		
Categoría	Cantidad	Porcentaje
Aciertos	5	83.3%
Errores	1	16.7%

Figura 32: Archivos CSV y PDF con los resultados del análisis.

Cada sesión de la aplicación generará un archivo CSV con los datos de la tabla mostrados en la pantalla de historial de predicciones, sin embargo, para una mejor legibilidad y presentación de los resultados está la opción de exportarlos en formato PDF incluyendo información como el rendimiento. Así se podrá entender de una manera más adecuada el funcionamiento de esta aplicación.

2.4.8. Pruebas

Para el uso de la aplicación que permita verificar el rendimiento del mejor modelo elegido por el algoritmo, se realizarán 2 tipos de pruebas:

- Mamografías malignas etiquetadas.
- Mamografías benignas etiquetadas.



 benign	17/2/2025 11:08:PM	Carpeta de archivos
 malignant	17/2/2025 11:08:PM	Carpeta de archivos

Figura 33: Imágenes listas para los 3 tipos de prueba.

El programa procesa las imágenes con corrección de gamma y las redimensiona a 512 x 512 con el fin de que el formato sea el adecuado para el análisis. También existen restricciones, ya que únicamente se podrán subir mamografías digitalizadas en formato JPJ, JPEG y PNG, otro tipo de imágenes que no sean mamografías no serán aceptadas.

```

def is_medical_image(self, image_path: str) -> Tuple[bool, str]:
    """
    Verifica si una imagen es una mamografía válida

    Args:
        image_path: Ruta de la imagen

    Returns:
        Tuple[bool, str]: (es_válida, mensaje)
    """
    try:
        # Verificar extensión del archivo
        if not self._validate_file_extension(image_path):
            return False, self.validation_config.INVALID_FORMAT_MSG

        # Cargar imagen
        image = Image.open(image_path).convert('RGB')

        # Verificar resolución mínima
        if not self._validate_resolution(image):
            return False, self.validation_config.SIZE_TOO_SMALL_MSG.format(
                image.size[0], image.size[1],
                self.config.MIN_RESOLUTION, self.config.MIN_RESOLUTION
            )
    
```

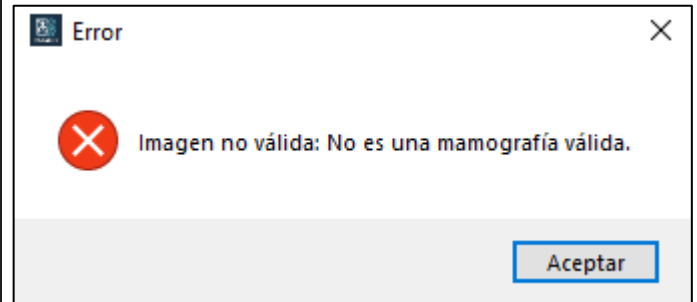


Figura 34: Función para verificar si la imagen es una mamografía.

Muestras malignas etiquetadas

Para las mamografías malignas etiquetadas, el programa obtiene las etiquetas verdaderas en base a las carpetas en donde estén alojadas las imágenes, hay que recordar que esta aplicación se desarrolla en base al entrenamiento y prueba del algoritmo con el fin de observar cuantos aciertos o equivocaciones tiene el mejor modelo elegido. Por esta razón las imágenes ya están previamente etiquetadas en las carpetas y la cantidad de muestras a probar serán 50. Para iniciar se da clic en el botón de “Cargar Imagen”, luego se abrirá el explorador de archivos y se podrá probar con cada imagen que se necesite, en este caso hasta completar 50 muestras.



Figura 35: Cargar imagen.

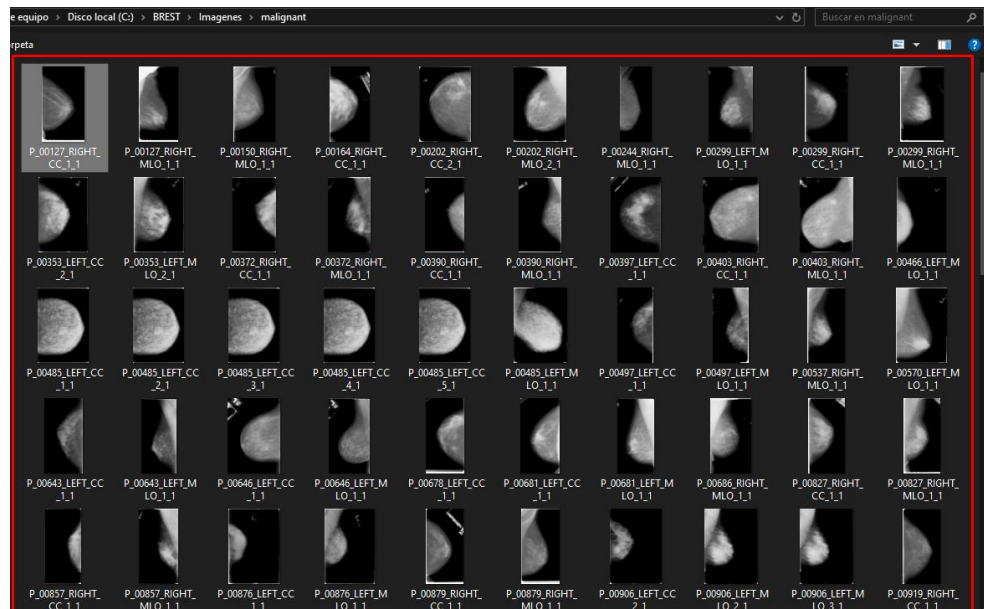


Figura 36: Imágenes malignas a probar.

Una vez se prueban las 50 imágenes se obtienen los resultados. En primer lugar, el programa clasificó bien 40 de las 50 imágenes dando un total de 80% en precisión del modelo.

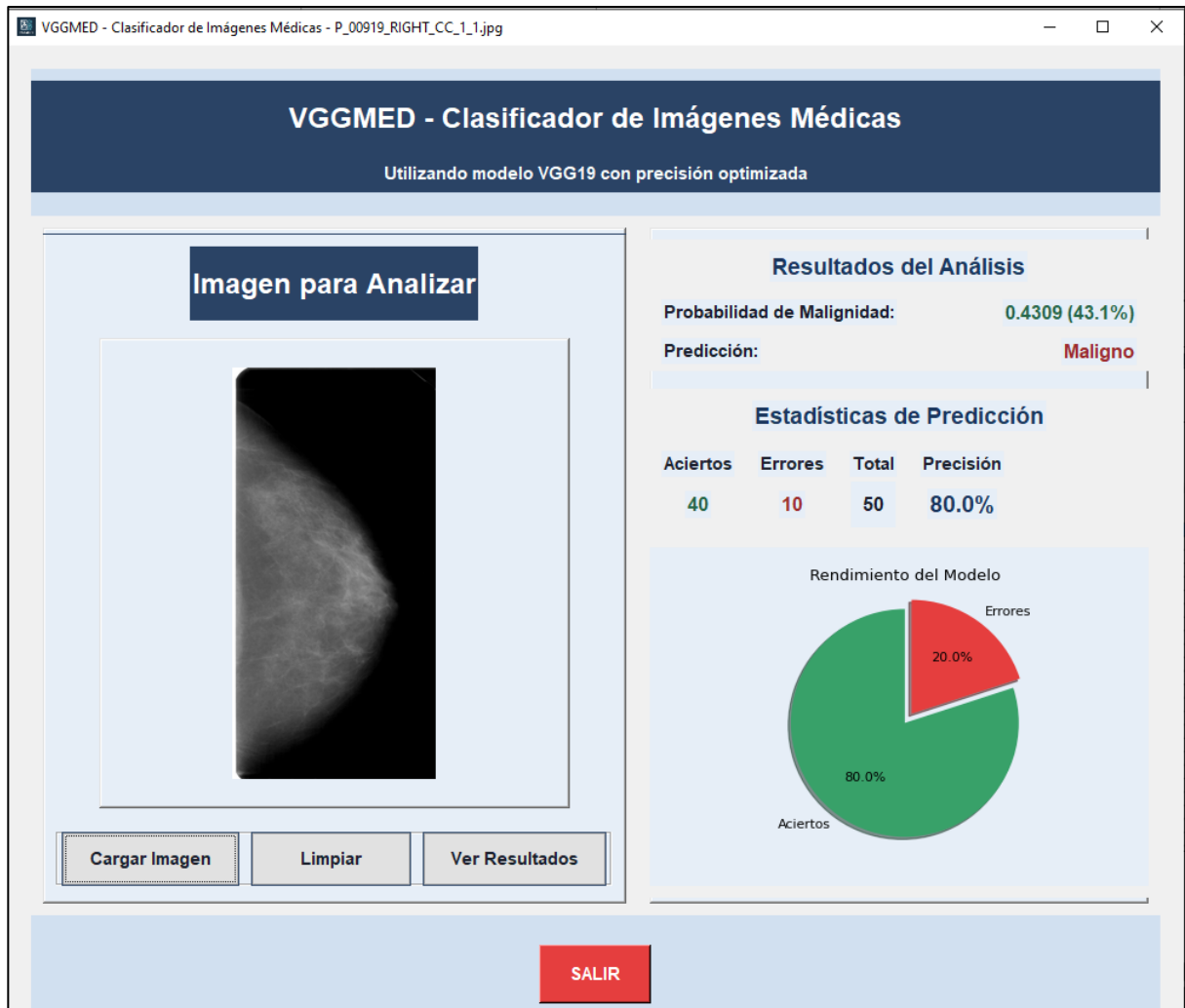


Figura 37: Pantalla principal con resumen de las 50 muestras malignas analizadas.

Para ver más a detalle el historial de predicciones se abre el botón de “Ver Resultados”. Los resultados indican que poniendo a prueba esta combinación hay un buen desempeño analizando las mamografías etiquetadas como malignas.

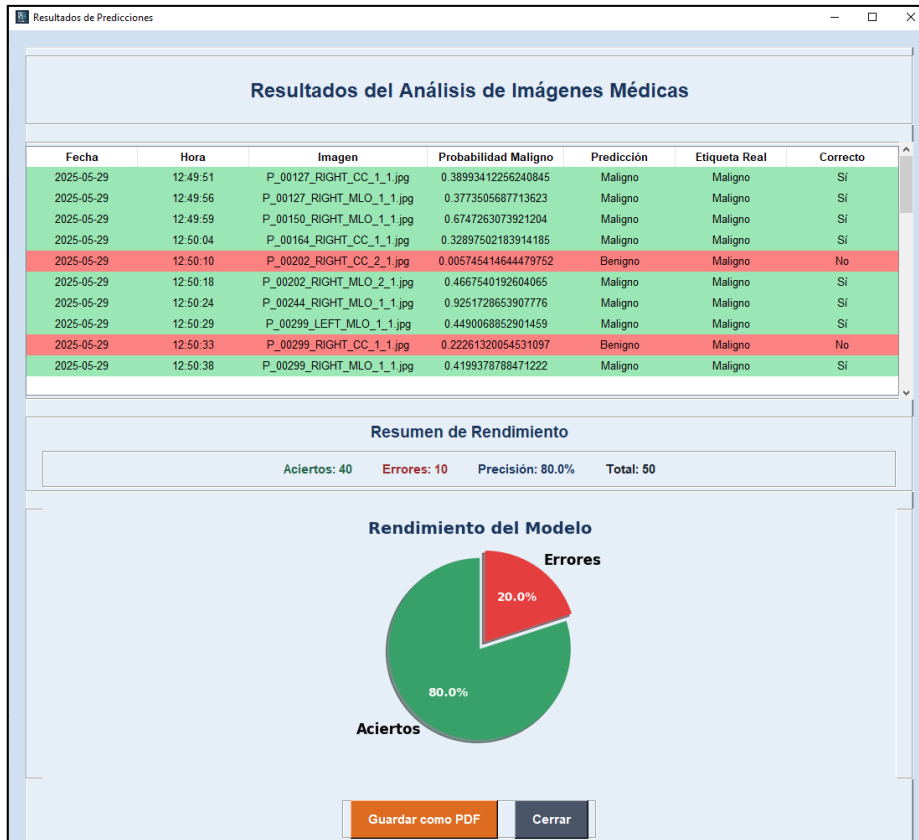


Figura 38: Pantalla de resultados con el historial de predicciones.

Posteriormente se da clic en el botón de “Guardar como PDF” para verificar el resumen de las pruebas.

2025-05-23	11:40:37	P_00646_LEFT_MLO_1_1.jpg	0.9016438722610474	Maligno	Maligno	Si
2025-05-23	11:40:46	P_00678_LEFT_CC_1_1.jpg	0.6112956404685974	Maligno	Maligno	Si
2025-05-23	11:40:56	P_00681_LEFT_CC_1_1.jpg	0.7418572902679443	Maligno	Maligno	Si
2025-05-23	11:41:04	P_00681_LEFT_MLO_1_1.jpg	0.9081246256828308	Maligno	Maligno	Si
2025-05-23	11:41:12	P_00686_RIGHT_MLO_1_1.jpg	0.5475165843963623	Maligno	Maligno	Si
2025-05-23	11:41:20	P_00827_RIGHT_CC_1_1.jpg	0.41646844148635864	Maligno	Maligno	Si
2025-05-23	11:41:41	P_00827_RIGHT_MLO_1_1.jpg	0.5030261874198914	Maligno	Maligno	Si
2025-05-23	11:41:55	P_00857_RIGHT_CC_1_1.jpg	0.7464845180511475	Maligno	Maligno	Si
2025-05-23	11:41:59	P_00857_RIGHT_MLO_1_1.jpg	0.5601249933242798	Maligno	Maligno	Si
2025-05-23	11:42:03	P_00876_LEFT_CC_1_1.jpg	0.259155809879303	Benigno	Maligno	No
2025-05-23	11:42:07	P_00876_LEFT_MLO_1_1.jpg	0.23450155556201935	Benigno	Maligno	No
2025-05-23	11:42:12	P_00879_RIGHT_CC_1_1.jpg	0.5339111685752869	Maligno	Maligno	Si
2025-05-23	11:42:33	P_00879_RIGHT_MLO_1_1.jpg	0.4972270131111145	Maligno	Maligno	Si
2025-05-23	11:42:55	P_00906_LEFT_CC_2_1.jpg	0.9161485433578491	Maligno	Maligno	Si
2025-05-23	11:43:06	P_00906_LEFT_MLO_2_1.jpg	0.7443951368331909	Maligno	Maligno	Si
2025-05-23	11:43:16	P_00919_RIGHT_CC_1_1.jpg	0.4309253394603729	Maligno	Maligno	Si
2025-05-23	11:43:23	P_00919_RIGHT_MLO_1_1.jpg	0.3774034082889557	Maligno	Maligno	Si

Resumen de Rendimiento

Categoría	Cantidad	Porcentaje
Aciertos	40	80.0%
Errores	10	20.0%

Figura 39: Reporte PDF con el resumen de la primera prueba.

Entonces, para esta primera prueba con mamografías malignas etiquetadas, se evidencia que después de analizar 50 imágenes el modelo logró acertar correctamente 40, demostrando su buen desempeño en este tipo de muestras y obteniendo una precisión del 80%, un puntaje aceptable pero no perfecto para el ámbito clínico.

Muestras benignas etiquetadas

Para las mamografías benignas etiquetadas, el programa obtiene las etiquetas verdaderas igual que en la prueba anterior, es decir por el nombre de la carpeta en la que se encuentran. Igualmente se probará con una cantidad de 50 imágenes y se inicia desde la pantalla principal cargando las imágenes una individualmente hasta completar la cantidad total a probar.

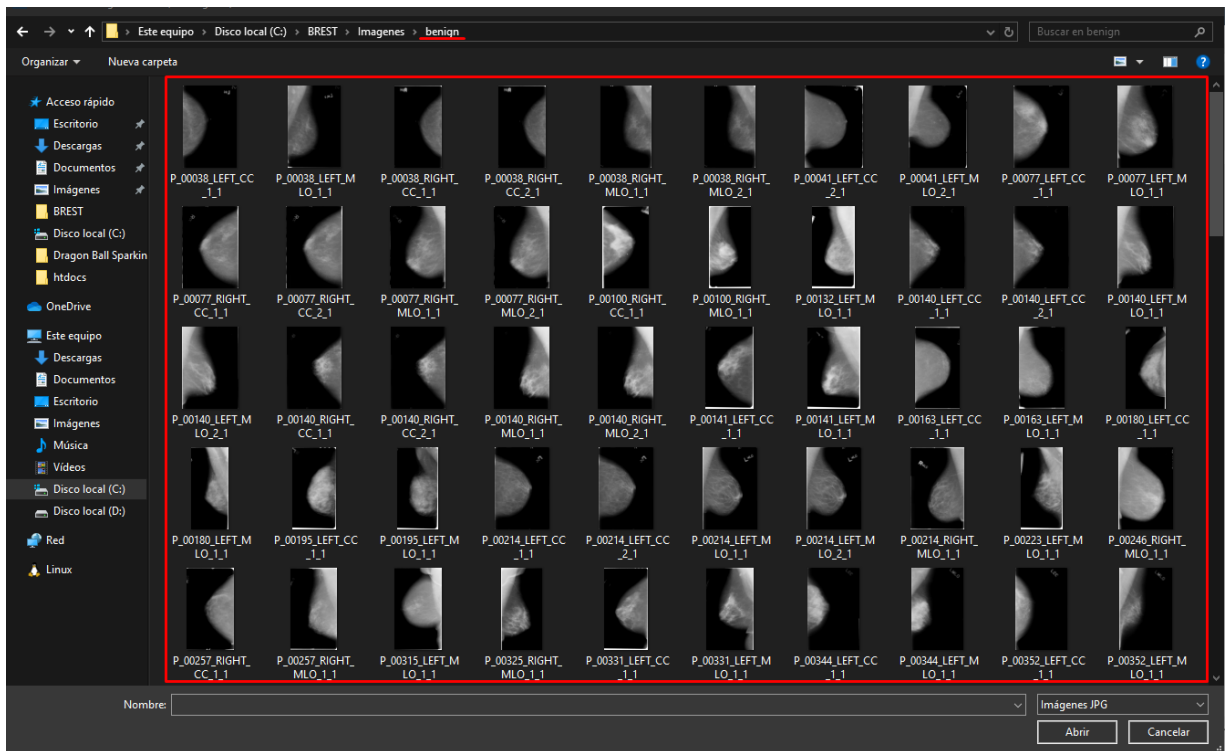


Figura 40: Imágenes benignas a probar.

Analizadas las 50 imágenes benignas, se puede ver que el rendimiento es similar a la prueba anterior. De las 50 muestras 39 fueron analizadas correctamente y 11 tuvieron error.

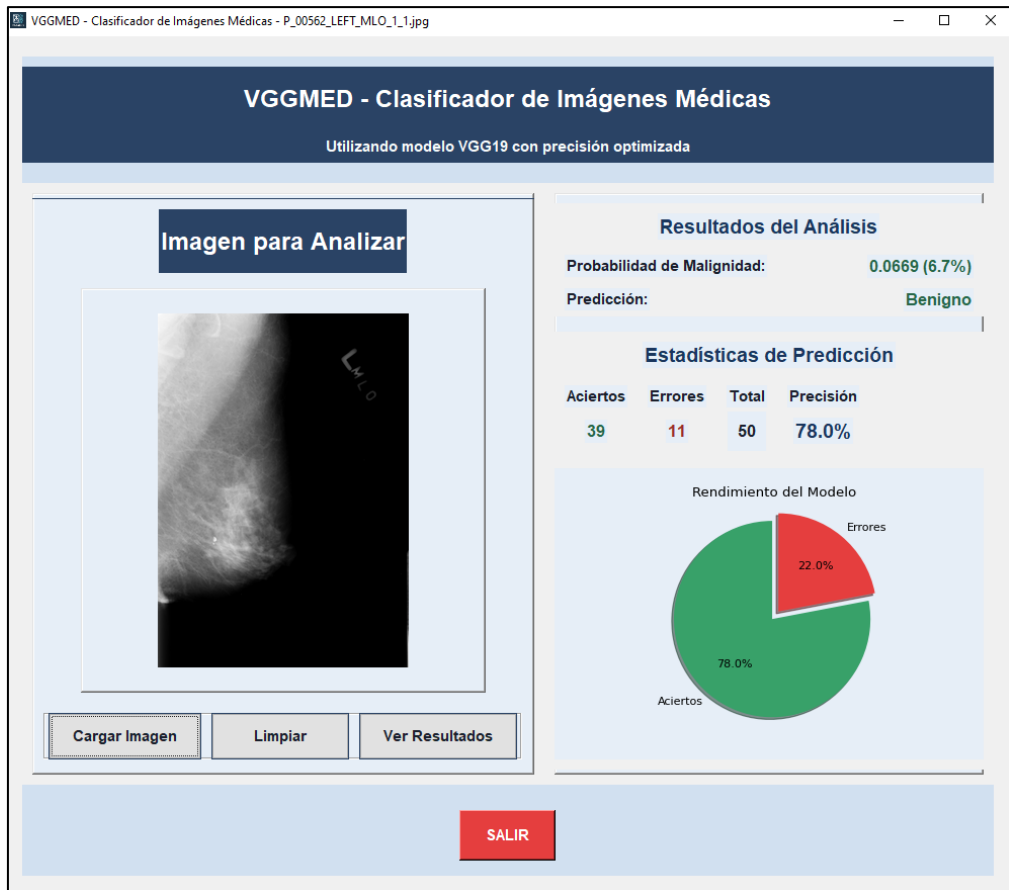


Figura 41: Pantalla principal con resumen de las 50 muestras benignas analizadas.

Gracias al gráfico pastel y la revisión en la pantalla de historial de predicciones se evidencia que el modelo tuvo una precisión del 78% y una tasa de error del 22%, no llega al 80% en el primer parámetro sin embargo se acerca a la prueba de imágenes malignas.

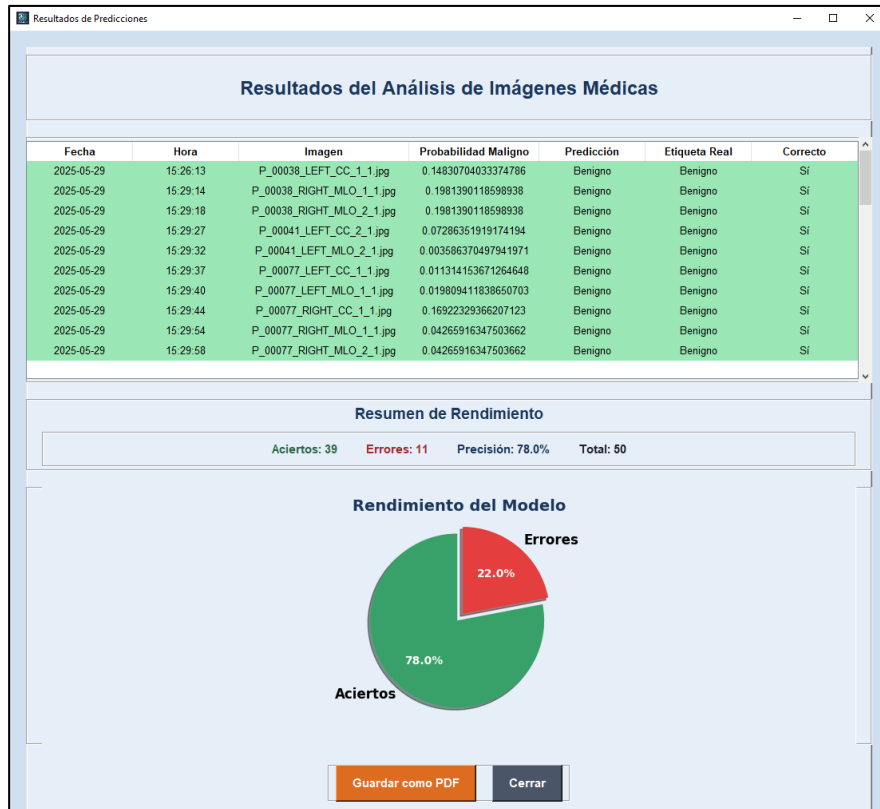


Figura 42: Pantalla de resultados con el historial de predicciones benignas.

Por último, se genera el respectivo archivo PDF para guardar los resultados de las pruebas realizadas con las mamografías benignas.

2025-05-29	15:33:07	P_00352_LEFT_MLO_1_1.jpg	0.050872739404439926	Benigno	Benigno	Si
2025-05-29	15:33:18	P_00353_LEFT_CC_1_1.jpg	0.24445325136184692	Benigno	Benigno	Si
2025-05-29	15:33:24	P_00353_LEFT_MLO_1_1.jpg	0.35135865211486816	Maligno	Benigno	No
2025-05-29	15:33:31	P_00368_LEFT_CC_1_1.jpg	0.0880749300122261	Benigno	Benigno	Si
2025-05-29	15:33:36	P_00368_LEFT_MLO_1_1.jpg	0.2557505667209625	Benigno	Benigno	Si
2025-05-29	15:33:41	P_00402_RIGHT_CC_1_1.jpg	0.004496744368225336	Benigno	Benigno	Si
2025-05-29	15:33:54	P_00402_RIGHT_MLO_1_1.jpg	0.046699561178684235	Benigno	Benigno	Si
2025-05-29	15:34:01	P_00460_LEFT_CC_1_1.jpg	0.003565120277926326	Benigno	Benigno	Si
2025-05-29	15:34:05	P_00460_LEFT_MLO_1_1.jpg	0.030749797821044922	Benigno	Benigno	Si
2025-05-29	15:34:14	P_00495_RIGHT_CC_2_1.jpg	0.012139791622757912	Benigno	Benigno	Si
2025-05-29	15:34:22	P_00495_RIGHT_CC_3_1.jpg	0.012139791622757912	Benigno	Benigno	Si
2025-05-29	15:34:33	P_00495_RIGHT_MLO_1_1.jpg	0.013509010896086693	Benigno	Benigno	Si
2025-05-29	15:34:42	P_00495_RIGHT_MLO_3_1.jpg	0.013509010896086693	Benigno	Benigno	Si
2025-05-29	15:34:49	P_00562_LEFT_CC_1_1.jpg	0.036569323390722275	Benigno	Benigno	Si
2025-05-29	15:34:59	P_00562_LEFT_CC_2_1.jpg	0.036569323390722275	Benigno	Benigno	Si
2025-05-29	15:35:08	P_00562_LEFT_MLO_1_1.jpg	0.06693989038467407	Benigno	Benigno	Si

Resumen de Rendimiento

Categoría	Cantidad	Porcentaje
Aciertos	39	78.0%
Errores	11	22.0%

Figura 43: Reporte PDF con el resumen de la segunda prueba.

En esta segunda prueba, el sistema experimentó el análisis de las mamografías benignas, tuvo un desempeño ligeramente inferior a la prueba con malignos por la diferencia de 1 imagen. No obstante, el desempeño es aceptable, de las 50 imágenes 39 fueron analizadas correctamente y 11 de manera errónea, un rendimiento similar a la prueba anterior. Esto demostraría que la combinación de la técnica de optimización de imagen Corrección de Gamma combinada con la red neuronal VGG19 tiene un balance adecuado para la clasificación de muestras malignas y benignas.

CONCLUSIONES

En este estudio se evaluó el rendimiento de varios modelos de redes neuronales convolucionales, como ResNet50, EfficientNetB0 y VGG19, combinados con técnicas de mejora de imagen (CLAHE, HE, Gamma) y optimizadores (Adam, SGD) con planificadores (StepLR, MultiStepLR). El análisis se basó en matrices de confusión que proporcionaron datos como precisión, sensibilidad, especificidad y tasas de falsos positivos y negativos. Una vez se analizó cada combinación con un total de 36 experimentos, el algoritmo eligió la mejor configuración que posteriormente sirvió para ser evaluada en una aplicación de predicción de imágenes malignas o benignas.

Después de todo este proceso se concluye que:

- El Data Set CBIS DDSM fue adecuado para entrenar y probar los modelos de arquitecturas de redes neuronales. Con base en el enfoque CRISP DM se limpiaron y normalizaron las imágenes, dividiendo el data set en 80% entrenamiento y 20% prueba. Posteriormente se aplicó aumento de datos para obtener como conjunto total 3.568 imágenes y se les aplicó un redimensionamiento de 224 x 224. Esto permitió tener un conjunto de información ideal para comenzar con el entrenamiento de los modelos que por los recursos computacionales se limitaron a 3 épocas, sin embargo, se lograron obtener resultados confiables.
- El uso de diferentes modelos de arquitecturas entrenadas permitió realizar un análisis comparativo más efectivo en el desempeño de redes neuronales convolucionales con enfoque en diagnóstico mamográfico. Los modelos mostraron diferentes comportamientos con las imágenes del Data Set CBIS DDSM, sin embargo, VGG19 con Corrección de Gamma y SDG + MultiStepLR logró alcanzar una precisión global del 71.16% y un AUC de 0.7666, superando al modelo del segundo puesto, EfficientNetB0 con CLAHE, que puntuó 69.32% en precisión global y 0.726 en AUC. Entre ambos modelos hay una diferencia del 1.84% para precisión y 0.04% en AUC, indicando que VGG19 para la clasificación de mamografías es superior (Ver tabla 1).

- La implementación de técnicas de mejora de imagen como CLAHE, HE y Gamma incrementó la precisión de los modelos. Entre estas, la Corrección de Gamma destacó como la técnica más consistente para mejorar el contraste de las imágenes, potenciando así el desempeño comparativo de los modelos. Específicamente, Gamma logró una precisión del 71.16%, en comparación con HE, que alcanzó el 68.46%, mostrando una diferencia del 2.7%. Igualmente, la combinación de Gamma con ResNet50 alcanzó una precisión del 68.60%, superando por 0.44% a VGG19 con CLAHE que obtuvo un puntaje de 68.16% (Ver tabla 1). Esto indica que si se aplica una técnica de optimización gráfica adecuada a las imágenes médicas, la reducción de falsos negativos sería mayor.
- La configuración de redes y técnicas de mejora de imagen más efectiva para la predicción de imágenes malignas o benignas es VGG19 con Corrección de Gamma y el optimizador SGD + MultiStepLR, alcanzando una precisión global del 71.16% y un AUC de 0.766 como se mencionó antes (Ver tabla 1). Esta combinación demostró un buen equilibrio entre especificidad y sensibilidad al ponerla a prueba en la aplicación. Las muestras malignas obtuvieron una precisión del 80% al clasificarse correctamente 40/50 (Ver figura 37), mientras que las benignas obtuvieron un puntaje de 78% con una diferencia del 2% al predecir bien 39/50 (Ver figura 41). Esto demuestra una mayor robustez y capacidad de análisis de esta combinación para la clasificación de mamografías.

RECOMENDACIONES

- Se recomienda extender la cantidad de imágenes del dataset con el fin de hacer más robusto al modelo elegido o probar con arquitecturas más recientes como DenseNet, Vit o ConvNeXt para la captura de patrones más complejos en mamografías y así aumentar el número de épocas de entrenamiento para la observación de convergencias más estables. De igual manera se recomienda el uso de GPUs de Nvidia ya que éstas tienen soporte para herramientas de machine learning como “Pytorch” y se puede trabajar con cada modelo a más épocas para mayor robustez del entrenamiento de cada modelo.
- En futuros desarrollos se sugiere explorar otras técnicas de mejora como Top-Hat filtering, Wavelet Transform o Realce por segmentación adaptativa, que podrían incrementar la vista a las imágenes en lesiones más sutiles. De igual manera, es importante realizar los estudios pertinentes con médicos radiólogos para la validación clínica de las mejoras aplicadas.
- Se recomienda utilizar la base del sistema desarrollado como herramienta para crear una aplicación más fuerte que ya incluya las historias clínicas de los pacientes y así poder entrenar cada modelo de una manera más específica, incluyendo más tipos de lesiones. La capacidad del sistema para evaluar las combinaciones de modelos, optimizadores, schedulers y técnicas gráficas lo convierte en un marco idóneo para ser adaptado a una mayor cantidad de conjuntos de datos médicos con el fin de asegurar reproducibilidad y escalabilidad en la investigación de la IA en salud.

REFERENCIAS

- [1] Q. Sonia, «Organización Panamericana de la Salud,» Globocan, 9 Abril 2024. [En línea]. [Último acceso: 2022].
- [2] J. J. Carballo Pacheco, Clasificación de imágenes médicas con técnicas de "Deep Learning", Grado en Ingeniería Informática en Ingeniería del Software, 2022.
- [3] Y. Castillo Montero, Clasificación de Anomalías en Mamografías mediante, Madrid: Universidad Politécnica de Madrid, 2021.
- [4] J. B. Robles Fajardo, Machine Learning para el Diagnóstico del Cáncer, Bogotá: Universidad Distrital Francisco José de caldas, 2020.
- [5] C. S. K. Lester, Modelo de Machine Learning para la Detección de Tumores Cerebrales, Guayaquil: Universidad Catolica Santiago de Guayaquil , 2022.
- [6] Dr. Federico Aguirre, Dr. Leandro Carballo, Dra. Ximena González, INTELIGENCIA ARTIFICIAL APLICADA A LA IMAGEN MÉDICA, Revista de Imagenología, 2021.
- [7] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R., CRISP-DM 1.0: Step-by-step data mining guide, 2000.
- [8] Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean & Richard Socher , «Deep learning-enabled medical computer vision,» *npj Digital Medicine*, vol. 39, n° 1, pp. 24-32, 2021.
- [9] Singh, Ajit Pal and Saxena, Rahul and Saxena, Suyash and Maurya, Neelesh Kumar, «Artificial Intelligence Revolution in Healthcare: Transforming Diagnosis, Treatment, and Patient Care,» *Asian Journal of Advances in Research*, vol. 7, n° AJOAIR.3556, pp. 241-263, 2024.

- [10] S. N. d. Planificación, *Plan de Desarrollo para el Nuevo Ecuador*, Gráficas IMAGO Cía. Ltda., 2024-2025.
- [11] Nima Tajbakhsh, Suryakanth Gurudu, Jae Y. Shin, R Todd Hurst, «Convolutional Neural Networks for Medical Image Analysis: Fine Tuning or Full Training?,» *IEEE Transactions on Medical Imaging*, 2016.
- [12] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Los Angeles : SAGE, 2009.
- [13] Geert Litjens 1, Thijs Kooi 2, Babak Ehteshami Bejnordi 2, Arnaud Arindra Adiyoso Setio 2, Francesco Ciompi 2, Mohsen Ghafoorian 2, Jeroen A W M van der Laak 2, Bram van Ginneken 2, Clara I Sánchez, «A survey on deep learning in medical image analysis,» *NIH NLM*, vol. 42, pp. 60-88, 2017.
- [14] «Mundo Estudiante,» Servicios Estudiantiles Integrales SL, 2016. [En línea]. Available: <https://www.mundoestudiante.com/metodo-hipotetico-deductivo/>. [Último acceso: 12 11 2024].
- [15] 7wghe, «Método Hipotético-Deductivo: El enfoque científico que revoluciona la resolución de problemas,» BenjaeLobo, 11 11 2023. [En línea]. Available: <https://benjaelobo.es/metodo-hipotetico-deductivo-2/>. [Último acceso: 12 11 2024].
- [16] R. Hernández-Sampieri, *Metodología de la investigación*, México: McGraw-Hill, 2014.
- [17] Shada Omer Khanbari, A. Haider, «Enhanced Mammography image for Breast cancer detection using LC-CLAHE technique,» *Semantic Scholar*, 2022.
- [18] Kyle Strand - Daniela Collaguazo, «Inteligencia artificial: conceptos básicos y aplicaciones en el desarrollo,» Banco Interamericano de Desarrollo, 25 2

2019. [En línea]. Available: <https://blogs.iadb.org/conocimiento-abierto/es/inteligencia-artificial/>. [Último acceso: 12 11 2024].

- [19] Sinno Jialin Pan and Qiang Yang Fellow, «A Survey on Transfer Learning,» *TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, vol. 22, n° 10, pp. 1345-1359, 2010.
- [20] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [21] M. Sánchez Raya, R. Rodríguez Macías y J.M. Andújar Márquez, «SISTEMA PORTABLE DE ADQUISICIÓN DE ADQUISICIÓN DE IMÁGENES DE DIAGNÓSTICO MÉDICO,» *Departamento de Ingeniería Electrónica, de Sistemas Informáticos y Automática*, pp. 393-396, 1998.
- [22] Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, «International evaluation of an AI system for breast cancer screening,» *Nature*, vol. 577, p. 89–94, 2020.
- [23] Babak Ehteshami Bejnordi, Mitko Veta , Paul Johannes van Diest , Bram van Ginneken, Nico Karssemeijer, Geert Litjens, «Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer,» *NIH National Library of Medicine*, vol. 22, pp. 2199-2210, 2017.
- [24] Kaggle, «CBIS-DDSM: Breast Cancer Image Dataset,» Kaggle, 2003. [En línea]. Available: <https://www.kaggle.com/datasets/awsaf49/cbis-ddsm-breast-cancer-image-dataset>. [Último acceso: 27 03 2025].
- [25] Yann LeCun, Yoshua Bengio & Geoffrey Hinton, «Deep learning,» *Nature*, vol. 521, p. 436–444, 2015.

- [26] IEEE, «2015 IEEE nuclear science symposium and medical imaging conference,» de *Transactions on Nuclear Science*, 2015.
- [27] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press, 2016.
- [28] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *DEEP LEARNING*, MA: Cambridge, 2016.
- [29] Kaiming He; Xiangyu Zhang; Shaoqing Ren; Jian Sun, *Deep Residual Learning for Image Recognition*, IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [30] Karen Simonyan, Andrew Zisserman, «Very Deep Convolutional Networks for Large-Scale Image Recognition,» *Arxiv*, vol. 6, 2014.
- [31] Karen Simonyan, Andrew Zisserman, «Very Deep Convolutional Networks for Large-Scale Image Recognition,» *arxiv*, vol. Computer Vision and Pattern Recognition, 2014.
- [32] Mingxing Tan, Quoc V. Le, «EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,» *arXiv, International Conference on Machine Learning, 2019*, vol. 5, pp. 1905-11946, 2019.
- [33] Beau Pollard; Phanindra Tallapragada, «Passive Appendages Improve the Maneuverability of Fishlike Robots,» *IEEE/ASME Transactions on Mechatronics*, vol. 24, n° 4, pp. 1586-1596, 2019.
- [34] Rafael C. Gonzalez • Richard E. Woods, *Digital Image Processing*, Pearson, 2008.
- [35] Stephen M. Pizer, «Adaptive histogram equalization and its variations,» *Science Direct*, vol. 39, n° 3, pp. 355-368, 1987.

- [36] Diederik P. Kingma, Jimmy Ba, «Adam: A Method for Stochastic Optimization,» de *3rd International Conference for Learning Representations*, San Diego, 2015.
- [37] S. Ruder, «An overview of gradient descent optimization algorithms,» *Added derivations of AdaMax and Nadam, arxiv logo*, vol. 2, 2016.
- [38] P. Domingos, «A Few Useful Things to Know About Machine Learning,» *communications of the acm* , vol. 55, n° 10, p. 78–87, 2012.
- [39] Mengtian Li, Ersin Yumer, Deva Ramanan, «Budgeted Training: Rethinking Deep Neural Network Training Under Resource Constraints,» 5 May 2019. [En línea]. Available: <https://arxiv.org/abs/1905.04753>. [Último acceso: 30 Jun 2020].
- [40] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun, «Dermatologist-level classification of skin cancer with deep neural networks,» *Nature*, n° 542, p. 115–118, 2017.
- [41] Deshan Yang, Miao Zhang, Xiao Chang, Yabo Fu, Shi Liu, Harold H. Li, Sasa Mutic, Ye Duan, «A method to detect landmark pairs accurately between intra-patient volumetric medical images,» *American Association of Physicists in Medicine (AAPM)*, vol. 44, n° 12, p. 6547–6558, 2017.
- [42] C. Shearer, «The CRISP-DM Model: The New Blueprint for Data Mining,» *JOURNAL OF DATA WAREHOUSING*, vol. 5, n° 4, pp. 13-22, 2000.
- [43] Li Deng, Dong Yu Authors Info & Claims, «Deep Learning: : Methods and Applications,» *Foundations and Trends in Signal Processing*, vol. 7, n° 3-4, pp. 197 - 387, 2014.

- [44] Yann LeCun, Yoshua Bengio & Geoffrey Hinton , «Deep learning,» *Nature*, p. 436–444 , 2015.
- [45] Xiang Yu, Jian Wang, Qingqi Hong, Raja Teku, «Transfer learning for medical images analyses: A survey,» *Journal of Biomedical and Health Informatics*, vol. vol. 24, n° no. 5, p. 1295–1304, 2020.
- [46] S M A Sharif, Rizwan Ali Naqvi, Mithun Biswas, Woong-Kee Loh, «Deep Perceptual Enhancement for Medical Image Analysis,» *Electrical Engineering and Systems Science* , p. 4826–4836, 2022.
- [47] Yutaro Iwamoto and Kyohei Takeda and Yinhao Li and Akihiko Shiino and Yen-Wei Chen, «Unsupervised MRI Super-Resolution Using Deep External Learning and Guided Residual Dense Network with Multimodal Image Priors,» *Electrical Engineering and Systems Science*, 2022.
- [48] Nada Bouchekout and Abdelkrim Boukabou and Morad Grimes, «Enhanced Secure Transmission of Medical Images through OFDM using Hyperchaotic Systems,» *Computer Science*, 2024.
- [49] «Understanding Epochs in Machine Learning,» DeepAI, 2023. [En línea].
- [50] F. Chollet, *Deep Learning with Python*, Manning Publications, 2018.
- [51] F. (n.d.), «Multimeter usage guide,» 2025.
- [52] Tamara Otzen & Carlos Manterola, «Técnicas de Muestreo sobre una Población a Estudio,» *Int. J. Morphol*, vol. 35, n° 1, pp. 227-232, 2017.
- [53] Goodfellow, I., Bengio, Y., & Courville, A., *Deep Learning*, MIT Press, 2016.

- [54] Wan Azani Mustafal and Mohamed Mydin M. Abdul Kader1, «A Review of Histogram Equalization Techniques in Image Enhancement Application,» *IOP Science home*, vol. 1019, n° 1, p. 126.

ANEXOS

Entrevista dirigida al Dr. Especialista en Diagnóstico Médico

Contexto:

La presente entrevista forma parte del trabajo de titulación “Desarrollo de un algoritmo Basado en Transfer Learning para el análisis de imágenes médicas de Cáncer de Mama”. Esta entrevista sirvió para obtener información sobre las características que debe tener una herramienta de inteligencia artificial para apoyar eficazmente en la práctica clínica.

Objetivo:

La entrevista tiene como objetivo recoger información relevante sobre la utilidad de técnicas de mejora de imagen y modelos de aprendizaje profundo en el contexto médico, así como las necesidades y desafíos que enfrentan los profesionales de la salud en la interpretación de imágenes diagnósticas.

1) ¿Qué características específicas debería tener una herramienta de inteligencia artificial para ser realmente útil en la práctica clínica diaria?

Tiene que ser útil en la práctica clínica diaria ya que debe ser precisa y confiable, capaz de ofrecer diagnósticos con un alto grado de certeza. Su integración debe ser fluida, es decir, debe interactuar sin problemas con los sistemas de gestión de historias clínicas electrónicas y con otros dispositivos utilizados por los médicos en su rutina.

2) ¿Considera que el uso de inteligencia artificial podría reducir los errores humanos en el diagnóstico? Si es así, ¿de qué manera?

No, si bien la IA es una herramienta rápida y útil para determinar características propias o similares que se conozca de los diferentes diagnósticos diferenciales del paciente, basados en sus signos y síntomas, pero existen herramientas y estrategias que se toman mediante exámenes físicos directos al paciente que pueden determinar el aislamiento y brevedad de las sintomatologías del paciente para aproximar en un diagnóstico más fijo.

3) ¿Cree que la implementación de técnicas de mejora de imágenes (como contraste adaptativo o corrección gamma) podría ayudar a mejorar la calidad de las imágenes diagnósticas?

Sí, en la diversidad de enfermedades que se pueden determinar en imágenes se pueden cometer errores en cuanto a la similitud de las patologías, como la aparición de tejidos hiperplásicos que en imagen podrían diferenciarse con quistes, objetos extraños, etc. Además, porque permitirá reducir el uso de pruebas de laboratorio patológicas que se extraen mediante intervención invasiva como biopsias.

4) ¿Qué tipo de imágenes o características considera más difíciles de interpretar (por ejemplo, imágenes de baja resolución, estructuras complejas)?

Las imágenes más difíciles de interpretar son generalmente aquellas de baja resolución o que contienen artefactos que distorsionan la visualización anatómica. Las estructuras complejas, como aquellas con múltiples tejidos superpuestos o áreas con mucha densidad, como el abdomen o el tórax, también son desafiantes.

5) ¿En qué medida influye la calidad de las imágenes (por ejemplo, contraste, nitidez) en su capacidad para hacer un diagnóstico preciso?

La calidad de las imágenes para determinar con claridad las estructuras anatómicas y detectar cambios patológicos pequeños que podrían no ser visibles en imágenes de menor calidad.

6) ¿Qué aspectos éticos o de responsabilidad cree que deben considerarse al implementar inteligencia artificial en el diagnóstico médico?

Por ejemplo, la transparencia: los sistemas de IA deben ser capaces de explicar cómo llegaron a un diagnóstico, para que los médicos puedan confiar en los resultados y asumir la responsabilidad si fuera necesario. Además, es crucial definir claramente quién es responsable en caso de un error, si el médico que utiliza la herramienta, el desarrollador de la IA o ambos

7) ¿Ha trabajado con redes neuronales o ha tenido alguna experiencia directa con sistemas basados en Machine Learning para el diagnóstico? ¿Cuál ha sido su experiencia?

No, aún no he tenido experiencia en utilizar sistemas basados en Machine Learning.

8) ¿Cree que la inteligencia artificial debería actuar como una herramienta complementaria o reemplazar ciertos aspectos del diagnóstico médico?

La inteligencia artificial debe actuar como una herramienta complementaria en el diagnóstico médico, no como un sustituto. La IA puede procesar grandes cantidades de datos rápidamente y proporcionar sugerencias valiosas, pero el juicio clínico humano es insustituible, especialmente en casos complejos donde la interpretación subjetiva y el conocimiento experiencial son clave.

9) ¿Cómo integraría un sistema de inteligencia artificial en su flujo de trabajo clínico? ¿Qué aspectos considera más importantes para su adopción?

Para integrar un sistema de inteligencia artificial en mi flujo de trabajo clínico, sería esencial que esté bien conectado con los sistemas de historias clínicas electrónicas y que funcione de manera rápida y eficiente, sin generar retrasos.

10) ¿Cuáles son los principales obstáculos para la adopción de nuevas tecnologías como la inteligencia artificial en su campo?

Los principales obstáculos incluyen el costo de implementación, la necesidad de infraestructura tecnológica avanzada y la resistencia al cambio por parte de algunos profesionales de la salud y también que la IA sea precisa y confiable, así como las cuestiones legales y éticas sobre la responsabilidad en caso de un error.

11) ¿Cree que los sistemas de Machine Learning y AI podrían ayudar a reducir los tiempos de diagnóstico y mejorar la eficiencia en su especialidad? ¿Por qué?

Sí, creo que los sistemas de Machine Learning y IA pueden ayudar a reducir significativamente los tiempos de diagnóstico y mejorar la eficiencia. Al

automatizar tareas repetitivas y analizar grandes volúmenes de datos rápidamente, permiten a los médicos concentrarse en casos más complejos y tomar decisiones mejor fundamentadas en menos tiempo.